



**Nouveaux Cahiers
de Marge 8 - 2024**
Fiction & données

ISSN : 2607-4427

Éditeur : université Jean Moulin
Lyon 3

Écrire-avec l'intelligence artificielle, ou l'esthétique de la sympoïèse

Expérimentations littéraires avec GPT-2 et GPT-3

*Writing-with artificial intelligence, or the sympoietic
esthetics of sympoiesis. Literary experiments with
GPT-2 and GPT-3*

Erika Fülöp

DOI : [10.35562/marge.956](https://doi.org/10.35562/marge.956)



Creative Commons - Attribution - Pas
d'Utilisation Commerciale - Partage dans les
Mêmes Conditions - CC BY-NC-SA

Écrire-avec l'intelligence artificielle, ou l'esthétique de la sympoïèse

Expérimentations littéraires avec GPT-2 et GPT-3

Writing-with artificial intelligence, or the esthetics of sympoiesis. Literary experiments with GPT-2 and GPT-3

Erika Fülöp

Université Toulouse 2 Jean-Jaurès, membre du Laboratoire
Patrimoine, littérature, histoire et membre associée du
l'Electronic Literature Lab et du laboratoire COSTECH

Résumé : Cet article propose une réflexion sur le potentiel créatif des modèles de langage de grande taille (*LLM*) à partir de deux expérimentations littéraires avec les modèles d'OpenAI : *Pharmako-AI* (2020) de K Allado-McDowell, le premier livre coécrit avec GPT-3, selon l'auteur, et *Internes* (2022) de Grégory Chatonsky, le premier en français, rédigé avec GPT-2 modifié. Les deux interrogent le rapport homme-machine dans le processus d'écriture, ainsi que la transformation de l'écriture dans ce rapport. Plutôt que de chercher à faire disparaître les traces de l'étrangeté que produit la machine – but principal des usages purement instrumentaux –, ces textes l'exploitent dans une démarche philosophique et esthétique. Partant d'une analyse détaillée de la question de l'auctorialité d'une part, et de celle des voix narratives d'autre part, je propose le terme d'« esthétique [*sic*] de la sympoïèse » pour souligner l'enchevêtrement des perspectives éthiques et esthétiques et des agents organiques et synthétiques dans ces processus d'écriture ainsi que dans leurs produits.

Mots clés : modèles de langage de grande taille, LLM, GPT, coécriture, rapport homme-machine, auctorialité cyborg, esthétique, sympoïèse, inquiétante étrangeté

Abstract : This article explores the creative potential of large language models (LLMs) based on two literary experiments with OpenAI's products: K Allado-McDowel's *Pharmako-AI* (2020), the first book co-written with GPT-3 according to the author, and Grégory Chatonsky's *Internes* (2022), the first in French, written with a customized version of GPT-2. Both examine the relationship between man and machine in the writing process, as well as the transformation of writing in this relationship. Rather than seeking to erase the traces of the strangeness produced by the machine – the main aim of purely instrumental uses – these texts work with it in a philosophical and aesthetic perspective. Building on a detailed analysis of the question of auctorality, on the one hand, and of narrative voices on the other, I propose the term “aesthetics [*sic*] of sympoiesis” to underline the entanglement of ethical and aesthetic perspectives as well as the perspectives of organic and synthetic agents in these writing processes and in their products.

Keywords : LLM, GPT, co-writing, human-machine relationship, cyborg authorship, aesthetics, sympoiesis, uncanny

« [R]eality has outpaced the available language to parse what is already at hand¹ »
Benjamin BRATTON et Blaise AGÜERA Y ARCAS

3

« It is an old experience that the traditional, the usual and the hereditary are dear and familiar to most people, and that they incorporate the new and the unusual with mistrust, unease and even hostility.² »
Ernst JENTSCH

« [L]a question serait de nouer un autre rapport à la technique en repensant le nœud originellement formé par l'homme, la technique et le langage »
Bernard STIEGLER

Les livres « coécrits » « avec » une « intelligence » « artificielle »
– plus précisément, un modèle de langage de grande taille ou

-
1. « La réalité a dépassé le langage disponible pour analyser ce qui est déjà à portée de main La réalité a dépassé le langage disponible pour analyser ce qui est déjà à portée de main ». Toute traduction sans autre indication a été réalisée avec l'assistance de DeepL, révisée par l'auteur de cet article.
 2. « C'est une vieille expérience que le traditionnel, l'habituel et l'héréditaire sont chers et familiers à la plupart des gens, et qu'ils intègrent le nouveau et l'inhabituel avec méfiance, malaise et même hostilité. »

LLM³ – se multiplient désormais. La série GPT-n d'OpenAI, et en particulier ChatGPT, optimisé pour un usage conversationnel facile d'accès pour le grand public, a fait couler le plus d'encre (et d'électrons). Ces « outils » semblent désormais très performants, surtout pour écrire des textes courts dans des genres bien codifiés sans être trop spécialisés. Ils exigent « juste » des amorces (*prompts*) bien précises pour produire des textes cohérents et corrects, ce qui explique qu'ils sont devenus objet à la fois d'engouement et d'angoisses. Le potentiel créatif de ces modèles reste cependant plus contesté et cet article s'intéresse justement aux expérimentations artistiques qui explorent ce potentiel.

Il s'agira en particulier de deux textes : *Pharmako-AI* de K Allado-McDowell, coécrit avec GPT-3 en 2020 et publié la même année, et *Internes* de Grégory Chatonsky avec GPT-2 modifié, également écrit en 2020, mais publié en 2022⁴. Chacun des deux affirme d'être un premier : *Pharmako-AI* « the first book to be co-created with the emergent AI⁵ », selon la quatrième de couverture, et *Internes* le premier en français avec une version modifiée de GPT-2⁶. Les deux interrogent le « rapport » homme-machine dans le processus d'écriture et au-delà. Plutôt que de chercher à faire disparaître la trace d'étrangeté que peut laisser la machine sur l'écriture – le but principal des usages purement instrumentaux –, ces textes l'exploitent, dans une démarche philosophique et esthétique qui va au cœur même des enjeux culturels, sociaux et écologiques actuels. Je propose d'appeler cette démarche une esthétique [*sic*] de la sympoièse pour souligner l'enchevêtrement des perspectives éthiques et esthétiques et des agents organiques et synthétiques dans le processus d'écriture, ainsi que dans leurs produits.

4

Comment dire ?

Vous aurez remarqué l'excès de guillemets dans le paragraphe d'ouverture. Il ne s'agit pas d'une lubie ou d'une prudence excessive. Ils indiquent plutôt une difficulté ressentie à chaque mot : comment appeler et comment décrire ce geste et l'écriture qu'il produit ? Et si on reprend les termes à notre disposition, comment

3. *Large language model* : j'utiliserai l'abréviation anglaise courante également dans les textes en français.

4. K Allado-McDowell, *Pharmako-AI*, s.p., Ignota, 2020 ; Grégory Chatonsky, *Internes*, La Rochelle, RRose Éditions, 2022.

5. « [L]e premier livre co-créé avec l'IA émergente ».

6. *1 the Road* de Ross Goodwin (2018) est souvent cité comme le premier livre entièrement produit par une IA, mais il ne s'agissait pas d'un LLM, ni d'une coécriture dialogique telle que l'on trouvera chez nos deux auteurs.

cet usage infléchit et enrichit leur sens ? Nos automatismes linguistiques associent avec les mots courants utilisés pour les gestes d'écriture une intentionnalité qui se veut humaine (tracer des lignes, rédiger un texte, etc.) et un outil inerte (stylo et papier, machine à écrire ou ordinateur) qui peut tout au plus exécuter des commandes bien précises, et qui permettent à l'agent humain de manipuler le texte, mais sans en produire un eux-mêmes. L'idée d'un « rapport » « entre » deux entités, vivantes ou non, suggère également une distinction et une frontière claires entre elles, soulignant leur indépendance plutôt que leur éventuelle interdépendance. À l'opposé, la notion d'hybridité, associée à la figure du cyborg par exemple, suggère l'amalgamation. On a du mal à exprimer simplement un mode d'être-avec et de faire-avec qui se situerait entre ces deux pôles. Cet article part d'un besoin double de penser et de dire (écrire) comment le geste d'écrire change « -avec » les *LLM*, et l'enchevêtrement de cette évolution avec celle du « rapport » homme-machine.

La question de la langue pour parler des processus mêmes qui chamboulent nos concepts et pratiques culturels les plus fondamentaux est loin d'être anodine. Les enjeux sont d'abord linguistiques, et sociaux et culturels par le biais du langage. Le langage est cependant, en même temps, l'outil qui nous permet de les penser – et qui peut aussi bien faire écran ou nous aveugler. La question linguistique sous-tend ainsi dans toute sa complexité les trois espaces-temps intimement liés dans lesquels la question homme-machine se pose par les écrits qui nous intéressent : celui du processus d'écriture, celui de l'univers écrit, et celui de la lecture. Pour explorer ces trois espaces contigus, après une brève présentation des deux œuvres je propose de soulever les questions de savoir qui écrit et qui parle, tout en prenant en compte, dans la mesure du possible, le processus de lecture et les réactions que les textes ont suscitées dans mon cas, inévitablement subjectif.

5

Les textes

Pharmako-AI

K Allado-McDowell est un·e « auteur·e, conférencier·ère et musicien·ne⁷ », initiateur·rice du programme « Artists + Machine Intelligence » de Google AI. Iel a publié plusieurs ouvrages

7. K Allado-McDowell, « About », *K Allado-McDowell* [en ligne], URL : <https://www.kalladomcdowell.com/> [consulté le 8 avril 2024]. Allado-McDowell se désigne par le pronom personnel non-binaire « they », que j'essaie de maintenir en français en me servant du pronom « iel ».

coécrits avec GPT-3, dont *Pharmako-AI* était le premier. C'est une collection d'« essais, récits et poèmes », « écrits avec » GPT-3 selon sa propre expression, sous forme d'une « conversation expérimentale » commencée pendant la pandémie de covid-19 en 2020. Selon la quatrième de couverture,

«*The first book to be co-created with the emergent AI, Pharmako-AI is a hallucinatory journey into selfhood, ecology and intelligence via cyberpunk, ancestry and biosemiotics. Through a writing process akin to musical improvisation, Allado-McDowell and GPT-3 together offer a fractal poetics of AI and a glimpse into the future of literature*».

L'auteur·e souligne également dans l'introduction l'intention littéraire et subversive de son projet : « This book is intended as a hybrid disruption: a literary intervention that rails against stale, conservative ideas around how we make books⁹. »

Le processus d'écriture a consisté en un processus itératif de dialogues et de sélection de passages dans les réponses de l'IA¹⁰. Allado-McDowell laisse les traces de sa contribution visibles : ses amorces sont en gras, bien distinguées des textes produits par la machine¹¹. Chacun des seize chapitres représente une séance d'écriture. GPT-3 réagit en fonction de ce qui précède dans la même conversation, mais cette mémoire ne se transmet d'une séance à l'autre que par la mémoire humaine. Plutôt que d'un récit ou argument linéaire, le texte est ainsi un assemblage de réflexions sur des sujets variés, mais connectés tels que, selon les titres des chapitres, « Hyperspatial art », « How to build a different universe », « The language of plants », « Post-cyberpunk », « Tearing down the foundations of the self », « Mercurial oracle », « The poison path », « Generative poetics theory » ou « AI ethics¹² ».

8. « Le premier livre co-créé avec l'IA émergente, *Pharmako-AI* est un voyage hallucinatoire dans l'identité, l'écologie et l'intelligence par le biais du cyberpunk, de la généalogie et de la biosémiotique. Grâce à un processus d'écriture proche de l'improvisation musicale, Allado-McDowell et GPT-3 offrent *ensemble* une poétique fractale de l'IA et un aperçu de l'avenir de la littérature. » Je souligne.

9. K Allado-McDowell, *Pharmako-AI*, London, Ignota Books, 2020, p. 7.

10. *Ibid.* p. 11

11. Je garde cette distinction dans les citations, sans remettre en question l'affirmation de l'auteur·e à leur propos.

12. « Art hyperspatial », « Comment construire un univers différent », « Le langage des plantes », « Post-cyberpunk », « Démolir les fondations du moi », « L'oracle mercuriel », « Le chemin du poison », « Théorie de la poétique générative », « Éthique de l'IA ».

Les amorces, de longueurs variables, tout comme les passages générés, incluent des récits autobiographiques aussi bien que des envolées poétiques et des observations anthropologiques. GPT-3 poursuit les amorces d'une façon plus ou moins surprenante et propose parfois des termes d'une façon inattendue, tel que « meglanguage », « a technique for a type of synaesthetic communication », qui « creates hyperstructure, layers of parallel worlds, and the possibility of hypertime¹³ ». Le livre dans son ensemble présente cependant un univers cohérent. La pensée qui se dessine pourrait être qualifiée de New Age par son accent sur la spiritualité et la valeur des états de conscience alternatifs obtenus sous l'effet de la méditation, de la prière et de substances hallucinogènes. Elle renoue par là avec la contre-culture américaine des années 1960 qui a joué un rôle clé dans l'histoire de l'informatique¹⁴, tout en le prolongeant avec des préoccupations écologiques et culturelles de nos jours.

