

La quatrième mémoire

Grégory Chatonsky
& Yves Citton

Les dites « intelligences artificielles » (IA) signifient moins le remplacement des facultés humaines que l'apparition d'une nouvelle catégorie de mémoire générant de façon récursive des discours, images et sons qui peuvent ressembler aussi bien à des « créations » qu'à des documents « authentiques ». Cette automatisation de l'expression et de la représentation à partir des données massives accumulées sur le Web explique pourquoi la question des arts, loin d'être anecdotique, est devenue consubstantielle aux modèles statistiques actuels. On peut y percevoir l'émergence d'un nouveau réalisme qui déstabilise les archives du passé comme les perspectives de futur. La nature contrefactuelle de ce réalisme alien trouble ce sur quoi nous fondions auparavant les indices de vérité.

Rétentions, protentions, distentions

Selon Bernard Stiegler, les différentes rétentions sont des manières de la mémoire qui permettent une réitération, un rappel, une répétition¹. La

rétention *primaire* est l'immédiateté de l'intuition ou de la perception, telle la perception d'une note de musique. La rétention *secondaire* est la temporalisation (de l'entendement) qui compare, anticipe et rappelle différents événements, par exemple des notes qui, en se suivant, forment une mélodie musicale. La rétention *tertiaire* est l'inscription sur un support matériel de ces événements en permettant leur répétition technique, par exemple sur un disque où la mélodie se trouve enregistrée.

Les développements techniques des dernières décennies ont permis le développement de ce qu'on peut considérer comme des rétentions *quaternaires*. Celles-ci résultent, d'une part, de la multiplication de rétentions tertiaires portant non tant sur des parties du monde (un texte imprimé, un objet photographié, un concert enregistré) que sur les *gestes et habitudes attentionnelles* qui sous-tendent nos rétentions primaires (ce sur quoi je clique ou swipe, les corrélations entre mes gestes digitaux et ceux d'autres utilisatrices, le temps que je passe sur tel ou tel contenu, etc.). Les rétentions quaternaires résultent, d'autre part, de la capacité à traiter ces (méta)données en très grand nombre (big data) par des

¹ Voir Bernard Stiegler, *La technique et le temps* (1994), Paris, Fayard, 2018.

processus de computation basés sur « l'induction statistique », c'est-à-dire sur une interaction dynamique et bidirectionnelle entre des consignes de traitement structurant (de haut) un ensemble de données et des attractions de proximité observées (depuis le bas) entre les agrégats de données.

Les rétentions quaternaires forment une *quatrième mémoire* en laquelle consiste le traitement statistique des rétentions tertiaires par des dites « intelligences artificielles ». Ce traitement ne vise plus le rappel à l'identique de ce qui a été. La quatrième mémoire se nourrit du passé des rétentions pour les possibiliser dans son espace latent et pour pouvoir les régénérer : ce ne sont pas les mêmes rétentions indicielles qui reviennent encore et encore, mais ce sont des rétentions *ressemblantes*. C'est la ressemblance elle-même, entendue comme représentation mimétique, qui est automatisée, marquant une nouvelle étape dans le complexe processus d'industrialisation.

En effet, l'IA est nourrie par des grands ensembles de données (dataset) qui lui permettent de calculer un *espace latent*, lequel définit des probabilités selon une logique bayésienne qui constitue un espace statistique structuré par les proximités et attractions observables au sein d'un ensemble de données. Sans entrer dans les détails, développés dans le remarquable ouvrage d'Anna Longo², *Le jeu de l'induction*, il faut retenir que si les images générées avec l'aide des réseaux de neurones sont crédibles et réalistes, ce n'est pas seulement qu'elles contiennent les probabilités tirées de beaucoup d'images passées qui ont été accumulées en à peine trente ans sur le Web (dont c'était sans doute la fonction préparatoire

à la manière d'un forage et stockage extractiviste dans nos mémoires), mais encore qu'elles contiennent potentiellement toutes les images à venir. Et c'est pourquoi elles peuvent être en même temps différentes et réalistes. Le réalisme devient alors l'anticipation crédible dans un espace inductif d'une image possible.

