

samedi 6 mai 2023

NUMERIQUE

Grégory Chatonsky : « IA : comprenons ce qui nous arrive plutôt que de le juger d'avance »

Par Benjamin Tainturier

SOCIOLOGUE

Comment penser l'intelligence artificielle et sortir à la fois des discours technosolutionnistes et technocritiques ? La philosophie tout comme l'art contemporain nous équipent bien, à condition d'accepter non pas de dominer la technique, de l'arraisonner à notre tour, mais d'adopter une posture humble, expérimentale, propice à la découverte des fascinants mondes latents que les réseaux de neurones induisent. Pourtant, au seuil de cette nouvelle Révolution, il ne faudrait pas perdre de vue que l'impasse écologique et l'intelligence artificielle sont intimement liées, et tracent de concert les voies de notre futur.

On a parlé de GPT3, de DALL-E, de ChatGPT, de GPT4, en attendant GPT5, 6, et quoi d'autre encore... ? L'intelligence artificielle (IA), terme largement impropre pour définir le développement technologique dont nous sommes les contemporains, accouche de toujours plus d'algorithmes, aujourd'hui capables de produire des images d'art, des textes de littérature, des scénarios, ou de réussir des concours d'entrée très sélectifs...

Comme à chaque fois qu'un tel surgissement se produit, et que de nouveaux possibles font irruption dans un monde trop peu normé pour les accueillir, nous traversons une période d'inflation de discours et de contre-discours ; les thuriféraires se pâment, les Cassandres se trouvent de nouvelles scènes...

Les premiers ne boudent pas leur plaisir à parler de bouleversement civilisationnel, les autres se rappellent qu'on n'aurait « jamais employé tant d'esprit à vouloir nous rendre Bêtes[1] ».

Dans cette période pauvre en balises et repères nous pouvons néanmoins compter sur les lumières d'expérimentateurs généreux, de fins connaisseurs, d'intellectuels, dont Grégory Chatonsky, fait partie. Artiste franco-canadien, pionnier du Netart en 1994, Grégory Chatonsky travaille depuis 2009 sur l'induction statistique, une des façons d'appréhender plus précisément l'intelligence artificielle. Il a exposé au Palais de Tokyo, au Centre Pompidou, a publié le premier roman en langue française co-écrit avec une intelligence artificielle, et a enseigné au Fresnoy, à Paris VIII, à l'École normale supérieure... Au centre de son approche, se trouve le concept d'espace latent : cet espace vectoriel très abstrait et bien connu des professionnels de l'IA est induit par la totalité des images, des textes, des informations que nous donnons à un réseau de neurones pour l'entraîner. Toutes ces données, créditées de coordonnées, se positionnent les unes par rapport aux autres dans un espace latent, continu, dans lequel il est possible de circuler pour découvrir de nouvelles formes intermédiaires, synthétiques... et autant de surprises ! À travers les collections de données qui nourrissent les algorithmes, et qui sont autant d'archives de nos vies, l'intelligence artificielle nous donne à contempler notre propre histoire, nos futurs possibles comme nos passés inadvenus. **BT**

Il y a quelques semaines, plusieurs chercheurs, entrepreneurs ou spécialistes de l'intelligence artificielle (IA) ont diffusé une pétition réclamant un moratoire sur la recherche fondamentale dans le domaine. Il faudrait stopper, pendant six mois, les progrès, le temps d'une discussion collective pour prendre la mesure de ce qui nous arrive et des risques encourus. La présence d'Elon Musk, le directeur général de Tesla Motors, dans la liste des signataires, a été particulièrement remarquée. Que pensez-vous de cette initiative ? Elle fait écho à un texte récent où vous regrettiez notre naïveté collective et notre confiance aveugle, loin de la prudence des débats des années 50 sur les nouvelles technologies. Finalement, ce moratoire ne rétablirait-il pas une saine inquiétude ?