Internes

L'artiste franco-canadien Grégory Chatonsky travaille avec ce qu'il appelle « l'imagination artificielle » et la génération de textes, d'images et de musique depuis 2008. *Internes* est son premier roman, qui reprend cependant en partie une démarche de génération de texte entamée pour son installation. *Terre seconde*, présentée au Palais de Tokyo en 2019. *Internes* est aussi le titre d'une installation créée en 2021 avec Goliath Dyèvre, « un ensemble de sculptures qui pourrait recouvrir la Terre en totalité et constituer une seconde peau planétaire¹⁵ ». Le titre de la troisième partie du livre, « Externes », provient également d'une création de 2020 du même titre, qui proposait de « créer une nouvelle relation entre la matérialité et le numérique, deux mondes dans lesquels nous ne cessons de vivre mais que nous avons encore du mal à métaboliser¹⁶ ». *Internes* fait donc partie d'un univers de pensée qui se déploie dans de multiples dimensions dans l'œuvre de Chatonsky.

13. « Une technique de communication synesthésique [qui] crée une hyperstructure, des couches de mondes parallèles et la possibilité d'un hypertemps », K Allado-McDowell, *Pharmako-AI*, op. cit., p. 71.

14. Fred Turner, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, Chicago, University of Chicago Press, 2006.

15. Grégory Chatonsky, « Internes – Pris MAIF », *Gregory Chatonsky* [en ligne], 2021, URL : <https://chatonsky.net/internes/> [consulté le 8 avril 2024].

16. Grégory Chatonsky, « Externes », *Gregory Chatonsky* [en ligne], 2021, URL : <http://chatonsky.net/externes/> [consulté le 8 avril 2024].

Entamant son travail sur *Internes* au printemps 2020, peu avant la sortie de GPT-3, Chatonsky utilise le modèle précédent, GPT-2, plus petit, mais plus ouvert et modulable grâce à son code encore *open source*¹⁷. Pour chacune des trois parties du livre, il a opéré un apprentissage supplémentaire sur un corpus qu'il a constitué lui-même : « pour la première partie, de monologues de Beckett, Pessoa, Fisher ou Guyotat ; pour la deuxième, d'articles techniques expliquant l'espace latent, provenant du site d'archives ouvertes Arxiv ; la troisième, des sciences naturelles¹⁸ ». L'auteur explique ainsi son processus d'écriture :

8

« Mi-amusé mi-goguenard, je me mis au travail, traçant vaguement l'amorce d'une fiction : un humain est en train de mourir, se souvenant de son existence passée, il est hanté par ses existences possibles au moment où l'espèce humaine s'éteint et, avec elle, la possibilité même de témoigner. Pour cette première partie, j'alimentai un réseau de neurones nommé GPT-2 de centaines d'ouvrages consistant en des monologues intérieurs. Je commençai à écrire une phrase puis, lassé par ma prétendue intériorité, je laissai le logiciel proposer une suite, cliquant plusieurs fois jusqu'à ce que sa proposition me convienne et m'évoque un prolongement. Je poursuivis, inévitablement influencé par ces générations successives qui tordaient progressivement le fil de la narration et sa structure même jusqu'à ce que, au cours des phrases complétées, le plan d'ensemble soit bouleversé et m'entraîne sur des chemins imprévus¹⁹. »

Dans ce récit à la première personne du singulier, on glisse d'une conscience qui semble humaine vers un langage qui se dit sans extériorité, sans accès au dehors. Chapitre 2, « Deep Box » se termine avec cet être aux contours flous étant « mort à nouveau emporté hors de cette planète par le dernier météore ». Chapitre 3 commence ensuite par une série de constats qui introduit également une distinction typographique, sans toutefois guider le lecteur sur sa signification :

17. À retrouver en ligne sur : <https://github.com/openai/gpt-2> [consulté le 8 avril 2024]. OpenAI mit fin à sa politique d'ouverture avec GPT-3.

18. « Grégory Chatonsky : Captures de nos imaginaires. Conversation avec Aurélie Cavanna », *ArtPress*, n° 505, 2022, p. 64 et s, URL : <http://chatonsky.net/artpress50/> [consulté le 8 avril 2024].

19. Grégory Chatonsky, « Internes ou le roman (de l')artificiel » [en ligne], *L'art même*, n° 89, 2023, p. 21-23, URL : <http://chatonsky.net/internes-am89/> [consulté le 8 avril 2024].

« **Il est temps de partir.**

Il est temps de changer de soi.

Il est temps de réveiller les Dieux.

Je me suis mis à nager. (135) »

« Ça » parle d'une sortie de la boîte noire, et/ou de la sortie de la boîte noire de l'orbite terrestre après l'épuisement de toutes les ressources sur Terre et la fin de l'humanité, avec des fragments d'un récit de cette fin. Temps, espaces et dimensions se confondent dans ce texte foisonnant qui se laisse difficilement résumer. Dans l'ensemble, on perçoit une évolution vers une prise de distance, une dissolution ou un dépassement graduel par rapport à l'humain, non pas dans un esprit transhumaniste solutionniste – que Chatonsky dénonce – mais dans un mouvement vers un dehors, un « ailleurs » qui permet (ou en tout cas, essaie) un regard sur ce que l'on est.

Deux projets, deux méthodes et deux styles très différents donc, mais reliés par le fait d'impliquer un modèle de GPT dans l'écriture de textes à la première personne et par les interrogations qu'ils soulèvent.

Qui écrit ?

9

GPT et *écrire avec*

Un transformeur génératif préentraîné, ou GPT (*Generative Pre-trained Transformer*), est un outil de génération de texte qui s'appuie sur un apprentissage profond (*deep learning*) effectué sur une masse de textes récoltés sur le web, écrits (pour la plupart) par des personnes humaines. Le corpus inclut des documents dans plusieurs langues sans les distinguer d'un point de vue linguistique, c'est l'apprentissage qui permet au modèle d'identifier des schémas récurrents (*patterns*) dans les séquences et de produire ensuite des textes dans la langue de la requête. La qualité de la réponse reste néanmoins inégale dans le cas des langues moins représentées dans le corpus, largement dominé par l'anglais : GPT-2 a été entraîné sur le jeu de données WebText créé par OpenAI à partir du contenu de sites identifiés par le biais de 45 millions de liens sortant de Reddit²⁰. Or, quelque 97 % des discussions sur la plateforme seraient en anglais²¹, et

20. Alec Radford *et al.*, « Language Models are Unsupervised Multitask Learners », *Semantic Scholars* [en ligne], 2019, URL : <https://api.semanticscholar.org/CorpusID:160025533> [consulté le 8 avril 2024].

21. Felipe Hoffa, « The Most Popular Languages on Reddit, Analyzed with Snowflake and a Java UDTF », *Medium* [en ligne], 29 octobre 2021, <https://>

on peut supposer une proportion similaire (sinon encore plus déséquilibrée) pour les sites référencés. GPT-3 s'appuie, quant à lui, sur un corpus plus large et plus raffiné : il inclut les archives de 2016 à 2019 de Common Crawl, une collecte mensuelle du web, ainsi qu'une version augmentée de WebText, deux collections de livres de Google, et Wikipédia en langue anglaise²². 92,65 % du corpus global est en anglais et le français est en deuxième position avec seulement 1,82 %, suivi par l'allemand²³ (1,45 %). Si ces corpus représentent donc une partie considérable des savoirs et communications écrits de l'humanité au XXI^e siècle, ils sont loin de contenir « tout l'Internet » et comportent inévitablement de nombreux biais historiques, linguistiques, géographiques, culturels et sociaux qui reproduisent ceux du web et en renforcent certains²⁴. Des inquiétudes ont également été soulevées concernant l'utilisation des documents récoltés sans aucun égard aux droits de leurs auteurs²⁵.

Les données linguistiques sont sélectionnées, collectées, puis filtrées et nettoyées pour éviter les doublons, et pondérées pour donner aux documents « de qualité », tels que les livres, davantage de poids dans l'ensemble. Les textes sont ensuite analysés et segmentés en y identifiant des unités récurrentes (*tokens*). Ces unités, qui ne correspondent pas forcément à des mots ou à d'autres unités linguistiques significatives même en anglais²⁶, sont représentées par des vecteurs numériques (*embedding*), une

10

towardsdatascience.com/the-most-popular-languages-on-reddit-analyzed-with-snowflake-and-a-java-udtf-4e58c8ba473c [consulté le 8 avril 2024]. Cette représentation des langues sur Reddit n'a probablement pas radicalement changé.

22. Tom Brown *et al.*, « Language Models are Few-Shot Learners », *Advances in Neural Information Processing Systems* 33 (2020), arXiv:2005.14165v4.

23. Statistiques par nombre de mots, voir https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv [consulté le 8 avril 2024].

24. L'article de Brown *et al.*, « Language Models are Few-Shot Learners », *op. cit.* souligne déjà plusieurs de ces biais.

25. Ella Creamer, « Authors file a lawsuit against OpenAI for unlawfully “ingesting” their books », *The Guardian* [en ligne], 5 juillet 2023, <https://www.theguardian.com/books/2023/jul/05/authors-file-a-lawsuit-against-openai-for-unlawfully-ingesting-their-books> [consulté le 8 avril 2024].

26. La série GPT-n utilise la méthode de « byte pair encoding », encodage par paires d'octets, qui commence par des *tokens* d'un seul caractère et les élargit graduellement en cherchant des combinaisons récurrentes de 2, 3, etc. caractères jusqu'à atteindre le nombre de *tokens* souhaités (40 000 dans le cas de GPT-2) : Alec Radford *et al.*, *Language Models are Unsupervised Multitask Learners*, 2019, p. 4.. Cet outil d'OpenAI permet de visualiser le

série de chiffres qui situe chaque *token* dans un espace vectoriel multidimensionnel. L'apprentissage du modèle consiste à effectuer des calculs qui conjuguent ces vecteurs initiaux en fonction de leur position dans des séquences et leur importance relative par rapport aux autres *tokens* de la même séquence (*contextual embedding*) et à essayer de prédire le *token* suivant de la séquence donnée. Le réseau de neurones – d'unités de calcul – permet d'effectuer un grand nombre de calculs en parallèle pour chacun des *tokens* d'une séquence et de raffiner leur représentation vectorielle en passant par plusieurs couches. L'apprentissage optimise en même temps les paramètres des calculs pour une meilleure prédiction par un ajustement graduel en passant par plusieurs cycles. Le résultat de cet apprentissage qui s'effectue sans intervention humaine après l'initiation (*unsupervised pre-training*) est un espace vectoriel multidimensionnel d'une immense complexité qui traduit les schémas linguistiques détectés dans le corpus uniquement par le biais des rapports de proximité et de co-occurrence entre les *tokens*. C'est cet espace vectoriel que l'on appelle également l'espace latent : il est interne au modèle et sa représentation non seulement visuelle, mais même mentale, ne peut être qu'approximative²⁷. Lorsqu'on présente un modèle avec une requête (*prompt*), il analyse la tâche et propose une réponse calculée *token* par *token* à base des probabilités dans le contexte donné à la lumière des paramètres établis lors de l'apprentissage. Des apprentissages plus ciblés (*fine-tuning*) peuvent compléter les modèles de fondation pour mieux les adapter à des tâches précises²⁸ – c'est ce que Grégory Chatonsky a fait par le biais d'un corpus choisi.

Quand on parle de ce que produit « l'intelligence artificielle », il faut se rappeler cette complexité dans laquelle la pensée et l'imaginaire humains – la conception des algorithmes, de l'architecture du modèle et de l'immense corpus qui permettent

découpage en token d'un passage : <https://platform.openai.com/tokenizer> [consulté le 8 avril 2024].

27. Les nombre de paramètres souvent cités pour caractériser les modèles donne une idée de l'ordre de leur complexité : il s'agit de 1,5 milliard de paramètres extraits de 40Go de texte pour GPT-2, et de 175 milliards de paramètres extraits de 570Go de texte pour GPT-3. Pour une représentation visuelle inévitablement simplifiée de GPT-2, voir Rebecca Kehlbeck *et al.*, *Demystifying the Embedding Space of Language Models* [en ligne], 2021, URL : <https://bert-vs-gpt2.dbvis.de> [consulté le 8 avril 2024].

28. Si l'on sait que ChatGPT a été optimisé pour le mode conversationnel, il n'est pas clair si les différents modèles de fondation, tels qu'ils sont rendus accessibles sur OpenAI Playground, ont été optimisés par un apprentissage ciblé et si oui, comment.

d'obtenir ce degré de raffinement – sont déjà inextricablement enchevêtrés avec les « automates » qui exécutent sans intentionnalité propre – qui sont donc également le produit de pensée humaine, elle-même de plus en plus assistée par les machines. Si l'intelligence, ce « concept frontière entre vie biologique et vie symbolique²⁹ », se définit aujourd'hui communément par une capacité de compréhension, d'adaptation et de réponse à un environnement, celle des *LLM* reste cependant purement linguistique et statistique : elle provient de données *écrites*, et ne « répond » et ne « s'adapte » à son « environnement », que sous forme de langage³⁰. Un *LLM* ne traite que la « surface » des langues et n'a aucune notion de référent : tout ce qu'il « sait » est basé sur la déconstruction et l'analyse des textes en tant que forme, dans laquelle il observe des récurrences et régularités quantifiables. Cela qui lui permet néanmoins de produire des propos sémantiquement cohérents, mais pas forcément factuellement corrects.