Peut-on encore parler de mémoire, de rétention, d'archive, d'histoire ? Les rétentions secondaires et tertiaires, comme mémorisation des perceptions et comme action de garder par-devers soi ce qu'on devrait diffuser, ne sont pas à proprement parler périmées, mais elles se trouvent dépassées par les dynamiques propres aux rétentions quaternaires. Celles-ci capturent, fraient et façonnent l'avenir en opérant un bouclage automatisé du futur sur la présence du passé – en un mixte que l'on pourrait faire relever de la *prétention* (à entendre comme un saute-mouton de rétentions en protentions). Une inversion significative s'opère en effet dans de nombreux cas : la mise en circulation semble tendre à précéder ce qui est retenu parce que cette mise en circulation dépend d'une externalisation dans les datacenters, dont les présupposés infrastructurels d'ordre commercial déterminent ce qu'il y a à retenir et la constitution des rétentions. En fournissant aux réseaux sociaux ce pour quoi ils sont faits, nous produisons nos mémoires en les adaptant.

Cette dynamique prétentionnelle qui structure nos comportements par un va-et-vient instantané entre rétentions (ce qui se garde du passé) et protentions (ce qui nous projette vers un futur préformé) est en train de subir une transformation majeure, épopéale. Nous proposons la notion de *distention* pour désigner l'époque qui s'ouvre avec la possibilisation des rétentions traitées par les IA. Il s'agit par-là, non seulement d'augmenter le volume ou la surface d'un corps en le soumettant à une très forte

² Anna Longo, *Le jeu de l'induction. Automatisation de la connaissance et réflexion philosophique*, Sesto San Giovanni, Mimesis, 2022.

tension, mais aussi de relâcher les liens qui resserrent un tout ou qui unissent plusieurs choses. La distention est une extension parce qu'à partir de rétentions tertiaires, elle multiplie encore plus le nombre de documents en créant une rétention de rétentions, une attention à l'attention, une mémoire de mémoire, en une inflation de *méta*-données portant sur des données.

Le changement de nom de Google à Alphabet et de Facebook à Meta est à lire comme un symptôme de cette *méta-isation* catalysée par les prétentions quaternaires. La distention est une rétention génétiquement récursive, de sorte qu'elle n'est pas, comme la rétention stieglerienne classique, la répétition d'un événement : elle est la répétition de la prétention accouplée avec elle-même, l'automatisation de sa mise en boucle autoproductrice. Cette répétition singulière permet de comprendre de quelle façon la ressemblance mimétique est répétée et automatisée comme telle.

Arpenter la disfactualité de l'espace latent

Il serait naïf d'annoncer la fin pure et simple des rétentions (classiques). Elles continuent à être constituées, mais l'apparition d'une quatrième mémoire qui est distendue, qui métabolise les rétentions passées et qui, se prenant pour son propre objet, devient exponentielle, change l'orientation des rétentions et la définition même de la mémoire comme de l'expérience.

Un exemple : en août 2023, un producteur basé à Nancy, Lnkhey, publie sur SoundCloud et YouTube³ un remix où la voix clonée d'Angèle, grâce au logiciel libre Retrieval-based-Voice-Conversion⁴, chante un morceau qu'elle

n'a jamais interprétée. Plusieurs millions de personnes l'écoutent. Angèle réagit sur TikTok : « je ne sais pas quoi penser de l'intelligence artificielle, j'trouve c'est une dinguerie, mais en même temps j'ai peur pour mon métier mdrrrrr⁵ ». Sur cette vidéo, elle chante en playback ce remix puis fait une moue amusée, comme si elle était prise de vertige face à cette voix ressemblante qui n'est pas la sienne.

Autre exemple : un film des frères Lumières upscalé à une définition de 3840x2160 pixels et 60 images/seconde. Le film n'a pas été « restauré », parce qu'il ne s'agit pas d'un retour au film dans son état d'origine, mais d'une « instauration » parce que des éléments originellement absents ont été ajoutés. L'effet est saisissant : le film n'a plus le réalisme de 1895, mais le grain d'une vidéo tournée en 1971 avec un Portapac Sony. La *complétion*, c'est-à-dire le fait de compléter un document historique afin de le réparer, entraîne un réalisme anachronique qui modifie la nature de l'archive, laquelle n'est plus déterminée par une origine. La complétion invente une historicité qui n'existe pas à l'origine, parce qu'elle a été alimentée par des images prises entre 1895 et aujourd'hui. Le résultat n'émane pas de la rétention la plus fidèle possible des données à capter en 1895, mais d'un mix modulable de données et de statistiquement probable, déterminé par l'espace latent du dataset.