Je n'ai pas de certitudes, mais voici l'état de ma réflexion à ce stade. Ce n'est pas la première fois qu'une demande de ralentissement se manifeste. Je crois qu'en 2017, il y avait déjà eu une pétition par rapport à l'intelligence artificielle, pour proposer de ralentir son développement et d'y réfléchir. Cela me suggère trois remarques. La première touche à l'ambivalence de ces deux pétitions : ceux qui demandent le ralentissement sont aussi des acteurs de l'intelligence artificielle, qui utilisent cet argument de manière ambiguë. On parle de personnes ou d'entreprises comme Elon Musk, Microsoft, Facebook... Il faut s'interroger sur leurs motivations car il pourrait bien s'agir là d'attirer l'attention sur l'intelligence artificielle, même négativement ; de montrer que les débats font rage, que le sujet doit se placer au centre de la société... Ce qui est une manière détournée de promouvoir l'IA.

Cette ambivalence rejoint un concept que j'avais utilisé au milieu des années 90 et qui me sert encore de grille de lecture de l'idéologie sociale. C'est un concept hérité du Jacques Derrida de *Spectres de Marx*, celui « d'enthousiasme conjuratoire », qui joue sur le double sens de « conjuration ». L'enthousiasme conjuratoire qualifie l'ambivalence d'un discours de conjuration, qui voudrait que les choses n'adviennent pas, qu'elles soient conjurées, et qui est le discours d'une conjuration – une conjuration désignant aussi une forme, non pas de complot, mais d'organisation – dédié à ce que quelque chose advienne en effet. Appliqué aux technologies, l'enthousiasme conjuratoire est le fait pour les technocritiques et les technosolutionnistes d'utiliser les mêmes structures, réversibles, comme deux facettes d'un même affect. Rappelez-vous du discours sur la réalité virtuelle dans les années 1990 : il y avait les Pierre Lévy qui étaient enthousiastes, et puis les conjuratoires, Virilio ou Baudrillard. Mais en même temps, et malgré les oppositions apparentes, il y avait une espèce de fascination commune aux deux groupes, car les technologies sont surtout une surface de projection avant d'être une entité autonome, qui existe en soi. À cette époque, de grandes discussions et déclarations tonitruantes sur la réalité drainaient d'énormes concepts métaphysiques : le *Gestalt*, l'arrondissement... Il me semble, de la même manière, que ces pétitions sur l'IA relèvent de l'enthousiasme conjuratoire. Pour comprendre mon troisième point de critique, il faut en revenir au texte

du moratoire, qui est catastrophique : l'IA remettrait en cause la vie sur Terre. Le texte donne l'impression de parler moins de l'intelligence artificielle que de l'extinction anthropocénique qui est actuellement à l'œuvre. Une vision très courante, mais qui ne va pourtant pas du tout de soi, postule que pour trouver une solution et prévenir un danger, il faut penser avant d'agir, dévalorisant l'action et valorisant la pensée : il faut se mettre en conclave, réfléchir, décider de ce qu'il faut faire et agir ensuite conformément à ce qu'on a décidé... Ce qui me semble très paradoxal. Cela rejoue l'histoire de la volonté de puissance en Occident et l'idée que la pensée compte en priorité, qu'il suffit de décider pour faire et que la réalité se réforme. On recycle ici de vieilles structures inspirées de l'idéalisme. En tant qu'artiste, je trouve qu'il y a d'autres façons de faire et je me situe beaucoup plus proche de personnes comme Yann Le Cun, et de sa réponse à cette pétition. Son argumentation, si je la déplace sur le plan artistique, fait la part belle à l'expérimentation, à l'expérimental : on peut penser en agissant et penser certains modes d'action comme des modes de pensée. Le défaut essentiel de la recherche sur l'IA n'est pas d'aller trop vite, mais pas assez vite : de nouveaux logiciels, de nouveaux articles, des codes sources sortent tous les jours et nous n'avons pas le temps de nous pencher sur toutes ces nouveautés, de les expérimenter, de les essayer pour en faire autre chose. Pour penser, il ne faut pas suspendre l'action.

Vous avez une formation de philosophe : comment faire pour penser ces nouvelles technologies sans que la réflexion ne soit polluée par cette agitation de concepts que vous évoquiez ? Vous mentionnez Derrida, il est vrai que l'hantologie est très présente dans votre propre analyse de l'IA ; la métaphysique heideggérienne et la notion d'historicité du *Dasein* entrent également bien en résonance avec ces collections latentes d'images possibles que l'IA permet de créer. Comment mobiliser ces abondantes références et dépasser la simple alternative entre technocritique et technosolutionnisme ?