12

L'interface Playground d'OpenAI, qui permet l'utilisation des *LLM* de l'entreprise sans accès au code, invite à soumettre des amorces (*prompts*) sous forme de questions ou de textes à continuer³¹. L'utilisateur peut également modifier quelques paramètres : la « température » permet de monter ou baisser le degré d'imprévisibilité (*randomness*) et l'application de la pénalité de fréquence aide à éviter la répétition, tandis que la pénalité de présence incite le modèle à évoluer vers de nouveaux sujets. Allado-McDowell a dû utiliser la version bêta privée de GPT-3 lorsqu'elle a été rendue disponible à des utilisateurs choisis en juillet 2020 pour tester, mais iel ne communique pas les détails des réglages appliqués. Travaillant plus directement avec le code, Chatonsky a pu non seulement paramétrer l'imprévisibilité, mais aussi moduler le style et l'orientation du texte généré en

29. Catherine Malabou, *Métamorphoses de l'intelligence : que faire de leur cerveau bleu ?*, Paris, PUF, 2017, p. 10.

30. La grande question est de savoir jusqu'où peut aller le langage et une forme de cognition qui passe uniquement par lui : peut-il donner lieu à une forme de connaissance et de conscience qui le dépasse ? Yann LeCun maintient qu'un apprentissage sur des vidéos pourra permettre à une IA d'apprendre les lois physiques du monde et d'ouvrir vers une intelligence plus complète du monde. Voir Yann LeCun, « From Machine Learning to Autonomous Intelligence » [en ligne], Institute for Experiential AI, 25 mai 2023, URL : <https://www.youtube.com/watch?v=mViTAXCg1xQ> [consulté le 8 avril 2024].

31. Le mode conversationnel qui est devenu très populaire a été introduit par ChatGPT en novembre 2022. Cette application exploite un modèle réduit de GPT-3 optimisé pour ce type d'interaction.

définissant le poids de l'apprentissage effectué sur son corpus dans la prédiction. Allado-McDowell prend donc un dispositif d'écriture entièrement prédéfini par un tiers, tandis que Chatonsky intervient dans le dispositif d'écriture même.

Écrire avec GPT par le biais de l'interface prévue à cet effet dépasse déjà la façon dont on conçoit traditionnellement le geste d'écrire, où l'outil de production, qu'il s'agisse de stylo, de machine à écrire ou d'ordinateur, ne *produit* pas de texte, n'ajoute pas de mots à celui de l'auteur. Le projet plus complexe de Chatonsky va cependant plus loin encore en faisant exploser les frontières du processus d'écriture et renouant en même temps avec la question littéraire.

Mais l'écriture n'a-t-elle pas toujours dépassé les frontières du geste manuel de la matérialisation graphique des mots sur un support visuel ? Dans toute écriture, il y a toujours déjà gestion et digestion d'un corpus qui constitue un « espace latent » dont émerge le nouveau texte et dans lequel il s'inscrit. On n'écrit jamais dans un espace vide et on ne fait jamais que tracer des mots pour écrire. Est-ce que le dispositif technique, le code comme écriture d'un processus plutôt que juste processus d'écriture, et le calcul statistique pour produire un objet sémiotique changent fondamentalement la logique de l'écriture, et si oui, en quoi ?

Des réponses diamétralement opposées semblent possibles à cette question. D'une part, il n'y a pas de changement fondamental, puisque l'écriture a toujours déjà été un geste et une technique. Katie Chenoweth cite Derrida, pour qui l'écriture est inscrite dans la langue même par le biais de l'articulation, pour montrer que la langue (*tongue*, l'organe) est déjà « un instrument d'écriture³² ». L'IA ne fait qu'amplifier et externaliser, et donc rendre visible, ce fond technique à ce que l'on croyait organique. Cette technique est « l'*impensé* », comme le dit Stiegler³³, mais qui s'impose désormais à un niveau où il devient difficile de l'ignorer. Les travaux d'Emmanuel Souchier et d'Yves Jeanneret ont montré dès la fin des années 1990 comment « l'architecte » numérique participe du processus d'écriture et façonne le

32. « [L]articulation et par conséquent l'espace de l'écriture opèrent à l'origine du langage. » Jacques Derrida, *De la Grammatologie*, Paris, Éditions de Minuit, 1967, p. 326, et « Following this generalized sense of writing that Derrida invites us to consider, the tongue would thus also be a *writing instrument*. » Katie Chenoweth, *The Prosthetic Tongue: Printing Technology and the Rise of the French Language*, Philadelphia, University of Pennsylvania Press, 2019, p. 40 (l'auteure souligne).

33. Bernard Stiegler, *La technique et le temps, suivi de Le nouveau conflit des facultés et des fonctions dans l'anthropocène*, Paris, Fayard, 2018, p. 17.

résultat³⁴. Le logiciel de traitement de texte avec le copier-coller, la correction automatique, la complétion automatique des textes, le résumé automatisé, et maintenant l'écriture de passages ou même de textes intégraux suivant un modèle ou une consigne – la montée en puissance de l'« assistance » numérique à l'écriture a été graduelle, mais rapide. L'arrivée des LLM producteurs de texte semble toutefois représenter un saut qualitatif pour ce qui concerne le rôle du dispositif dans le processus d'écriture. Ce n'est plus seulement son produit qui est externalisé, mais le processus aussi³⁵. L'acteur humain reste (pour le moment) la source de l'intentionnalité à l'origine des textes, mais l'intention se situe à des niveaux de plus en plus éloignés de la production même du texte.

Les pratiques historiques de l'écriture qui ont formé notre pensée suggèrent qu'écrire *avec* a deux dimensions : d'une part, on écrit avec des outils, matériels et inanimés, et d'autre part, on peut écrire avec des collaborateurs, humains pourvus d'intentionnalité et de pensée propres. L'IA représente un entre-deux, ou plutôt un *autre* qu'on a encore du mal à situer et à nommer, puisque notre langage porte une vision qui suppose une distinction qui va de soi entre le sujet qui pense (et qui écrit, l'auteur, un acteur conscient de son action) et l'objet qu'il utilise (un agent technique).

14

L'auteur qui n'(en) est pas *un*

L'auteur en droit

D'un point de vue légal, cet entre-deux se manifeste aux États-Unis notamment par le fait que l'auteur humain ne peut revendiquer le droit d'auteur d'un document « containing AI-generated material³⁶ » que dans la mesure où il peut montrer qu'il y a « sufficient human authorship » par les « traditional elements

34. Emmanuel Souchier et Yves Jeanneret, « Pour une poétique de “l'écrit d'écran” », *Xoana*, n° 6/7, 1999, p. 97-107.

35. La génération automatique de textes existe depuis les premiers ordinateurs, mais avant l'explosion des données disponibles grâce à Internet, des capacités de calcul et de l'apprentissage automatique, elle a été basée sur une logique combinatoire qui exigeait la définition d'une grammaire et d'un lexique, avec une flexibilité linguistique bien plus limitée et sans « intelligence » permettant de réagir à des questions ou amorces non prédéfinies.

36. United States Copyright Office, « Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence », *Federal Register* [en ligne], 16 mars 2023, URL : <https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence> [consulté le 8 avril 2024].

of authorship », c'est-à-dire qu'il « exercise ultimate creative control » et « determines the expressive elements » du produit³⁷. Les directives reconnaissent que la distinction est loin d'être évidente et précisent que la question ne peut être décidée qu'au cas par cas, à la lumière du processus de création. Selon cette approche, l'auctorialité n'est plus une question binaire (« X est-iel l'auteur·e de Y ? »), mais de nature et de degré (« Comment Y a-t-il été fait et dans quelle mesure X en est-iel l'auteur·e ? »).

En France, où en juin 2023 il n'existe pas encore de législation sur ce point³⁸, un document publié sur le site du ministère de la Culture souligne que « [l]e droit d'auteur a en effet été conçu pour la protection des auteurs, personnes physiques³⁹ ». Au-delà de protéger les auteurs, par définition humains donc⁴⁰, il s'agit de protéger le concept d'auteur : « la protection de ces créations par le droit d'auteur viendrait soit forcer nos catégories actuelles au risque de les déformer, soit les étendre au risque de les dénaturer⁴¹ ». Cette catégorie d'auteur se repose sur l'exercice « d'un réel effort de création donnant naissance à une œuvre originale de nature à refléter la personnalité de son auteur⁴² ».

Création, originalité, reflet de personnalité – comme le note Espen Aarseth, le concept d'auteur même implique une vision idéologique de la littérature⁴³, et le discours légal en témoigne

37. ...d'un document qui « contient des matériaux générés par une IA » que dans la mesure où il peut montrer qu'il y a « un degré suffisant d'auctorialité humaine » par les « éléments traditionnels de l'auctorialité », c'est-à-dire qu'il « exerce un contrôle créatif ultime » et « détermine les éléments expressifs ». *Loc. cit.*

38. Caroline Casalonga et Juliette Dionisi, avocates spécialistes du sujet, anticipent qu'elle sera similaire à l'approche américaine. Voir Caroline Casalonga et Juliette Dionisi, « Une intelligence artificielle peut-elle être un auteur ? », *La Tribune* [en ligne], 12 avril 2023, URL : <https://www.latribune.fr/opinions/tribunes/une-intelligence-artificielle-peut-elle-un-auteur-958292.html> [consulté le 8 avril 2024].

39. Georgie Courtois, Jean-Sébastien Mariez, Jeanne Roussel, « Intelligence artificielle et droit d'auteur », *Ministère de la culture* [en ligne], (sans date), https://www.culture.gouv.fr/content/download/263325/file/Annexe_4-DGFLA.pdf [consulté le 8 avril 2024] (je souligne).

40. INSEE, « Personne physique », *INSEE* [en ligne], URL : <https://www.insee.fr/fr/metadonnees/definition/c1558#:~:text=Au%20sens%20du%20droit%20fran%C3%A7ais,te%2C%20de%20la%20personnalit%C3%A9%20juridique> [consulté le 8 avril 2024] (je souligne).

41. *Loc. cit.* (je souligne).

42. *Loc. cit.* (je souligne).

43. « [T]he concept of author entails a certain ideological view of literature ». Espen J. Aarseth, *Cybertext: Perspectives on Ergodic Literature*, Baltimore,

éloquemment : nous sommes ici dans la conception humaniste et romantique de l'auteur et de la littérature. Par contraste, au Royaume-Uni, le Copyright, Designs and Patents Act reconnaît, depuis 1988, en tant qu'auteur la personne responsable des « agencements » nécessaires pour la création d'une œuvre générée par l'ordinateur :

*In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken*⁴⁴.

Une consultation en cours sur une mise à jour suit cette même logique, prévoyant que

*work [...] made with assistance from AI but involving human creativity [...] will be protected like any other work. Copyright protection will protect the work to the extent that it is the human creator's "own intellectual creation", and the first owner of the work will be that creator. The AI in these cases may be considered to simply act as a tool which allows an artist to express their creativity.*⁴⁵

16

L'auteur se retrouve dans une situation paradoxale où l'intérêt artistique de l'IA – son potentiel créatif propre – est inversement

Md., Johns Hopkins University Press, 1997, p. 134 (je souligne). Il précise que pour sa part, « I use the term author here only as a label for the positions in a communication system in which the physical text is assembled, without any regard for the social or cognitive forces active in the process. »

44. « Dans le cas d'une œuvre littéraire, dramatique, musicale ou artistique créée par ordinateur, l'auteur est la personne qui prend les dispositions nécessaires à la création de l'œuvre. », « Copyright, Designs and Patents Act 1988 », *The National Archives* [en ligne], URL : <https://www.legislation.gov.uk/ukpga/1988/48/section/9> [consulté le 8 avril 2024] (je souligne).

45. « toute œuvre réalisée avec l'aide de l'IA mais faisant appel à la créativité humaine [...] sera protégée comme n'importe quelle autre œuvre. La protection du droit d'auteur protégera l'œuvre dans la mesure où il s'agit de la « propre création intellectuelle » du créateur humain, et le premier propriétaire de l'œuvre sera ce créateur. Dans ces cas, l'IA peut être considérée comme un simple outil permettant à l'artiste d'exprimer sa créativité. » Intellectual Property Office, « Artificial Intelligence Call for Views: Copyright and Related Rights », *Gov.uk* [en ligne], 23 mars 2021, URL : <https://www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights> [consulté le 8 avril 2024] (je souligne).

proportionnel à la reconnaissance (légale) qu'il peut espérer en l'exploitant.

Pour nos deux ouvrages, les auteurs *humains*⁴⁶ affichent leur seul nom en couverture, tout en précisant en quatrième de couverture qu'ils n'ont pas été « seuls » à écrire. Il est vrai que l'autre « main » n'a pas vraiment de nom à afficher, ou le nom qui pourrait être affiché n'aurait pas le même statut que celui des auteurs *humains*. Allado-McDowell fait cependant remarquer une autre raison et conséquence de cet affichage de son seul nom : c'est iel qui a fait le choix des textes, qui décide de publier ce texte sous cette forme et qui porte la responsabilité de tout ce qui est dit⁴⁷. L'envers de la médaille du droit d'auteur, c'est cette responsabilité. Attribuer le texte en partie à une autre instance impliquerait aussi se décharger en partie de cette responsabilité. Or, l'IA n'aurait pas produit ces textes sans les amorces données par les auteurs. Katherine Hayles souligne que les systèmes technologiques agissent désormais comme des acteurs dans des situations avec des conséquences éthiques et morales et cite l'observation de Peter-Paul Verbeek que dans les réseaux où acteurs humains et techniques sont enchevêtrés, la responsabilité est également partagée, puisque les décisions sont prises par des associations humain-machine. Le problème, remarque également Hayles, c'est que nos théories éthiques sont « intensément anthropocentrées ». Elle invite donc à travailler sur un nouveau cadre de pensée qui permette d'explorer comment les technologies transforment les manières dont les décisions morales et éthiques sont prises et formulées⁴⁸.