Le réalisme change de nature et devient *disfactuel*, le préfixe « dis- » signifiant ici la séparation, la différence, la cessation ou le défaut au sein même du factuel, c'est-à-dire des faits. Cette disfactualité touche en dernière instance à la facticité, c'est-à-dire à la contingence de la corrélation entre la pensée et les faits qu'elle

³ www.youtube.com/watch?v=EiV1YxtbfcE

⁴ <https://github.com/RVC-Project/Retrieval-based-Voice-Conversion-WebUI/blob/main/docs/fr/README.fr.md>

⁵ www.tiktok.com/@angele_vl/video/7265090543191936288

visé. Les images sont factices, mais cette facticité vient affecter la réalité en son ensemble, et c'est pourquoi elle est disfactuelle : *elle disloque quelque chose du dedans*. Ce faisant, elle corrode la factualité sur laquelle est basée notre confiance en notre pouvoir d'exercer une certaine maîtrise sur le monde.

Avec les intelligences artificielles, ce qu'on retient, ce qu'on traite, ce qu'on métabolise, ce sont toutes les formes passées de rétentions tertiaires une fois qu'elles ont été numérisées massivement sous forme binaire et rendues par là même intercompatibles, traitables, traduisibles. Cette période, on peut l'associer au Big Data, comme projet de numérisation de la culture, et au Web 2.0, comme participation de chacun à cette mémorisation. Cette période n'était en fait qu'un acte préparatoire à l'induction statistique. Nous voyons peut-être apparaître une nouvelle forme de réalisme, un réalisme de réalisme, qui permettrait de mieux comprendre la multiplication des vérités alternatives et contrefactuelles que la simple promesse d'une démarcation, de plus en plus difficile à tenir, entre vérité et fiction.

Cette nouvelle métabolisation peut s'analyser en *six étapes* : **1°** les rétentions quaternaires (qui enregistrent nos gestes attentionnels, nos réactions interprétatives, nos réélaborations créatives) se trouvent **2°** accumulées dans d'énormes bases de données, pour y être **3°** triées, calculées, associées, approximées par une puissance inédite de computation basée sur l'induction statistique, dont **4°** on fait émerger des espaces latents, à partir desquels **5°** une quatrième mémoire génère des préentions **6°** sous forme d'objets esthétiques (une nouvelle chanson d'Angèle) qui ne sont ni de « vrais » enregistrements, ni de « vraies » créations, mais des entités inédites – qu'on peine à qualifier (modèles, *proxies*, produits de synthèse, *deep fakes*, preuves de concepts ?).

L'aliénation de la crédibilité

L'espace latent est notre nouvel espace culturel, dont les produits sont disfactuels. La chanson d'Angèle existait avant d'exister réellement. Elle existait comme une statistique ou, selon les cas, un possible. Il a fallu qu'elle naisse à la réalité par cette reprise de reprise. C'est la signification ontologique du post déjà cité : « Quand l'IA d'Angèle sur Saiyan devient enfin réalité », où le « d' » qui sépare et relie l'IA à Angèle exprime cette prétérition du latent culturel. Tout existe avant d'exister. Il y a, dans cette étrange anticipation disfactuelle, un nouveau pacte complice avec le public. C'est Kaaris s'amusant de sa propre IA⁶, et ce n'est nullement là la propriété d'un fantasmagique remplacement : c'est la distance à soi, une étrangeté bien connue de la modernité, un décalage de notre appareillage. Grâce à l'accumulation du passé par les supports matériels de mémoires tertiaires, nous produisons quelque chose qui n'avait jamais eu lieu, mais qui ressemble étrangement à tout ce qui pourrait avoir lieu : Kaaris chantant l'Inspecteur Gadget ou un animé de Disney. Ce possible a déjà sa forme de réalité, mais toute l'intelligence culturelle de notre époque est dans cet amusement partagé entre les chanteuses et leurs publics, dans cette nouvelle répétition où nous interprétons ce possible encore non-effectué mais qui pourtant a déjà (eu) lieu en tel point de l'espace latent.

Si, jusqu'à présent, notre culture et son partage étaient déterminés par les mémoires tertiaires, fruits de la période industrielle, nous entrons assurément dans une nouvelle époque avec les mémoires quaternaires où le contrat esthétique pourrait être celui de *l'aliénation* : nous reproduisons des machines qui nous

⁶ https://x.com/booska_p/status/1689304038393683968?s=20

reproduisent. L'espace latent devient un espace des possibles qui contient le passé, mais aussi, sans doute, une partie de l'avenir et de l'incalculable. Car on pourrait bien, par exemple, faire une photographie avec un appareil quelconque et l'envoyer dans une IA pour vérifier qu'elle existe déjà et la trouver, la retrouver. Il ne s'agit plus seulement d'une numérisation qui rend discrètes sous forme de 0 et 1 (en les découpant par échantillonnage) des variations pouvant être recombinaisons à l'environnement (comme produits de synthèse). Il s'agit à présent de prétentions statistiques, qui distendent nos prétentions en informant ce qui n'existe pas encore selon la soif de profits commerciaux de plateformes exploitant leur accès privilégié à l'économie de nos attentions.