Je travaille depuis 2009 sur l'induction statistique – terme que je préférerais ici à « intelligence artificielle », et je l'ai mis au centre de ma pratique depuis 2014, car, pour moi, ce dont on parle ne se réduit pas à une technique, mais ouvre un monde qui charrie aussi une manière de concevoir les rapports sociaux, les choses dans un espace bayésien...

ChatGPT a vraiment déclenché une discussion globale qui affecte le corps social et qui ne relève pas seulement d'un effet de mode – ce que j'ai rarement vu, à ce point, dans le cas d'une technologie. Mais cette discussion mobilise à tort et à travers des concepts extrêmement compliqués, très ambivalents d'un point de vue historique. Beaucoup de ceux qui s'expriment sur le sujet n'ont pas eu le temps requis pour faire les bonnes lectures, pour bien assimiler ces concepts extrêmement difficiles, et ne comprennent pas en profondeur les techniques sur lesquelles ils s'expriment. Cela demande un vrai temps de travail pour assimiler que l'intelligence artificielle mobilise l'ensemble de l'histoire occidentale, dans son destin. Il ne s'agit pas seulement de comprendre la technique mais de comprendre aussi notre histoire, ce qui nous est extrêmement complexe puisque cette histoire embrasse le réchauffement climatique. Il me semble en effet que les deux lignées historiques, IA et réchauffement climatique, sont totalement liées.

Je suis très frappé de constater qu'en France, entre les années 40 et 60, on parlait beaucoup de cybernétique : de nombreuses revues s'en saisissait, Lacan a écrit un texte absolument insensé sur la cybernétique... Les textes de Norbert Wiener, d'Alan Turing, sont d'une richesse incroyable et d'une ambiguïté fondamentale. Il faut revenir au test de Turing, que nous n'avons pas bien compris. Turing explique que la pensée n'existe pas, que c'est un bourdonnement dans la tête ; les deux versions du test de Turing, sont des textes sur l'intelligence en tant que simulacre. Ensuite la cybernétique a perdu quelque chose au moment de la seconde cybernétique de Marvin Minsky, et de l'approche générative chomskyenne : on a abordé l'intelligence sous forme de modélisation hypothético-déductive des compétences, disons, ce qui a complètement échoué... On assiste aujourd'hui à un retour de la première cybernétique par l'intermédiaire des systèmes d'induction statistique. L'autre point important consiste à sortir de la fausse alternative entre technosolutionnisme et technocriticisme qui n'a aucun intérêt à mon avis : comprenons ce qui nous arrive plutôt que de le juger d'avance.

Du point de vue des références plus larges, il faut bien admettre que Heidegger a donné une dimension à la question de la technique comme peu d'autres avant lui dans l'histoire de la pensée. J'essaie de poursuivre ce

chemin, qui n'est pas le chemin propre de Heidegger, mais celui d'autres après lui comme Derrida, comme le premier Bernard Stiegler, celui des deux premiers tomes de *La technique et le temps*. Il me semble que cette tradition, très française en définitive, Leroi-Gourhan, Simondon, Stiegler, Derrida, Lyotard aussi... celle de la déconstruction, a fourni des outils absolument pertinents et originaux quand on les applique à la question de la technique. Nous avons besoin de la penser de manière contre-intuitive, sans quoi on risque non pas de la penser mais d'être agis par des mots qui nous dépassent complètement et dont on ne comprend pas le sens. Il faut sortir de cette espèce de pensée agitée, éditorialiste où il faut prendre des décisions, parler très fort et prendre des décisions alors qu'on n'a pas pleinement compris de quoi on parlait.

Comment votre approche spécifique de l'IA par l'espace latent se traduit-elle dans vos projets artistiques actuels ? Je pense notamment à « Un Été au Havre », dont le lancement aura lieu le 24 juin ?