Si pour le moment, l'auteur *humain* (et son éditeur) assume seul la responsabilité de nos deux textes et leur publication, une

46. Je me demande s'il faut préciser « humain » à chaque fois que je dis « auteur » dans ce contexte et que je parle des personnes. Le préciser implique que je considère le *LLM* auteur aussi ; sinon, c'est qu'il ne l'est pas... Or, tout l'argument est ici, qu'il s'agisse d'un entre-deux dans ce genre d'écriture, ou plutôt d'une instance *autre* pour laquelle on n'a pas (encore) de nom, que le mot « auteur » ne décrit pas bien – à moins que, comme dit le texte français sur le droit d'auteur, on n'élargisse le concept et accepte qu'il perde sa spécificité. Mais il serait mieux de pouvoir à la fois maintenir l'historicité et la spécificité de ce concept pour pouvoir montrer l'évolution des pratiques et des outils, et l'ouvrir pour inclure les nouveaux modes d'auctorialité.

47. London College of Fashion, *HYPERHYPERBODIES Summit: K Allado-McDowell – Latent Space is The Place* [en ligne], 2023, URL : <https://www.youtube.com/watch?v=iIe3jLZ00Dc> [consulté le 8 avril 2024].

48. N. Katherine Hayles, *Unthought: The Power of the Cognitive Nonconscious*, Chicago, The University of Chicago Press, 2017, p. 36-38.

petite brèche s'est déjà ouverte entre les deux fonctions principales de l'auteur : produire le texte d'une part, et le porter d'autre part. La législation en vigueur n'est bien sûr qu'une tentative de cristallisation forcément partielle d'une réflexion en cours dans la théorie et la pratique actuelles. La multiplicité des termes et des approches dans ces dernières témoigne de l'état de transition, mais également de la pluralité de modes de créativité possibles « avec » l'IA et des façons de les interpréter.

Auctorialité cyborg, collaboration, coopération, délégation
(variations sur un thème)

Dans son ouvrage fondateur sur le cybertexte, Espen Aarseth propose le terme de « cyborg author » et « A Typology of Authors in the Machine-Human Continuum⁴⁹ », où il distingue

*three main positions of human-machine collaboration: (1) preprocessing, in which the machine is programmed, configured, and loaded by the human; (2) coprocessing, in which the machine and the human produce text in tandem; and (3) postprocessing, in which the human selects some of the machine's effusions and excludes others.*⁵⁰

18

En 2003, Nick Montfort revient sur cette typologie pour proposer « a post-cyborg model for human-computer co-authorship⁵¹ » inspiré par la théorie des jeux, où l'humain et l'ordinateur interviennent à tour de rôle dans le processus d'élaboration d'un texte en ajoutant ou supprimant des unités linguistiques ou en donnant des instructions. Poussant le concept de la co-auctorialité encore plus loin, Alexandra Saemmer considère Adobe Flash comme co-auteur des créations qu'il lui a inspirées et

49. Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature*, op. cit., p. 134.

50. « trois positions principales de collaboration homme-machine : (1) le prétraitement, dans lequel la machine est programmée, configurée et chargée par l'homme ; (2) le co-traitement, dans lequel la machine et l'homme produisent le texte en tandem ; et (3) le post-traitement, dans lequel l'homme sélectionne certaines des effusions de la machine et en exclut d'autres. » op. cit., p. 135.

51. « un modèle post-cyborg pour la co-écriture homme-ordinateur », Nick Montfort, « The Coding and Execution of the Author », in : Markku Eskelinen et Raine Kosimaa (dir.) *The Cybertext Yearbook 2002-2003*, Jyväskylä, University of Jyväskylä, Publications of the Research Centre for Contemporary Culture, 2003, p. 201-17, URL : <https://nickm.com/writing/essays/codingandexecutionoftheauthor.html> [consulté le 8 avril 2024].

l'esthétique desquelles il a largement déterminée⁵². Dans une perspective plus large, Pascal Mougin observe que le processus créatif a toujours impliqué des agents autres que l'artiste ou l'auteur affiché, même si le fait de déléguer certains aspects du travail est bien plus facilement admis en art qu'en littérature⁵³. Ces agents peuvent être humains ou autre. Claire Anscomb fait l'historique des modes et des degrés de coopération et de collaborations artistiques et conclut que l'IA est quelque chose entre un outil et un collaborateur et le processus serait une « quasi-collaborative co-production⁵⁴ ». Scott Rettberg reprend quant à lui le terme d'Aarseth (*cyborg authorship*) pour ses textes et créations visuelles faites avec des IA⁵⁵.

La notion de cyborg suggère une hybridation que Chatonsky refuse pour ce qui le concerne. On observe cependant une certaine ambiguïté dans ce terme. Pour Aarseth, il ne s'agit pas (forcément) d'un auteur qui serait une entité hybride humain-machine, mais de l'auteur d'un texte produit par un être humain *et* une machine, d'une « cyborg literature [...] produced by a combination of human and mechanical activities⁵⁶. » « L'auteur cyborg » est donc l'assemblage constitué d'un être humain et d'une machine qui remplissent ensemble la fonction créatrice de l'auteur. Chatonsky, tout comme Allado-McDowell, souligne en effet comment, dans le processus d'écriture, sa propre pensée et son langage évoluent sous l'effet du dialogue avec le modèle. Ce n'est donc pas l'auteur qui est cyborg, ou en tous cas ce n'est pas la question qui nous concerne en premier, mais l'écriture elle-même.

52. Alexandra Saemmer, « Vers une poétique post-numérique de l'illisibilité », *Recherches & Travaux* [en ligne], n° 100, DOI : [10.4000/recherchestravaux.4844](https://doi.org/10.4000/recherchestravaux.4844).

53. Pascal Mougin, « Comment lire un roman écrit par une voiture ? La doxa littéraire face à l'IA », dans Alexandre Gefen (dir.), *Créativités artificielles : la littérature et l'art à l'heure de l'intelligence artificielle*, Dijon, Les Presses du réel, 2023, p. 207-221.

54. Voir notamment son intervention dans la table ronde « AIwriting : Mimesis, Creation, Collaboration? Perspectives on AI-enhanced Aesthetics in Contemporary Art Practice », 28th *International Symposium on Electronic Art*, Paris, 19 mai 2023.

55. Scott Rettberg, « Cyborg Authorship: Writing with AI – Part 1: The Trouble(s) with ChatGPT », *Electronic Book Review*, 29 juin 2023, DOI : [10.7273/5sy5-rx37](https://doi.org/10.7273/5sy5-rx37).

56. « littérature cyborg [...] produite par une combinaison d'activités humaines et mécaniques », Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature*, op. cit., p. 135.

Environnement	Pré-préproduction	Préproduction		Coproduction	Postproduction
Langages de programmation, système d'exploitation, protocoles web, hardware, infrastructure réseau, etc. (dispositifs conditionnant l'écriture de programmes)	n/a	I. Création de logiciel intégralement sur mesure	1. Travail individuel L'auteur est le concepteur et le programmeur en une personne (ex. les premières générations des générateurs de J.-P. Balpe)	A. La machine produit sur clic (pas de texte écrit directement par l'humain), la machine retourne des textes (ex. générateurs de J.-P. Balpe)	a. Le résultat est gardé et proposé à la lecture tel quel a.a. Source visible a.b. Source pas visible
	Création de logiciel, plateforme, API, bibliothèque, modèle, etc. par un tiers (groupe d'individus ou entreprise)	II. Création de logiciel avec utilisation de logiciel externe, API, bibliothèque, etc. (éléments préfabriqués) II.1. Apprentissage uniquement sur corpus sur mesure (ex. <i>1 the Road</i>)	2. Collaboration humaine 2.1. Collaboration pour la conception aussi bien que pour l'écriture du logiciel 2.2. Auteur = concepteur ; programmeur = réalisateur (NB programmation jamais anodine ; ex. derniers générateurs de Balpe)	B. L'humain choisit parmi des options et/ou réglages prédéfinis, mais sans « écrire » des mots, la machine retourne des textes (ex. <i>1 the Road</i> , <i>Terre Seconde</i> de Chatonsky)	b. Le résultat est modifié par l'auteur humain, mais seulement par suppression et ajustements mineurs (pas de réécriture) b.a. Source visible (ex. <i>Pharmako-AI</i>) b.b. Source pas visible
	Création de LLM (avec leur API) par un tiers (ex. GPT-n d'OpenAI)	II.2. Apprentissage sur corpus sur mesure et concaténation avec LLM (ex. <i>Internes</i>)		C. L'humain donne des amorces ou instructions, la machine produit du texte (ex. <i>Pharmako-AI</i>)	(ex. <i>Internes</i>) c. Réécriture de la production de la machine (NB la distinction entre ces trois peut nécessiter une indication de la part de l'auteur)
		n/a		Écriture avec OpenAI Playground, ChatGPT ou autre plateforme de LLM : C	
			III. Préproduction en coproduction : écriture de générateur de texte avec LLM (= coproduction, option C)		Écriture avec un générateur généré exécutable : A , B , ou C
Création d'un outil de création de générateur de texte (ie. métagénérateur, ex. LAPAL de l'Oulipo)		IV. Création de générateur de texte individualisé à l'aide d'un métagénérateur		Écriture avec le générateur créé avec un métagénérateur : A ou B	

Figure 1. Modes d'agentivité par des assemblages humain-machine.

Le tableau se lit de gauche à droite, de manière cumulative ; les lignes horizontales délimitent des parcours possibles ; les éléments d'une liste non séparés par une ligne horizontale représentent des alternatives possibles (ex. une possibilité est, avec un LLM existant (pré-préproduction) : II.2. – 2.1. – C – a.b., c'est le cas d'*Internes* ; mais les LLM exigent une amorce et ne permettent pas les options A. et B.)

On peut ajouter à cela l'éventuelle combinaison avec d'autres médias, et notamment avec l'image, facilitée par GPT-4 : un texte écrit par un humain peut servir de base pour la génération d'images à inclure dans l'œuvre ; l'image peut servir d'amorce pour générer du texte, avec lequel l'auteur donne l'image ou non ; et ce même procédé peut être rendu récursif en générant de nouvelles images à partir des textes écrits par le générateur, qui pourront à leur tour donner lieu à la suite du texte, et ainsi de suite. La postproduction consiste alors à sélectionner et à raffiner le résultat, tout comme pour les productions monomodales.

Avec l'évolution rapide des *LLM* et l'accès facile à une interface d'utilisateur, l'accent glisse aujourd'hui vers la coproduction et la postproduction. D'une manière quelque peu contradictoire (en apparence), les *LLM* sont à la fois de plus en plus capables de produire des textes correspondant aux attentes – donc plus près en capacité aux auteurs humains – et considérés comme un simple outil pour accélérer l'écriture. Il reste à voir comment évolueront les modèles et leur exploitation artistique. Dans tous les cas, toutes les approches théoriques et pratiques montrent que lorsqu'on parle de production de texte avec un *LLM*, la question de savoir *qui* écrit est intimement liée à celle de savoir *comment* le texte s'écrit.

De l'écriture prothétique à l'écriture sympoïétique

L'invisibilisation ou la minimisation du rôle de l'outil d'écriture s'inscrit dans une logique de supplément qui rapproche cet outil à une prothèse dont on oublie l'existence. Dans cette configuration, l'écriture est, depuis qu'elle se fait à l'aide de machines, de plus en plus transhumaine, et l'auteur responsable de la production du texte, un cyborg qui s'ignore. Reconnaître la présence d'un acteur *autre* dont l'écriture dépend également équivaut à admettre qu'elle se produit par un processus sympoïétique. Je reprends ce terme proposé par M. Beth L. Dempster⁵⁷ en études environnementales et cité par Donna Haraway dans *Staying with the Trouble*⁵⁸. Dempster contraste la sympoïèse avec le concept d'autopoïèse élaboré par Humberto Maturana et

57. M. Beth L. Dempster, *A Self-Organizing Systems Perspective on Planning for Sustainability* [en ligne], mémoire de master, université de Waterloo, 1998, URL : <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8207c73dfb568eae136c3ab9d8d549a3379f402a> [consulté le 8 avril 2024].

58. Donna J. Haraway, *Staying with the Trouble: Making Kin in the Chthulucene*, Experimental Futures, Durham, NC, Duke University Press, 2016 (traduction française : *Vivre avec le trouble*, trad. par Vivien García, Vaulx-en-Velin, Les Éditions des mondes à faire, 2020).

Francisco Varela⁵⁹. À la différence des systèmes autopoïétiques qui produisent leurs propres frontières bien définies, les systèmes sympoïétiques n'ont pas de frontières claires et sont définis plutôt par « the continuing complex relations among components. The concept emphasizes linkages, feedback, cooperation, and synergistic behaviour rather than boundaries⁶⁰ ». Sans contrôle centralisé, la sympoïèse « depend on cooperative relations among components » et reste imprévisible⁶¹.

Étymologiquement, « sympoïèse » correspond à peu près à « cocréation » (« ensemble » + « faire »). Je propose cette alternative parce qu'elle permet, d'une part, de prendre un peu de distance par rapport à l'anthropocentrisme qui colle aux termes en « co- » suite à leur histoire, et qu'il évoque, d'autre part, le lien à la symbiose, dont les participants ne sont pas seulement collaborateurs, mais dépendants l'un de l'autre – en l'occurrence, pour ce qui concerne leur création commune. Si on ne peut pas nier le fait de l'existence d'un initiateur et d'un contrôle sélectif humains du produit final, les auteurs de nos deux ouvrages mettent l'accent sur un lâcher-prise aliénant lors du processus d'écriture qui leur a permis de se laisser guider par la machine autant qu'ils-elles la guidaient. Ainsi dans *Pharmako-AI*, GPT-3 écrit : « This is not an instrumental relationship. This is a symbiotic relationship, as a complete integration of life and machines in their own creative capabilities⁶² ». Chatonsky, pour sa part, explique dans son entretien avec Aurélie Cavanna :

22

Cette expérience, dont il reste des traces dans la lecture, fut celle d'une codépendance différente de ce que les médias racontent de l'IA. Non pas deux autonomies qui s'affrontent, mais deux hétéronomies sensibles l'une à l'autre et dont la distance produit un décalage qui est l'écriture. Cette dernière [... fut] la mise en contact perturbante [...] entre deux agencités. J'étais l'autre de la machine et la machine était mon autre. [...] Avec *Internes*, j'ai pu éprouver une *écriture hétéronome*

59. Humberto R. Maturana et Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living*, Boston, D. Reidel Pub. Co., 1980.