Ce réalisme très particulier – disfactuel – émerge d'une quatrième mémoire qui appartient au passé du spectateur et, en même temps, au futur des images instaurées. Elle n'est pas simplement un autre résultat de la causalité : elle entre un possible sur un donné, en même temps qu'elle hante le second par le premier, sapant les fondements de notre croyance indiciaire (si j'entends la voix d'Angèle, c'est qu'Angèle a dû chanter). Il est donc indispensable de replacer ce qui pourrait apparaître comme une simple innovation technologique – avec sa litanie de nouveautés, passant de GAN, Clip, Disco Diffusion, Zoetrope à Dall-E 2, Imagegen, Parti, etc. – dans le contexte général d'une incertitude envers la factualité où, selon certains sondages, 40 % des 18-24 ans aux États-Unis disent penser que la Terre est plate. Cette latence devrait bien sûr être liée au complottisme, aux fausses nouvelles, à cette étrange démocratisation expressive de l'opinion où tout le pensable semble devoir être pensé par quelqu'un, et où chacun semble ne penser que pour réagir à ce qu'il croit que l'autre pense dans une anticipation bayésienne sans fond.

Nous avons encore à nous orienter dans cette culture de l'espace latent et dans l'émotion paradoxale qui nous saisit lorsque nous écoutons, et réécoutons, la voix d'Angèle, puis que nous revenons à la voix de l'IA, que nous allons et venons entre les deux, ne sachant décider de notre émotion et du monde qui ainsi nous traverse. C'est un nouveau réalisme et une nouvelle historicité dont les nouvelles structures sont en train d'émerger – aliénant non tant nos identités que la crédibilité même de notre monde culturel.

Du possible disfactuel au réalisme contrefactuel

Face aux rengaines omniprésentes et étouffantes du remplacement par les IA, Angèle et son public jouent une autre partie que celle de la lamentation effarée. Les sentiments y sont mêlés. Il y a sans doute un peu de peur, d'étonnement. Mais il y a surtout un amusement dans le jeu infini des simulacres et des ressemblances – autre nom pour désigner la culture – ce que n'auront jamais compris ni les pasteurs technocritiques ni les prêtres humanistes. L'IA n'y est pas pensée d'avance, comme s'il suffisait de la réfléchir correctement pour fixer la manière dont il faut la réformer, l'encadrer, la mettre dans une législation ou dans un tuyau, avec une entrée et une sortie, des embranchements, toute une logistique qui est finalement un logos et qui aura toujours un coup de retard. Les IA – à lire ici comme nos Intelligences Aliénées – y sont expérimentées : nous les aliénons et elles nous aliènent⁷. Dans ce cas, elles ont bien appris à

⁷ Dans un livre important consacré au buzz à l'ère des *Large Language Models*, Dominique Quessada suggère de réécrire les IA en *IAI*, pour Intelligence *alien*, « une intelligence qui vient du silicium » : « pour la première fois une intelligence émerge d'elle-même à partir d'un autre support que le carbone » (*Parasite. Essai sur le bruit digital*, Paris, PUF, 2023, p. 166).

chanter comme Angèle et cette dernière leur a en quelque sorte répondu en reprenant « leur » chanson commune (affolant du coup les grands prêtres des droits d'auteur). Nous avons été les secrets témoins de cet écho sismique. Nous pouvons être les explorateurs et les agents (plus ou moins secrets) d'aliénations expérimentales.

Après l'apogée de l'accumulation hypermnésique des supports de mémoire par leur numérisation et leur enregistrement dans les centres de données – ultime étape de la reproductibilité benjaminienne – notre époque industrialise, avec les IA, la ressemblance elle-même par le possible. C'est sans doute la raison pour laquelle les IA – ces questions qui traversent et bouleversent tant de domaines de l'activité humaine – ont été si fréquemment abordées dans les médias et auprès du grand public par « la question de l'art ». Cette dernière concentre en effet symboliquement dans la modernité le propre de l'humanité, ainsi que le mystère de son intériorité qui, nous le savons, fut le procès d'une construction de la subjectivité en Occident, allant jusqu'à la volonté de puissance et le nihilisme.