C'est un projet assez étendu, sur trois ans, avec l'idée que la ville du Havre devienne l'espace latent de la ville du Havre, pour concrétiser un Havre alternatif. J'ai travaillé sur des images d'archive, afin de revisiter l'histoire de la ville et de proposer une nouvelle vision de la Révolution industrielle, que la ville aurait traversée et pendant laquelle la technique et la nature se seraient totalement mélangées. Ce projet, qui s'est fait main dans la main avec le bailleur social Alcéane, va aboutir à l'accrochage de grandes images, presque comme des fresques, produites avec des IA sur [des façades d'immeubles](#), plutôt sur les quartiers périphériques, des images de dix mètres de haut, mises en place courant juin, pour une ouverture le 24. Une seconde facette du projet raconte l'histoire des années 50, 60 et 70 au Havre par l'intermédiaire d'environ 25 000 [cartes postales uniques](#). J'ai commencé à découvrir dans l'espace latent, une sorte de fiction : après la reconstruction de la ville, l'architecte Auguste Perret aurait fait un projet d'utopie politique, fondé sur la paresse, sur des activités de lecture dans les parcs, de siestes... Dans cette utopie, la vie sociale s'organise autour de sculptures violettes qui émergent dans l'océan et que les jeunes vont pêcher et sortent de l'eau. Ces formes violettes qui sont apparues dans l'espace latent vont trouver une traduction physique, sous forme de sculptures, et se matérialiser, réellement, au Havre. Pour aboutir à ces images et ces cartes

postales, je n'ai pas commandé l'espace latent, je m'y suis déplacé grâce aux *prompts*[2], en tapant dedans grâce au langage ; puis j'ai vu apparaître ces formes violettes. Le troisième volet du projet prend place dans les bâtiments publics et reprend une esthétique « art pompier », avec une référence à Robert Hubert et à ses vues imaginaires du Louvre : les moments révolutionnaires sont aussi des moments de représentation des ruines. Je présente des environnements administratifs **en ruines**, dans lesquels on trouve toujours ces formes violettes étranges. Le projet « Un Été au Havre », du point de vue global, démarre donc par l'image, puis se matérialise par des formes physiques, qui donneront lieu à des rituels, à des performances, pour donner corps à l'espace latent dans lequel on avait au départ pioché les images. On fait rentrer un espace contrefactuel dans la réalité, en s'intéressant à ce qui n'existe pas, en s'y aliénant, ce qui mène à l'émancipation de ce qui existe.

Je travaille aussi en Arabie Saoudite, dans le cadre d'une résidence à Al-'Ula, un endroit qui m'intéresse notamment pour interroger nos propres préjugés occidentaux. Les projets sont liés à la manière d'aborder l'héritage préislamique qui est lacunaire, abîmé par le temps et parfois détruit par l'Islam, pour arriver à reconstruire un passé contrefactuel qui n'a pas existé grâce à l'IA. L'autre idée consiste à générer de nouvelles formes à partir de traditions artisanales à l'œuvre en Arabie Saoudite en créant des séries infinies d'œuvres uniques. Cela vise à décentrer le regard, à décoloniser les techniques en m'intéressant à un lieu politique en train d'inventer une modernité non occidentale et peut-être, pour reprendre ce concept de Yuk Hui, une « cosmotechnique ».

Pourquoi la discussion publique fétichise-t-elle tant l'IA, la recouvrant de multiples affects : tantôt l'émerveillement, tantôt la crainte... ? Il semble qu'on ait ici affaire à autre chose qu'une simple panique morale ? Par exemple, plutôt que d'engager une discussion de fond, on aime à se focaliser sur des faits divers, comme l'histoire tragique de ce chercheur Belge s'étant donné la mort après une conversation avec une intelligence artificielle.

Oui mais ces histoires sont anecdotiques... On en connaît déjà d'autres, en grand nombre. Je pense à ce couple de *gamers* en Corée, qui avaient laissé mourir de faim leur propre enfant parce qu'un jeu vidéo les happait... La

difficulté aujourd'hui est de conceptualiser plus généralement ce qui nous arrive sans tomber dans la collection d'anecdotes. Cette tâche se trouve d'autant plus compliquée qu'elle mobilise l'ensemble de notre histoire. ChatGPT a vraiment fonctionné comme une espèce de révélateur pour la totalité de notre corps social ; l'aspect conversationnel, il me semble, a focalisé le plus d'attention, car les *large language models* existaient déjà avant ChatGPT, qui les implémente spécialement dans une formule conversationnelle. Cette nouveauté s'est vue particulièrement recouverte d'affects et ChatGPT passer pour un « oracle », car les réseaux de neurones prévoient des séquences.