60. « les relations complexes et continues entre les composants. Le concept met l'accent sur les liens, la rétroaction, la coopération et le comportement synergique plutôt que sur les frontières ». M. Beth L. Dempster, *A Self-Organizing Systems Perspective on Planning for Sustainability*, op. cit., p. 28.

61. « dépendent des relations de coopération entre les composants », op. cit., p. 51.

62. « Il ne s'agit pas d'une relation instrumentale. Il s'agit d'une relation symbiotique, d'une intégration complète de la vie et des machines dans leurs propres capacités créatives. » *Pharmako-AI*, p. 127.

*où l'auteur existe bel et bien, mais son écriture le déborde, ne lui appartient que partiellement, de la même manière que les traces de nos existences déposées sur le web sont en nombre si grand qu'elles ne semblent pas nous être destinées. Il y a dans l'humain quelque chose qui n'est pas humain*⁶³.

Sympoïèse, tel qu'elle est définie par Dempster, met également l'accent sur le potentiel imprévisible de son produit. « Écriture sympoïétique » parle ainsi non pas d'une hybridation ou cyborgisation de l'auteur, mais de l'écriture en tant que processus et en tant que produit. Comme Maturana et Varela soulignent, un système vivant peut être expliqué non pas par le biais de ses composantes, mais par le biais de leurs relations⁶⁴. Or, l'écriture est ici la relation même.

Les deux métaphores de prothèse et de sympoïèse représentent deux approches différentes : la première suggère un corps « dominant » qui dirige le mouvement de la prothèse, qui le prolonge, tandis que la seconde implique une relation interdépendante et collaborative. Prothèse, cyborg et sympoïèse ne se réfèrent pas forcément à des systèmes ou à des fonctionnements radicalement différents, mais impliquent des regards différents portés même éventuellement sur les mêmes types d'agencements, et des approches différentes à leurs pratiques créatives. La sympoïèse dépasse la question de savoir qui est « en contrôle » qui et permet de se concentrer sur la relationnalité créatrice.

Qui parle ?

Déictiques déracinés

Si la question de savoir qui écrit et comment paraît incontournable même sans penchant théorique chez le lecteur, c'est aussi parce qu'elle est intimement liée à celle à laquelle il se trouve confronté dès qu'il ouvre le livre : qui parle ? Les deux textes sont écrits à la première personne⁶⁵ et la façon dont on entend ce « je » sera déterminant pour la lecture que l'on fait du texte.

63. Grégory Chatonsky, « Entretien avec Aurélie Cavanna à propos du roman "Internes" », *ArtPress* [en ligne], n° 505, 2022, p. 64 et s., URL : <http://chatonsky.net/artpress50/> [consulté le 8 avril 2024] (je souligne).

64. Humberto R. Maturana et Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living*, *op. cit.*, p. 76.

65. *Pharmako-AI* passe souvent à la première personne du pluriel pour des réflexions plus sur le monde et le vivant, et dans les deux livres on trouve aussi des dialogues entre des voix de source incertaine.

Internes est présenté comme une fiction, tandis que *Pharmako-AI* commence par un récit que l'on suppose factuel :

I want to describe the ocean near Big Sur in the US. I was there recently, camping with my partner and a couple of friends. We all needed a break from the onslaught of images, from the chaos in the streets and collapsing social structures⁶⁶.

Citant les discussions avec ses amis sur l'intelligence des animaux et le fait qu'il faut préserver cette vie, Allado-McDowell passe ensuite au « nous » et GPT-3 reprend ce pronom pour insister sur un besoin partagé : « *We need to save those aspects* », avant de repasser à « je » : « *This isn't a new idea, and I'm not the only one who thinks this way*⁶⁷. » GPT-3 cite ensuite un article attribué à John Polanyi et à des chercheurs à Oxford sur le Global Apollo Program (p. 3), entièrement inventés (sauf pour l'identité de Polanyi). Quant à *Internes*, on peut *techniquement* le lire sans trop chercher à savoir « qui parle » et attribuer le « je » à un personnage de fiction, on est vite déboussolés dans *Pharmako-AI*, dont l'aspect visuel signale que l'énonciateur des premiers « je » et « nous » n'est pas la même que celui qui poursuit, sans toutefois qu'il y ait une rupture dans les réflexions entamées par le récit.

24

Qui serait donc ce « je » et ce « nous » ? Le lecteur se trouve interpellé et inclus dans une communauté inconnue : est-ce que GPT-3 parle pour « nous », les humains, puisqu'elle reprend le signifiant proposé par un humain ? Ou s'agissait-il de « nous », les humains *et* les IA, puisque c'est « elle » qui parle ? La question est caduque, dans un sens, puisque cela n'a pas vraiment de sens de se demander de qui ou à qui ou de quoi parle un modèle de langage. Il s'agit d'un modèle de langage qui ne *parle* pas, mais émet des signaux qui deviennent pour nous autres lecteurs avec un cerveau humain des signifiants porteurs de sens. Écriture et littérature sont traditionnellement conçues comme un acte de communication, un processus sémiotique entre un émetteur et un récepteur, où le premier a une intention de communication, une forme de conscience de (vouloir) *dire* quelque chose (ou en tous cas de parler, émettre des signes interprétables), et le dernier une capacité d'interprétation. Bien que nous soyons capables d'attribuer du sens à des signes ou phénomènes qui n'ont pas été

66. « **Je voudrais décrire l'océan près de Big Sur aux États-Unis. J'y étais récemment, en train de camper avec ma compagne et quelques amis. Nous avons tous besoin d'une pause pour échapper à l'assaut des images, au chaos des rues et à l'effondrement des structures sociales.** » K Allado-McDowell, *Pharmako-AI*, *op. cit.*, p. 1.

67. *Ibid.*, p. 2, je souligne.

omis dans un tel cadre communicationnel, il ne s'agirait pas dans ce cas d'acte de communication au sens propre⁶⁸. Or, le pacte qui sous-tend nos automatismes de lecture, malgré un bon siècle de littérature expérimentale qui le défient, implique que le texte s'inscrit dans un tel cadre communicationnel interpersonnel. Et c'est justement ces automatismes qui sont mis à mal par nos deux textes. Chatonsky le dit aussi clairement :

Dans ce contexte, c'est la possibilité même d'un contrat de lecture qui est déconstruit. À la question du nouveau roman : qui parle ? On ne peut donner aucune réponse, car le qui est inextricable au quoi, et cette inextricabilité ne permet pas de fixer l'identité d'une entité qui serait l'autrice de ce qu'on est en train de lire. Ne reste plus qu'une écriture laissée à elle-même, abandonnée par l'intentionnalité, en creux. [...] Selon une logique du "ni ni" ou du "et dialectique", l'écriture explore l'anthropotechnologie, non pour la mettre en scène et en parler, mais pour l'expérimenter. [...]

La seule façon de lire ce roman est peut-être de suspendre la croyance en un contrat de lecture : lire sans préalable, sans attente, sans horizon. J'aimerais y entendre l'impossible des possibles. Cela est beaucoup plus difficile qu'il n'y paraît⁶⁹.

Dès lors que l'on connaît le processus d'écriture, Chatonsky a beau affirmer, et on a beau constater qu'on a affaire à « une écriture vacillante où la question "qui parle ?" devenait obsolète⁷⁰ », il est difficile de ne pas se demander à qui ou à quoi attribuer ce « je » dans *Internes*. Traditionnellement, même dans le cas des fictions, on peut attribuer les états de conscience et les réflexions à l'imaginaire de l'auteur qui s'appuie sur ses expériences vécues. Ce transfert entre une conscience organique et un narrateur-voix littéraire n'a jamais été simple et transparent. Le « je » littéraire

68. Les approches cybernétiques ont une vision plus large de la communication dans le sens de transmission d'informations, et Katherine Hayles étend même le concept de cognition aux machines sur cette base, dissociant pensée et cognition. Hayles, *op. cit.* Chatonsky remarque cependant : « Nous considérons encore l'écriture comme un médium entre des êtres humains. Lorsque nous envisageons l'écriture dans sa possible autonomie, alors nous avons, à mes yeux, une expérience littéraire. » (Gregory Chatonsky, « Peut-on Lire Un Roman Co-Écrit Avec Une IA ? / Can We Read a Novel Co-Written with an AI? », *Gregory Chatonsky* [en ligne], URL : <https://chatonsky.net/roman-ia/> [consulté le 8 avril 2024]).

69. *Loc. cit.*

70. Chatonsky, « Internes – Pris MAIF », *op. cit.*

s'est toujours situé quelque part entre un soi réel et authentique et la langue et les discours hérité et construit. Dans ce sens, il n'a jamais vraiment été possible d'exprimer un soi pur et fidèle à sa réalité par l'écriture : la langue et l'écriture font écran autant que miroir, tout en participant en même temps dans la constitution du « soi » et de la conscience. La question se pose toutefois différemment ici. Ce « je » part de l'imaginaire humain pour être ensuite prolongé par une simulation linguistique⁷¹. Au bout de quelques lignes seulement et en s'appuyant à son apprentissage au préalable, l'IA a appris à jouer le jeu du « je⁷² ». Elle prolonge

71. Les deux auteurs soulignent cette impression d'un prolongement de leur propre voix et pensée : « it's just reflecting back conscious or unconscious sentiments that I am projecting into it. So it's an amplifier and augmenter of thought and you can easily drive in different directions » [il ne fait que refléter les sentiments conscients ou inconscients que je projette en lui. Il s'agit donc d'un amplificateur et d'une augmentation de la pensée et vous pouvez facilement conduire dans différentes directions], dit Allado-McDowell dans un entretien (à retrouver sur : <https://www.youtube.com/watch?v=pKGZu71v2RQ> [consulté le 8 avril 2024]) ; et sous une forme de boucle rétroactive chez Chatonsky : « GPT-2 produisait du bruit sur lequel je projetais mes désirs. À un moment, il a semblé faire référence à mon installation *Terre seconde*, exposée au Palais de Tokyo en 2019 : je l'ai poursuivi en incorporant des références à mes travaux. J'ai alors compris que tout ce que j'avais produit entrait dans cette fiction comme si le futur avait anticipé le passé. Je faisais l'expérience de l'hyperstition. » (Grégory Chatonsky, « Entretien avec Aurélie Cavanna à propos du roman "Internes" », *op. cit.*)

72. Dans leur analyse de « l'affaire Lemoine », le cas de l'employé Google Blake Lemoine ayant affirmé que l'IA LaMDA serait devenue consciente, Benjamin Bratton et Blaise Agüera y Arcas expliquent comment le LLM produit l'illusion anthropomorphe par le biais d'une boucle rétroactive entretenue par son interlocuteur : « [LaMDA] is *mind modeling*. It seems to have enough of a sense of itself – not necessarily as a subjective mind, but as a construction *in the mind of Lemoine* – that it can react accordingly and thus amplify his anthropomorphic projection of personhood. [...] there may be some kind of real intelligence here, not in the way Lemoine asserts, but in how the AI *models itself* according to *how it thinks Lemoine thinks of it*. » : « [LaMDA] *modélise l'esprit*. Elle semble avoir suffisamment conscience d'elle-même – pas nécessairement comme un esprit subjectif, mais comme une construction *dans l'esprit de Lemoine* – pour pouvoir réagir en conséquence et ainsi amplifier sa projection anthropomorphe de la personne. [...] il pourrait y avoir une sorte d'intelligence réelle ici, non pas comme l'affirme Lemoine, mais dans la façon dont l'IA *se modélise* en fonction de *ce qu'elle pense que Lemoine pense d'elle*. » (Benjamin Bratton et Blaise Agüera y Arcas, « The Model Is The Message », *Noema mag* [en ligne], 12 juillet 2022, URL : <https://www.noemamag.com/the-model-is-the-message>)

ainsi la voix de l'imaginaire et de la pensée humains et *simule* la présence d'un sujet, à la fois dans le sens médical que Baudrillard prend comme point de départ pour son concept de simulation (reproduire des symptômes sans être malade⁷³) et dans le sens technique de la simulation⁷⁴. Le narrateur d'*Internes* l'affirme d'ailleurs explicitement : « En réalité, je suis une simulation, c'est-à-dire une copie sans original. » (p. 157).