Dans un autre post de Tik Tok, on peut lire « La boucle est bouclée ». Ce n'est pas seulement que nous apprenons aux IA à créer des images, des textes et des sons qui nous ressemblent : c'est que nous leur ressemblons et que, par rapport aux discours réactionnaires, nous ne désirons rien d'autre que d'activement nous aliéner. Nous ne croyons ni dans le fait de rendre lisibles les IA par la transparence du code, ni dans l'acte de couper et de nous séparer de ces flux pour regagner une imaginaire autonomie et souveraineté fantasmée. Nous voulons expérimenter que ce que nous croyons exister est aussi un produit de la technique et de sa reproduction paradoxale. Nous en sommes la reprise. En métabolisant toute l'histoire de nos

supports de mémoire, les IA, entendues comme « nos » Intelligences Aliénées – et ce « nos » très problématique mérite sans doute de scinder les humains en dominants et dominés, en blancs et non-blanches – sont en train de constituer une nouvelle mémoire, où le passé et le futur ne sont plus chronologiques, mais semblent se répondre l'un à l'autre en échangeant leur rôle.

Ce sont ici les ruminations prophétiques de Brian Massumi sur les différences entre le possible, le probable, le virtuel et le potentiel qu'il faut remobiliser pour prendre la mesure de ce qui nous arrive⁸. L'induction statistique joue avec un possible (partiellement) maîtrisé par les probabilités. La chanson d'Angèle-qu'Angèle-n'a-pas-chantée mobilise le probable pour réaliser d'autres possibles. Sa disfactualité s'inscrit toutefois parfaitement dans les protentions de l'esthétique dominante (sinon dans celle d'un commerce musical faisant un hommage hypocrite au sacro-saint « droit d'auteur »). Seule l'expérimentation – toujours vertigineuse – avec nos aliénations peut espérer tirer du simple disfactuel le *potentiel* de transformation inhérent à la contrefactualité. Le défi des images, des sons et des textes à l'heure des IA génératives n'est pas tant d'être « originaux », « nouveaux », « vrais », « innovants » ou « beaux » (autant de termes et de valeurs qui ont pris un méchant coup de vieux en quelques années seulement). Il est bien plutôt d'être *contrefactuels* : d'arpenter (artisanalement) l'espace latent pour en manifester (automatiquement) des rétentions crédibles de réalités non-advenues parce que contraires aux faits des dominations en place.

Incarnar ces aspirations latentes envers des faits contredisant les protentions de l'ordre régnant : n'est-ce pas ce qu'on appelait

⁸ Brian Massumi, « On the Superiority of the Analogue », *Parables for the Virtual*, Durham, Duke University Press, 2002.

« révolution » au cours du siècle dernier ? Pour reprendre une distinction sur laquelle insiste aujourd'hui Pierre-Damien Huygue, les usages expérimentaux, non-instrumentaux, « artistiques », des IA génératives ne relèvent pas tant de l'*agir* politique (*prattein*) que du *faire* artistico-artisanal (*poiein*)⁹. Non pas tant « faire la révolution » (au sens d'agir pour qu'une révolution advienne) que *fabriquer des objets contrefactuels*, qui fassent voir, entendre, penser avec la force du réalisme de quels contre-mondes nos sociétés sont porteuses. Que les *deep fakes* nous fassent craindre un monde de « post-vérité » – pendant politique du fantasme économique du remplacement de l'humain

par la machine – témoigne certainement d'un problème bien réel : il est essentiel de pouvoir préserver un certain rapport social de confiance envers notre accès à la factualité. Pas d'information possible sans crédibilité. Mais nos angoisses et nos croisades anti-conspirationnistes témoignent tout autant de l'incapacité des forces progressistes à comprendre le potentiel politique du réalisme contrefactuel – au moment où les forces réactionnaires exploitent sans vergogne les ressorts de l'irréalisme disfactuel. Quoique souvent oiseux, nos débats sur les IA n'auront pas été vains s'ils nous aident à identifier – et à investir – ce (pas si) nouveau terrain de luttes.

⁹ Pierre-Damien Huyghe, « Qu'est-ce que faire dans l'urgence? », enregistrement du séminaire 2022-2024, <http://pierredamienhuyghe.fr/recherches.html#urgence>