L'un des aspects les plus frappants, quand on pense aux chocs produits sur le corps social, a été ce renvoi de la notion d'intelligence vers nous-mêmes ; nous nous sommes attribués un rôle de juge. Finalement, ChatGPT a servi d'embranchement au modèle foucaldien de pouvoir et de savoir : quand je discute avec ChatGPT, je présuppose mon intelligence, et je l'attribue ou non à la machine – c'est un test de Turing au fond. Bien sûr, je vais désattribuer l'intelligence à Chat GPT pour me l'attribuer en sous-main à moi-même. ChatGPT permet d'allouer une intelligence aux êtres humains au moment où ils deviennent de plus en plus idiots. Tous les faux tests que l'on voit publiés dans des revues, où l'on soumet ChatGPT à des essais éclairs pour montrer combien cette technologie est bête mériteraient d'être comparés à des énoncés humains trouvés sur Twitter où on lit des choses bien plus idiotes... Cette vision de ChatGPT comme une espèce d'enfant qu'il faudrait former, incapable de parler comme un adulte, s'est imposée. L'une des racines du problème tient au fait que l'intelligence résulte d'un phénomène d'attribution : on discute de l'attribuer ou non à des animaux surtout pour nous en réserver un certain monopole indiscutable. Car nous sommes juges et partis en matière d'intelligence. En creux se déploie toute la question coloniale, celle de l'altérité, au regard du pouvoir d'attribution ou non de l'intelligence qu'ont les dominants. Nous conservons comme coordonnées du problème une conception instrumentale de la technique comme prolongement de la main, soumise à notre volonté. Ce produit historique ne va pas du tout de soi, précisément parce que nous faisons la technique autant qu'elle nous fait. Il faudrait radicalement changer cette conception instrumentale et utilitariste de la technique, qui conduit à ce que

j'appelle un « technosadisme » : tous ces tests faits sur ChatGPT pour montrer combien l'outil est stupide me rappellent cette vidéo de Boston Dynamics où on voit un ingénieur frapper un « *cyberdog* », un robot que l'entreprise produit, pour voir le robot, déséquilibré, se remettre à marcher. Le technosadisme qualifie cette insistance sur la bêtise de la machine pour nous convaincre, en regard, de notre intelligence. Turing avait déjà été conscient de ce risque de sadisme, du risque de se donner soi-même comme modèle, en vertu de cette relation entre savoir et pouvoir. Le test de Turing contourne le technosadisme, justement en refusant de dire qui est humain et qui ne l'est pas.

Pour bien comprendre néanmoins que tout n'est pas technologique dans la technologie, et que ces affects que l'on projette expliquent aussi une partie de ce à quoi nous faisons désormais face, il faut peut-être évoquer la série des *Contrefaits*, un ensemble de pièces qui prennent pour point de départ des tableaux très célèbres et les détournent. Malgré l'effet déformant de l'IA sur ces œuvres, elles demeurent tout à fait reconnaissables. Qu'est-ce qui est en jeu dans ce travail ?

C'est un projet déjà qui date un peu, fait avec les premiers systèmes de *prompts*, sur les questions de *mimésis* et d'air de famille. J'ai choisi sur un site internet les cinquante œuvres les plus connues de l'histoire occidentale, puis en ai passé des images dans une IA qui les décrivait au moyen d'un texte : c'est une première traduction de l'image en mots. J'ai ensuite fait de ce texte un *prompt*, injecté dans une autre IA pour aboutir à une nouvelle image. Étonnement, cette image finale garde une ressemblance avec la première. Qu'est-ce que la ressemblance ? Elle incorpore des différences, sinon elle n'est plus ressemblance mais copie, répétition à l'identique. La ressemblance tient à la fois de la différence et de la répétition. Ces ressemblances entre les images suggèrent très puissamment que l'intelligence artificielle, ou les systèmes d'induction statistique, prennent appui sur de la culture humaine. L'air de famille entre les images en entrée et en sortie du processus est lié au fait que l'opération de traduction se fonde sur un *dataset* contenant le patrimoine commun de l'espèce humaine et les données qu'elle a accumulées. Ce n'est pas avec la machine que nous discutons, c'est avec un espace latent qui est constitué de cultures humaines métamorphosées. Par cet intermédiaire, le passé peut se

poursuivre dans le présent et dans le futur.