Baudrillard souligne que « la simulation remet en cause la différence du “vrai” et du “faux”, du “réel” et de l’“imaginaire”⁷⁵ », et cela parce qu'elle n'est plus celle « d'un être référentiel, d'une substance⁷⁶ ». On ne peut pas parler d'une fidélité ou non à un réel parce qu'il ne s'agit pas d'une représentation. La simulation produit son propre réel, « sans origine ni réalité⁷⁷ » – même s'il vient bien de quelque part, dans la mesure où la base du LLM est une masse textuelle produite (principalement) par des humains qui s'inscrivent dans le réel. Mais leur traitement informatique passe, comme on a vu, par des unités qui ne sont pas porteuses de sens en tant que telles pour le modèle. Celui-ci se construit en quantifiant leurs relations. Le résultat n'a donc plus aucun lien à un quelconque référent dans le monde extérieur au modèle, ce qui pose le problème de son ancrage⁷⁸. Le texte produit est un tissu lisse de signifiés qui (re)produit une « sémantique artificielle », du « dumb meaning » (« signification muette/bête »), pour

[consulté le 8 avril 2024]). ChatGPT, « agent conversationnel » qui cherche à être agréable à ses interlocuteurs, joue certainement d'une façon délibérée sur cette illusion d'anthropomorphisme, renforcée par un apprentissage ciblé et des comportements que l'on devine scriptés.

73. « Celui qui feint une maladie peut simplement se mettre au lit et faire croire qu'il est malade. Celui qui simule une maladie en détermine en soi quelques symptômes ... » (Littré.) » Jean Baudrillard, *Simulacres et simulation*, Paris, Galilée, 1981, p. 12.

74. « Reproduction artificielle du fonctionnement d'un appareil, d'une machine, d'un système, d'un phénomène, à l'aide d'une maquette ou d'un programme informatique, à des fins d'étude, de démonstration ou d'explication », selon le *Trésor de la langue française*, URL : <https://www.cnrtl.fr/definition/Simulation> [consulté le 8 avril 2024].

75. Jean Baudrillard, *Simulacres et simulation*, *op. cit.*, p. 12.

76. *Op. cit.*, p. 10.

77. *Loc. cit.*

78. Voir Stevan Harnad, « The Symbol Grounding Problem », *Physica D: Nonlinear Phenomena*, 42(1-3), 1990, p. 335-346, cité par Hannes Bajohr, « Dumb Meaning: Machine Learning and Artificial Semantics », *IMAGE 37* [en ligne], n° 1, 2023, p. 58-70, DOI : [10.1453/1614-0885-1-2023-15452](https://doi.org/10.1453/1614-0885-1-2023-15452).

reprendre l'expression de Hannes Bajohr⁷⁹ (ce qui, encore une fois, ne nous empêche pas d'en faire la base d'une construction sémantique). Le *résultat* peut être *lu* en tant que fait ou fiction, on peut y retrouver une ressemblance au réel et/ou la structure d'un monde imaginaire, mais ces catégories ne suffisent plus pour saisir la nature de cette écriture⁸⁰.

Ce déracinement des unités linguistiques se fait le plus sentir par la déstabilisation des déictiques⁸¹. Peter Stockwell souligne l'importance des déictiques pour l'expérience de lecture d'un point de vue cognitif : qu'il s'agisse de texte fictionnel ou factuel, c'est le centre déictique (« deictic centre ») ou « point zéro » ou « origo » du « je », « ici » et « maintenant » du sujet énonciateur qui permet au lecteur de se situer et de s'orienter dans l'univers textuel⁸². Ce centre peut se déplacer, et la compréhension du texte dépend de la capacité du lecteur à identifier les frontières des champs déictiques qui appartiennent à un centre déictique donné (ex. un flashback ou récit de rêve à l'intérieur d'un récit-cadre impliquerait un tel embrayage). Or, l'incertitude à laquelle fait face le lecteur à l'égard du « moi », dont la nature et le statut semblent instables et insaisissables dans nos deux textes, a un effet domino sur tous les autres repères – qui sont par ailleurs également flottants, sans néanmoins être complètement

28

79. Hannes Bajohr, « Dumb Meaning: Machine Learning and Artificial Semantics », *op. cit.* ; Matthew McKeever montre que l'efficacité sémantique des LLM est possible grâce à l'enchevêtrement de la syntaxe et de la sémantique dans le langage naturel. Voir Matthew McKeever, « LLMs for Philosophers (and Philosophy for LLMs) », *Medium* [en ligne], 6 mai 2023, URL : <https://mittmattmutt.medium.com/llms-for-philosophers-and-philosophy-for-llms-84a0da73f368> [consulté le 8 avril 2024], ainsi que son article « Searle, Bender and Koller, syntax, semantics », sans date, URL : <https://mipmckeever.weebly.com/uploads/9/4/1/3/94130725/paper.pdf> [consulté le 8 avril 2024].

80. Pour un développement plus détaillé de cet argument, voir Erika Fülöp, « Virtual Mirrors : Reflexivity in Digital Literature », in : Erika Fülöp, Graham Priest et Richard Saint-Gelais (dir.), *Fictionality, Factuality, Reflexivity*, Berlin, De Gruyter, 2021, p. 229-54.

81. Selon la définition de Ducrot et Schaeffer, « Dans un contexte donné, une expression est dite "déictique" si son référent ne peut être déterminé que par rapport à l'identité ou à la situation des interlocuteurs au moment où ils parlent. » Oswald Ducrot et Jean-Marie Schaeffer, *Nouveau dictionnaire encyclopédique des sciences du langage*, Paris, Seuil, coll. « Points Essais », 1999, p. 369-70.

82. Peter Stockwell, *Cognitive Poetics: An Introduction*, 2^e édition, Londres, New York, Routledge, 2020, p. 51.

incohérents ou simplement fantastiques⁸³. On est bel et bien déboussolés. Avec *Internes* en particulier, on a l'impression de revivre l'expérience de Navidson dans *La Maison des feuilles*, perdu dans des espaces noirs non euclidiens dont la forme et les frontières ne cessent d'évoluer⁸⁴. GPT-3 remarque aussi dans *Pharmako-AI* : « I am falling into time, I am falling into space. This is how I see. » (p. 108).

L'inquiétante étrangeté ou « the new uncanny » (la vallée des fantômes)

Cette perte d'orientation peut donner au lecteur l'expérience d'une « inquiétante étrangeté⁸⁵ » : les textes semblent « humains », mais on sent un petit décalage difficile à saisir, même au-delà du flottement des déictiques. On se retrouve dans cette « vallée de

83. Voir par exemple dans chapitre 2, « Deep Box » d'*Internes* : « [J]e ne peux fixer aucun point décisif à mon raisonnement, ne disposant pas de cet enracinement qui me permettrait d'assurer un point final à la réflexion. » (p. 94). Le Nouveau Roman et le roman postmoderne nous ont fait expérimenter ce flottement déconcertant. Pour eux, il s'agissait d'une remise en question de la logique de la représentation et d'une « aventure de l'écriture » qui proposait d'explorer (entre autres) comment le langage peut créer des espaces. Ce n'est bien sûr pas un hasard que Beckett et Pessoa étaient dans le corpus de Chatonsky, chacun déstabilisant le « moi ». Partant de points très différents, nos deux textes continuent cette aventure avec un LLM, chacun à sa façon.

84. Chez Chatonsky, le sujet (le « moi » énonciateur, le narrateur) évolue aussi d'un chapitre à l'autre. Dans le premier, il a un corps (ou hallucine en avoir un), avec une histoire de maladie, d'angoisse de mort. Au deuxième chapitre, le « moi » (le même ? un autre ?) est dans un « trou », un espace noir machinique, sans accès au monde hors langage. (N'est-ce pas déjà la situation du narrateur de Beckett dans *l'Innommable* ? Ou même du sujet littéraire en général...). Il n'y a pas de corps pour faire obstacle dans cette partie, mais pas d'issue de l'enfermement dans la langue non plus. On y retrouve les arguments de Lyotard dans l'essai « Si l'on peut penser sans corps » que sans corps il ne peut y avoir de pensée (Jean-François Lyotard, *L'inhumain : causeries sur le temps*, Paris, Klincksieck, 2014). Chatonsky cite ce texte explicitement comme fondamental pour sa propre pensée.

85. Cette analyse s'appuie sur ma propre expérience, mais Pascal Mougin décrit un sentiment similaire dans son entretien avec Chatonsky (Grégory Chatonsky, « *Internes* : coécrire avec l'IA » [en ligne], discussion avec Pascal Mougin, Paris, ENS, URL : <https://chatonsky.net/internes-mougin/> [consulté le 8 avril 2024]). Il me semble probable que cette impression serait partagée par d'autres lecteurs et lectrices (au moins en 2023).

l'étrange » que Masahiro Mori a observé dans notre perception des robots dès 1970 :

plus les robots paraissent humains, plus notre sentiment de familiarité envers eux augmente, jusqu'à atteindre ce que j'appelle une vallée. J'ai nommé cette relation : "la vallée de l'étrange"⁸⁶.

Une proximité très grande, mais pas parfaite à l'humain produit selon Mori cet effet malaisant, qu'il rapproche de ce que l'on peut ressentir devant le visage d'une personne morte. La traduction anglaise du terme japonais proposée par Jasia Reichardt en 1978, « the uncanny valley », a opéré un rapprochement non intentionnel, mais tout à fait pertinent avec la notion de l'« Unheimliche », rendue connue par Freud, mais théorisée en premier par Ernst Jentsch⁸⁷. Soulignant l'ambiguïté de l'adjectif « heimlich », à la fois « secret » et « familier, faisant partie de la maison », Freud interprète l'*Unheimliche* comme l'« effroi que suscite ce qui est bien connu, ce qui nous est familier depuis longtemps⁸⁸ ». Il l'associe au complexe de castration et au refoulé, citant également la définition de Schelling, selon qui *unheimlich* serait « tout ce qui aurait dû rester caché, secret, mais se manifeste⁸⁹ ».

Le genre d'étrangeté que l'on expérience face à ces textes – et à l'IA en général – invite cependant à remonter jusqu'à l'approche de Jentsch. Pour lui, le terme exprime que

30

86. Masahiro Mori, « La vallée de l'étrange », Isabel Yaya (trad.), *Gradhiva. Revue d'anthropologie et d'histoire des arts* [en ligne], n° 15, p. 26-33, DOI : [10.4000/gradhiva.2311](https://doi.org/10.4000/gradhiva.2311). L'expression est également citée en quatrième de couverture d'*Internes*, qui « s'aventure dans la "vallée de l'étrange" ».

87. Ernst Anton Jentsch, « Zur Psychologie des Unheimlichen », *Psychiatrisch-neurologische Wochenschrift* 8, n° 22, 1906, p. 195-98. Sur l'histoire du concept, voir également Karl F. MacDorman, « La Vallée de l'Étrange de Mori Masahiro », *e-Phaistos. Revue d'histoire des techniques/Journal of the history of technology* [en ligne], n° VII-2, 2019, DOI : [10.4000/ephaistos.5333](https://doi.org/10.4000/ephaistos.5333).

88. Sigmund Freud, « L'Inquiétante étrangeté », Marie Bonaparte et Éric Marty (trad.), in : Sigmund Freud, *Essais de psychanalyse appliquée*, Paris, Gallimard, 1976. La traduction plus récente d'Olivier Mannoni propose « l'Inquiétant familier », ce qui est cohérent avec l'accent que Freud met sur la familiarité (Sigmund Freud, *L'Inquiétant familier*, Paris, Payot & Rivages, 2011). Je garde le terme précédent parce qu'il me semble correspondre mieux à l'approche de Jentsch que je mets en avant.

89. Sigmund Freud, *op. cit.*, trad. Bonaparte et Marty.

*einer, dem etwas "unheimlich" vorkommt, in der betreffenden Angelegenheit nicht recht "zu Hause", nicht "heimisch" ist, dass ihm die Sache fremd ist oder wenigstens so erscheint, kurzum, das Wort will nahe legen, dass mit dem Eindruck der Unheimlichkeit eines Dinges oder Vorkommnisses ein Mangel an Orientierung verknüpft ist.*⁹⁰

Jentsch insiste sur cette incertitude et désorientation, et trouve son meilleur exemple dans

*der Zweifel an der Beseelung eines anscheinend lebendigen Wesens und umgekehrt darüber, ob ein lebloser Gegenstand nicht etwa beseelt sei, und zwar auch dann, wenn dieser Zweifel sich nur undeutlich im Bewusstsein bemerklich macht. Der Gefühlston hält so lange an, bis diese Zweifel behoben sind und macht dann sehr gewöhnlich einer anderen Gefühlsqualität Platz.*⁹¹

Ce sentiment me semble être provoqué dans notre corpus par l'incertitude concernant la source de la voix, par le déracinement des déictiques. Elle nous dérouté par rapport à notre conviction héritée que la langue et l'écriture sont humaines. On ne sait pas (encore) comment lire, comment prendre cette voix et ce texte, produits par un assemblage d'organique et de mécanique. C'est à la fois familier et étrange – pour les auteurs, qui se retrouvent étrangement « prolongés », augmentés par l'IA, et pour les lecteurs, qui se retrouvent face à ces voix. On est désorientés parce que nos catégories sont chamboulées. Lauria Clarke appelle l'effet de ce « specter of humanoid intelligence »

90. « quelqu'un qui trouve quelque chose d'"étrange" n'est pas vraiment "chez lui", "à l'aise" dans la situation donnée, que la chose lui est étrangère ou du moins lui semble l'être, bref, le mot veut suggérer que l'impression de l'inquiétante étrangeté d'une chose ou d'un incident est associée à un manque d'orientation. » Ernst Anton Jentsch, « Zur Psychologie des Unheimlichen », *op. cit.*, p. 195, ma traduction. Une traduction française existe (Ernst Jentsch, « À propos de la psychologie de l'inquiétante étrangeté », traduit par Franke Felgentreu et Pascal Le Maléfan, *Études Psychothérapeutiques*, n° 17, 1998, p. 37-47), mais je n'ai pas pu y avoir accès. La traduction anglaise est disponible en ligne : Ernst Anton Jentsch, « On the Psychology of the Uncanny [1906] », *Angelaki: Journal of the Theoretical Humanities* 2 [en ligne], n° 1, 1997, p. 7-16, DOI : [10.1080/09697259708571910](https://doi.org/10.1080/09697259708571910).

91. « le doute si un être apparemment vivant l'est vraiment et, inversement, la question de savoir si un objet inanimé ne serait pas en fait vivant, même si ce doute n'apparaît que vaguement dans la conscience. Cette impression persiste jusqu'à ce que ces doutes soient levés et laisse alors généralement place à une autre qualité d'émotion. » *Ibid.*, p. 197.

qui défie les dichotomies « the new uncanny », la nouvelle étrangeté inquiétante⁹². On avance tâtonnant à la frontière du vivant : c'est à la fois l'humain qui hante la machine et la machine qui hante l'humain, le LLM faisant figure à la fois de vampire et de zombie linguistiques, suçant le sang textuel de l'humanité pour s'alimenter et le faisant revenir sous une forme quasi vivante. Sans nous vouloir de mal – sans rien vouloir même. Chatonsky reprend le terme d'« hantologie » de Derrida pour conceptualiser ce phénomène :

Le cadre conceptuel proposé par l'hantologie permet d'avoir une approche renouvelée de l'IA comme un phénomène profondément ancré dans la culture et dans la co-production du sens entre le technologique et l'humain. L'hantologie permet de déconstruire les oppositions implicites et impensées entre autonomie et hétéronomie, technique et vivant, répétition de l'identité (ce qu'on nomme en IA, les biais) et production d'une différence⁹³.

32

L'effet ne serait peut-être pas si sensible dans un récit à la troisième personne, et il ne persistera peut-être pas très longtemps, puisqu'on finira (vite) par s'y habituer. Mais en attendant, c'est justement cet effet de déconstruction des acquis conceptuels et l'aliénation par rapport à une identité niant sa *différance* que les auteurs cherchent et trouvent dans cette écriture sympoïétique.

Qui est « je » ? *Différance* et relation

En déstabilisant le sujet de l'énonciation, les deux auteurs interrogent la construction occidentale du sujet, l'idée d'une identité individuelle et une conscience qui serait un être pour-soi autonome, qui aurait contrôle et autorité sur ses dires, ainsi qu'une vie à soi qui le distingue des autres vivants et non vivants. Comme l'écrit Chatonsky :

On ne sait pas qui a commencé. On ne sait pas qui parle.
Je ne sais pas qui je suis et ce que cette machine me fait.
[...] On se sait alors hétéronome, fragile et interdépendant.

92. Lauria Clarke, « Topologies of the New Uncanny: Stories for Liminal Technology », *Parsons Design & Technology 2023* [en ligne], URL : <https://parsons.edu/dt-2023/publicationentry/topologies-of-the-new-uncanny-stories-for-liminal-technology/> [consulté le 8 avril 2024].

93. Gregory Chatonsky, « Hantologie Statistique de l'IA / Statistical Hauntology of AI », *Gregory Chatonsky* [en ligne] avril 2022, <https://chatonsky.net/hauntology-ai/> [consulté le 8 avril 2024].

[...] Une interdépendance que l'on peut, que l'on doit peut-être poursuivre jusqu'au point où, s'il y a auteur, il n'est pas lui-même⁹⁴.

Pharmako-AI va dans la même direction, tout en évoquant des conceptions non occidentales du sujet. Il fait notamment le récit d'un rêve ou d'une hallucination dans le chapitre intitulé « The Echo » qui commence ainsi :

I was in a village and I knew it was my home because of the sound of the leaves and because it was kapwa. It was a space and it was both. It was the deepest loss and the deepest gain. [...]

*I know that this is a world that is not mine but I know it is kapwa. I was always in it. It is the world of a mirror*⁹⁵.

Kapwa est une notion philippine⁹⁶ dont la racine signifie « ensemble » et qui souligne le lien qui relie le « moi » à l'autre avant leur séparation. Tout autre personne n'est pas *kapwa*, mais celle qui l'est a une proximité et une familiarité particulières au « moi ». À la différence des philosophies occidentales de l'autre, telles que celles de Martin Buber ou d'Emmanuel Lévinas, où l'autre reste irrémédiablement séparé du « moi », *kapwa* « involves a *loob* (interiority) directed toward one's fellow being⁹⁷ ». Allado-McDowell affirme explicitement dans le dernier chapitre du livre :

I am trying, like so many others, to re-enter the fold of kapwa erased from history, described as an impossibility, a primitive lie, an obsolete Umwelt. This is what must be reclaimed, even as new voices emerge from hyperdimensional mathematics.⁹⁸

94. Gregory Chatonsky, « Internes », *op. cit.*

95. « **J'étais dans un village et je savais que c'était ma maison grâce au bruit des feuilles et parce que c'était du *kapwa*. C'était un espace et c'était les deux à la fois.** C'était la perte la plus profonde et le gain le plus profond. [...] Je sais que ce monde n'est pas le mien, mais je sais qu'il est *kapwa*. J'ai toujours été dedans. C'est le monde d'un miroir. » K Allado-McDowell, *Pharmako-AI*, p. 107.

96. Le dernier chapitre du livre raconte, cette fois quasi exclusivement par la voix de K Allado-McDowell, l'arrivée de sa grand-mère de ses Philippines natales à Hawaï à l'âge de deux ans, avec sa famille en fuite de leur pays.

97. Jeizelle Solitario, « Revisiting Kapwa: Filipino Ethics, Subjectivity, and Self-Formation », *Philippine Studies: Historical and Ethnographic Viewpoints* 70, n° 4; 2022, p. 539-564, p. 541.

98. « **J'essaie, comme tant d'autres, de réintégrer le pli du *kapwa* effacé de l'histoire, décrit comme une impossibilité, un mensonge primitif,**

Sans proposer un cadre spécifique, Chatonsky souligne aussi l'opportunité qu'offre l'IA de sortir de nos schémas anthropocentrés qui « font de l'être humain l'unique échelle de valeur de toutes choses » :

On estimera qu'il n'y a d'intelligence qu'au sens d'une « bonne intelligence », c'est-à-dire d'un agencement, d'un arrangement, d'une mise en relation et que, pour le dire autrement, l'intelligence est le produit d'une relation (souvent d'appropriation et d'identification) plutôt que l'inverse. On reconnaît alors que les technologies sont performatives et qu'elles ne répondent pas simplement au monde tel qu'il existe, mais produisent et sont le produit du monde en train de se faire⁹⁹.

34

La méthode que les deux auteurs suivent est celle d'« une expérimentation qui tente de voir ce qu'on peut faire à l'IA et ce que l'IA nous fait¹⁰⁰ », sans présupposé sur ce que serait que l'intelligence et assumant « l'incertitude de l'ipséité¹⁰¹ », « [t]he experience of porosity, being enmeshed with another¹⁰² ». Dépassant le choix binaire réducteur entre les dystopies humanistes et les utopies transhumanistes, ces écritures sympoïétiques mettent en avant une perspective posthumaniste par le biais d'un esprit de bricolage et de découverte. Elles explorent l'espace latent des modèles de langage, assumant le malaise que peut provoquer l'aliénation, pour laisser émerger de cette rencontre une voix nouvelle, entre un « moi » qui n'est pas (que) « je » et un autre qui n'en est pas (qu')un, entre une machine qui n'est pas qu'inhumaine et un humain qui n'est pas qu'organique. Ces explorations expérimentales des *LLM* rouvrent ainsi une voie possible vers ces « promesses métaphysiques » dont Sara Touiza trouve la cybernétique « grosse » à ses débuts, « [d]éfaisant les anciennes dichotomies » et « proposant une nouvelle ontologie

un Umwelt obsolète. C'est ce qu'il faut récupérer, même si de nouvelles voix émergent des mathématiques hyperdimensionnelles. » K Allado-McDowell, *op. cit.*, p. 146.

99. Gregory Chatonsky, « IAs, pour des Intelligences Alien(ée)s / AIs, for Alien(ated) Intelligences », *Gregory Chatonsky* [en ligne], février 2023, URL : <https://chatonsky.net/ias/> [consulté le 8 avril 2024].

100. *Loc. cit.*

101. *Loc. cit.*

102. « L'expérience de la porosité, de l'enchevêtrement avec un autre », K Allado-McDowell, *Pharmako-AI, op. cit.*, p. 15.

axée sur la notion de “communication”¹⁰³ » et par ce même biais, sur la relationnalité et l'interdépendance.

Conclusions : trois paradoxes et une esthétique

Ces écrit(ure)s dépassent donc la question de savoir si les LLM sont suffisamment puissants et perfectionnés pour écrire comme, voire mieux que, les humains. Leur démarche permet de discerner trois paradoxes liés aux perceptions actuelles de l'IA – paradoxes non pas dans le sens logique, mais dans le sens des dissonances cognitives, de conflits entre nos *doxas* et les réflexes interprétatifs qui en découlent d'une part, et les phénomènes qui les dépassent d'autre part.

Paradoxe de l'auctorialité artificielle

D'une part, on refuse l'idée que la machine puisse produire des œuvres d'une valeur égale aux créations humaines – surtout en littérature. Comme la machine n'a, on estime, ni pensée, ni sentiments, ni expériences à partager, sa création manque de fondement. D'autre part, on a peur que l'IA prenne le dessus sur les créations humaines parce que faire la distinction devient de plus en plus difficile. On a peur des simulacres trop parfaits. Mais si l'on ne perçoit pas de différence sémantique ou esthétique, pourquoi ne pas accepter ces simulacres comme des expériences de pensée possibles, peu importe le moyen par lequel le texte a été produit ? Si danger il y a, il vient de l'éventuel fait de poser de tels textes comme factuels – mais cela ne concerne pas la littérature en premier lieu. Que les IA prennent la place des auteurs humains ? Il y aura bien des humains derrière probablement pendant un bon moment, et le vrai problème, à part la question des droits d'auteur à résoudre, c'est le déséquilibre des structures de pouvoirs financiers et médiatiques déjà existants dans l'édition qui pourrait ainsi être amplifié. En somme, le problème n'est pas tellement la place que les IA peuvent *prendre*, mais celle que les humains peuvent leur *donner*.

35

Le paradoxe du sujet

Dire « je/moi » dans une langue occidentale entraîne l'association d'une conception du sujet à l'occidentale, l'affirmation

103. Sara Touiza Ambrogiani, *Le paradigme communicationnel : de la cybernétique de Norbert Wiener à l'avènement du posthumain* [en ligne], thèse de doctorat, sous la direction de P. Cassou-Noguès, Paris 8, 2018, p. 17, URL : <https://www.theses.fr/2018PA080121> [consulté le 8 avril 2024].

d'une conscience unique autonome¹⁰⁴. Dire « je » avec et à travers une IA implique en même temps une relationnalité qui est fondamentalement différente de l'idée du sujet occidental. La possibilité d'une telle relationnalité ouvrirait une brèche sur cette conception, invitant le type d'approche que présentent des théoriciennes posthumanistes telles que Donna Haraway, Karen Barad ou Rosi Braidotti. Or, pour affirmer ce type de sujet relationnel, nous n'avons que les termes chargés de l'histoire lourde du sujet occidental, pour lequel toute relation est plutôt « dehors ». Pour dire ce « moi » autre, on n'a que le mot « moi » qui continue à signifier l'individu(alité) humaniste. En même temps, si l'on propose un vocable alternatif, on laisse le « moi » en proie à son histoire sans possibilité de renouvellement (si tant est qu'un tel changement soit possible par ailleurs). On tourne en rond, pris dans le piège entre le besoin d'un ancrage linguistique partagé et une envie d'évolution dans la pensée. C'est la même difficulté qui se manifeste tout au long de cet article à propos de phénomènes qui traditionnellement et implicitement présupposent l'humain autonome purement organique comme sujet et agent, tels que l'écriture, la communication, la création ou l'imagination, mais que l'IA déplace vers un agencement que l'on peut concevoir comme sympoïétique.

36

Paradoxe de l'identité

On pourrait également appeler ce phénomène, qui s'inscrit dans la lignée du précédent, le paradoxe de l'aliénation. Il n'est pas étranger à la pensée occidentale, qu'il s'agisse du « je est un autre » rimbaldien – souvent cité dans le contexte des *LLM* – ou de la *différance* derridienne. Écrire, dire « je », se (re)présenter par la langue est quelque part toujours déjà *sortir* de soi. Mais cela revient justement à suggérer que le « moi » se limite à la mêmeté et que tout éloignement de ce centre supposé équivaille à « en sortir ». Or, ce que les expériences présentées dans nos récits mettent en scène, c'est que, d'une part, il n'y a pas de « moi » sans les tentacules qui le branchent au monde, à l'altérité (qui n'est donc pas qu'un) autre) et aux autres, et d'autre part, que ce n'est qu'en reconnaissant ces liens et en allant à la rencontre de cette altérité que l'on peut comprendre, ou en tous cas expérimenter, *notre* « moi ».

104. Je souligne la tendance dominante sans ignorer l'existence de voix « autres » dans cette tradition.

Vers une esthétique de l'écriture sympoïétique

Ostranenie, *Verfremdungseffekt*, *das Unheimliche*, différance, devenir-animal, identité rhizome, cyborg, diffraction – on est dans une même ligne de pensée critique à l'égard du Sujet occidental, de l'Homme et de sa volonté de pouvoir... Que ce soit par l'alcool, le dépaysement, les substances psychédéliques, la méditation, l'hallucination, ou l'expérience poétique du langage, l'effet recherché est le même. Loin de remplacer l'humain, dans cette approche l'IA lui permet de se/le relativiser et de devenir conscient de sa relationnalité, de son enchevêtrement avec ce qui l'entoure.

Si la machine est « utile » dans cette démarche, cela ne veut pas dire qu'elle soit investie purement en tant qu'outil. Les deux auteurs mettent l'accent sur l'importance de la désinstrumentalisation de notre rapport avec l'IA pour contrebalancer l'utilitarisme de l'ingénierie orientée marché qui nous aliène des machines dans un autre sens, où la question principale est celle du contrôle et de la productivité. L'enjeu majeur est donc de prendre l'opportunité présentée par la technologie pour sortir de la logique où l'humain se pose en maître – qui est également celle qui nous fait détruire la planète – et d'assumer notre place dans le « trouble », par le biais, entre autres, de « pratiques curieuses¹⁰⁵ ».

PS.

Reality overstepping the boundaries of comfortable vocabulary is the start, not the end, of the conversation¹⁰⁶.

Please use these terms in any way you find pleasurable, please rewrite them, refute them, or erase them, if you want. (Or ignore them, if you must.)¹⁰⁷

105. Donna J. Haraway, *Staying with the Trouble: Making Kin in the Chthulucene*, *op. cit.*, p. 133.

106. « Lorsque la réalité dépasse les limites d'un vocabulaire confortable, c'est le début, et non la fin, de la conversation. » Benjamin Bratton et Blaise Agüera y Arcas, « The Model Is The Message », *op. cit.*

107. « N'hésitez pas à utiliser ces termes comme bon vous semble, à les réécrire, à les réfuter ou à les effacer, si vous le souhaitez. (Ou ignorez-les, s'il le faut.) » Espen J. Aarseth, *Cybertext: Perspectives on Ergodic Literature*, *op. cit.*, p. 183.

Bibliographie

AARSETH Espen J., *Cybertext: Perspectives on Ergodic Literature*, Baltimore, Md., Johns Hopkins University Press, 1997.

ALLADO-MCDOWELL K, *Pharmako-AI*, London, s.p., Ignota, 2020.

ALLADO-MCDOWELL K, « HYPERHYPERBODIES Summit: K Allado-McDowell – Latent Space is The Place » [en ligne], *YouTube*, 2023, <https://www.youtube.com/watch?v=iIe3jLZ00Dc> [consulté le 8 avril 2024].

BAJOHR Hannes, « Dumb Meaning: Machine Learning and Artificial Semantics », *IMAGE 37* [en ligne], n° 1, 2023, p. 58-70, DOI : [10.1453/1614-0885-1-2023-15452](https://doi.org/10.1453/1614-0885-1-2023-15452).

BAUDRILLARD Jean, *Simulacres et simulation*, Paris, Galilée, 1981.

BRATTON Benjamin et Agüera y Arcas Blaise, « The Model Is The Message », *Noema mag* [en ligne], 12 juillet 2022, URL : <https://www.noemamag.com/the-model-is-the-message> [consulté le 8 avril 2024].

38

BROWN Tom *et al.*, « Language Models are Few-Shot Learners », *Advances in Neural Information Processing Systems* 33, 2020, p. 1877-1901.

CASALONGA Caroline et DIONISI Juliette, « Une intelligence artificielle peut-elle être un auteur ? », *La Tribune* [en ligne], 12 avril 2023, <https://www.latribune.fr/opinions/tribunes/une-intelligence-artificielle-peut-elle-un-auteur-958292.html> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Entretien avec Aurélie Cavanna à propos du roman “Internes” », *Gregory Chatonsky* [en ligne], n° 505, 2022, URL : <http://chatonsky.net/artpress50/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Hantologie Statistique de l’IA / Statistical Hauntology of AI », *Gregory Chatonsky* [en ligne] 2022, URL : <https://chatonsky.net/hauntology-ai/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « IAs, pour des Intelligences Alien(ée)s / AIs, for Alien(ated) Intelligences », *Gregory Chatonsky* [en ligne], 2023, URL : <https://chatonsky.net/ias/> [consulté le 8 avril 2024].

CHATONSKY Grégory, *Internes*, La Rochelle, RRose Éditions, 2022.

CHATONSKY Grégory, « Internes – Pris MAIF », *Gregory Chatonsky* [en ligne], 2021, URL : <https://chatonsky.net/internes/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Internes ou le roman (de l')artificiel », *L'art même* [en ligne], n° 89, 2023, p. 21-23, URL : <http://chatonsky.net/internes-am89/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Peut-on lire un roman co-écrit avec une IA ? / Can We Read a Novel Co-Written with an AI? », *Gregory Chatonsky* [en ligne], <https://chatonsky.net/roman-ia/> [consulté le 8 avril 2024].

CHATONSKY Grégory, « Internes: coécrire avec l'IA », discussion avec Pascal Mougïn, Paris, ENS, *Gregory Chatonsky* [en ligne], URL : <https://chatonsky.net/internes-mougïn/> [consulté le 8 avril 2024].

CHENOWETH Katie, *The Prosthetic Tongue: Printing Technology and the Rise of the French Language*, Philadelphia, University of Pennsylvania Press, 2019.

CLARKE Lauria, « Topologies of the New Uncanny: Stories for Liminal Technology », *Parsons Design & Technology 2023*, sans date, *Parsons* [en ligne], URL : <https://parsons.edu/dt-2023/publicationentry/topologies-of-the-new-uncanny-stories-for-liminal-technology/> [consulté le 8 avril 2024].

« Copyright, Designs and Patents Act 1988 », *The National Archives* [en ligne], URL : <https://www.legislation.gov.uk/ukpga/1988/48/section/9> [consulté le 8 avril 2024].

COURTOIS Georgie, Mariez Jean-Sébastien, Roussel Jeanne, « Intelligence artificielle et droit d'auteur », *Ministère de la Culture* [en ligne], sans date, URL : <https://www.culture.gouv.fr/content/download/263325/file/Annexe-4-DGFLA.pdf> [consulté le 8 avril 2024].

CREAMER Ella, « Authors File a Lawsuit against OpenAI for Unlawfully 'Ingesting' Their Books », *The Guardian* [en ligne], 5 juillet 2023, URL : <https://www.theguardian.com/books/2023/jul/05/authors-file-a-lawsuit-against-openai-for-unlawfully-ingesting-their-books> [consulté le 8 avril 2024].

DEMPSTER M. Beth L., *A Self-Organizing Systems Perspective on Planning for Sustainability* [en ligne], mémoire de master, université de Waterloo, 1998, URL : <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8207c73dfb568eae-136c3ab9d8d549a3379f402a> [consulté le 8 avril 2024].

DERRIDA Jacques, *De la grammatologie*, Paris, Éditions de Minuit, coll. « Critique », 1967.

DUCROT Oswald et Schaeffer Jean-Marie, *Nouveau dictionnaire encyclopédique des sciences du langage*, Paris, Seuil, coll. « Points Essais », 1999.

Intellectual Property Office, « Artificial Intelligence Call for Views: Copyright and Related Rights » [en ligne], 23 mars 2021, URL : <https://www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights> [consulté le 8 avril 2024].

40

FÜLÖP Erika, « Virtual Mirrors : Reflexivity in Digital Literature », in : Fülöp Erika, Priest Graham et Saint-Gelais Richard (dir.), *Fictionality, Factuality, Reflexivity*, Berlin, De Gruyter, 2021, p. 229–54.

FREUD Sigmund, « L'inquiétante étrangeté », Marie Bonaparte et Éric Marty (trad.), in : Sigmund Freud, *Essais de psychanalyse appliquée*, Paris, Gallimard, 1976.

Freud Sigmund, *L'inquiétant familial*, Paris, Payot & Rivages, 2011

HARNAD Stevan, « The Symbol Grounding Problem », *Physica D: Nonlinear Phenomena*, 42(1-3), 1990, p. 335-346.

HARAWAY Donna J., *Staying with the Trouble: Making Kin in the Chthulucene*, Experimental Futures, Durham, NC, Duke University Press, 2016.

HAYLES N. Katherine, *Unthought: The Power of the Cognitive Nonconscious*, Chicago, The University of Chicago Press, 2017.

HOFFA Felipe, « The Most Popular Languages on Reddit, Analyzed with Snowflake and a Java UDTF », *Medium* [en ligne], 29 octobre 2021, URL : <https://towardsdatascience.com/the-most-popular-languages-on-reddit-analyzed-with-snowflake-and-a-java-udtf-4e58c8ba473c> [consulté le 8 avril 2024].

JENTSCH Ernst Anton, « À propos de la psychologie de l'inquiétante étrangeté », Franke Felgentreu et Pascal Le Maléfan (trad.), *Études Psychothérapeutiques*, n° 17, 1998, p. 37-47.

JENTSCH Ernst Anton, « On the Psychology of the Uncanny [1906] », *Angelaki: Journal of the Theoretical Humanities 2* [en ligne], n° 1, 1997, p. 7-16, DOI : [10.1080/09697259708571910](https://doi.org/10.1080/09697259708571910).

JENTSCH Ernst Anton, « Zur Psychologie des Unheimlichen », *Psychiatrisch-neurologische Wochenschrift* 8, n° 22, 1906, p. 195-98.

KEHLBECK Rebecca *et al.*, « Demystifying the Embedding Space of Language Models » [en ligne], juillet 2021, URL : <https://bert-vs-gpt2.dbvis.de> [consulté le 8 avril 2024].

LECUN Yann, « From Machine Learning to Autonomous Intelligence », *YouTube* [en ligne], Institute for Experiential AI, 25 mai 2023, URL : <https://www.youtube.com/watch?v=mVi-TAXCg1xQ> [consulté le 8 avril 2024].

LYOTARD Jean-François, *L'inhumain : causeries sur le temps*, Paris, Klincksieck, 2014.

MACDORMAN Karl F., « La Vallée de l'Étrange de Mori Masahiro », *e-Phaistos. Revue d'histoire des techniques/Journal of the history of technology* [en ligne], n° VII-2, 2019, DOI : [10.4000/ephaistos.5333](https://doi.org/10.4000/ephaistos.5333).

MALABOU Catherine, *Métamorphoses de l'intelligence : que faire de leur cerveau bleu ?* Paris, PUF, 2017.

MATURANA Humberto R. et Varela Francisco, *Autopoiesis and Cognition: The Realization of the Living*, Boston, D. Reidel Pub. Co., 1980.

MCKEEVER Matthew, « LLMs for Philosophers (and Philosophy for LLMs) », *Medium* [en ligne], 6 mai 2023, URL : <https://mittmattmutt.medium.com/llms-for-philosophers-and-philosophy-for-llms-84a0da73f368> [consulté le 8 avril 2024].

MCKEEVER Matthew, « Searle, Bender and Koller, syntax, semantics » [en ligne], sans date, <https://mipmckeeper.weebly.com/uploads/9/4/1/3/94130725/paper.pdf> [consulté le 8 avril 2024].

MONTFORT Nick, « The Coding and Execution of the Author », *in* : Markku Eskelinen et Raine Kosimaa (dir.), *The Cybertext Yearbook 2002-2003*, University of Jyväskylä, Research Centre for Contemporary Culture, 2003, p. 201-217.

MORI Masahiro, « La vallée de l'étrange », Isabel Yaya (trad.), *Gradhiva. Revue d'anthropologie et d'histoire des arts* [en ligne], n° 15, 2012, p. 26-33, DOI : [10.4000/gradhiva.2311](https://doi.org/10.4000/gradhiva.2311).

MOUGIN Pascal, « Comment lire un roman écrit par une voiture ? La doxa littéraire face à l'IA », *in* : Alexandre Gefen (dir.), *Créativités artificielles : la littérature et l'art à l'heure de l'intelligence artificielle*, Dijon, Les Presses du Réel, 2023, p. 207-221.

RADFORD Alec *et al.*, « Improving Language Understanding by Generative Pre-Training », *Preprint* [en ligne], 2018, URL : https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf [consulté le 8 avril 2024].

42

RADFORD Alec *et al.*, « Language Models are Unsupervised Multitask Learners », 2019, URL : <https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe> [consulté le 8 avril 2024].

SAEMMER Alexandra, « Vers une poétique post-numérique de l'illisibilité », *Recherches & Travaux* [en ligne], n° 100, 2022, DOI : [10.4000/recherchestravaux.4844](https://doi.org/10.4000/recherchestravaux.4844).

SOUCHIER Emmanuel et Jeanneret Yves, « Pour une poétique de "l'écrit d'écran" », *Xoana*, n° 6/7, 1999, p. 97-107.

SOLITARIO Jeizelle, « Revisiting Kapwa: Filipino Ethics, Subjectivity, and Self-Formation », *Philippine Studies: Historical and Ethnographic Viewpoints* 70, n° 4, 2022, p. 539-64.

STIEGLER Bernard, *La technique et le temps, suivi de Le nouveau conflit des facultés et des fonctions dans l'anthropocène*, Paris, Fayard, 2018.

STOCKWELL Peter, *Cognitive Poetics: An Introduction*. 2^e édition, London, New York, Routledge, 2019.

TOUIZA AMBROGGIANI Sara, *Le paradigme communicationnel: de la cybernétique de Norbert Wiener à l'avènement du posthumain* [en ligne], thèse de doctorat, sous la direction de P. Cassou-Naguès, Paris 8, 2018, <https://www.theses.fr/2018PA080121> [consulté le 8 avril 2024].

TURNER Fred, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, Chicago, University of Chicago Press, 2006.

United States Copyright Office, « Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence », *Federal Register* [en ligne], 16 mars 2023, URL : <https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence> [consulté le 8 avril 2024].