Le concept de « tra(ns)duction », que j'hérite de Simondon, fonctionne ici à plein : à partir du moment où nous avons tout numérisé, les textes, les images, les volumes, et tout traduit en 0 et 1, alors toutes les productions numériques vont garder un air de famille, puisque nous pouvons tout traduire en tout. Je pense que la culture et l'art vont être, dans les années qui viennent, plongés dans cet immense air de famille. Voilà un gros défi pour les artistes du futur : naviguer dans cet air de famille généralisé, transformé en statistiques, dans lequel on peut anticiper des données. Tous les systèmes de tra(ns)duction, qui pointent donc vers la traduction, vont être des images à venir de la culture telle qu'elle s'est passée. C'est quelque chose qui se fait à la suite du *pop art*, et qui propose une forme d'esthétique qui est peut-être la seule à la hauteur des trois décennies pendant lesquelles nous avons archivées des données personnelles jusqu'à transformer la notion même d'histoire et la possibilité même de faire une histoire au regard de cette quantité de données que nous avons, qui dépasse nos capacités de préhension et d'analyse.

En creux, c'est bien la question du style qui se pose. S'extasier lorsqu'une IA génératrice d'images, comme DALL-E de l'entreprise OpenAI ou comme les outils proposés par Midjourney, est capable de créer une représentation des « extraterrestres de Saturne dans le style de Rembrandt » c'est bien déconnecter le style d'un artiste, Rembrandt, des objets figurés, ici les extraterrestres. Pourtant ce style n'existe pas *in abstracto*, il ne se manifeste jamais qu'incarné dans des sujets, peints et représentés dans les tableaux. Nos usages de l'intelligence artificielle prennent largement appui sur notre méconnaissance ou notre incompréhension de l'histoire de l'art ?

Je voudrais nuancer. Quand des experts discutent pour savoir s'il faut attribuer un nouveau travail à Rembrandt, ils doivent bien distinguer le style du peintre de son objet ! Finalement le *prompting* pose des questions analogues à l'attribution. Je suis consterné par la naïveté des productions très populaires que l'on obtient avec des outils comme ceux de Midjourney, par leur style reconnaissable, leur manque de pensée...L'attribution d'un style est le degré le plus simple de la ressemblance : au fond, pour faire un arbre dans le style de Rembrandt, on plie l'un sur l'autre deux paramètres de

l'espace latent, le paramètre « style de Rembrandt » et la figure vectorielle de l'arbre. Il y a tellement mieux à faire... On peut créer des choses singulières et qui ne seraient pas seulement des répétitions à condition de savoir se déplacer dans l'espace latent, et nous manquons d'expérimentations, pour l'apprendre et dépasser une vision instrumentale de la technique.

Écrire un *prompt* et retrouver une ressemblance entre l'image produite et le texte n'a rien de surprenant ! Dans les *Contrefaits*, le résultat m'a au contraire profondément surpris ! Les outils de Midjourney utilisés comme vous le décrivez n'ont rien d'intéressant, ils réalisent simplement une ambition télépathique : on voudrait que la machine puisse lire ce qui se trame dans notre tête et pas plus ; ne surtout pas être surpris. L'exploration de l'espace latent se situe à l'opposé de la vision expérimentale qui veut dominer la technique, la rendre fidèle à ce que nous voulons d'elle plutôt que de l'expérimenter. Cette conception me semble extrêmement dangereuse, et finalement assez proche du rapport dominateur que nous entretenons avec l'environnement et qui produit le réchauffement climatique.

Est-ce à dire qu'un des futurs possibles de l'IA serait la « désinvention » ? **Devant l'impact que ces technologies ont sur les ressources en métaux rares, devant leur consommation bien trop importante d'énergie, faudra-t-il tout abandonner, l'entraînement des réseaux de neurones et le stockage des données nécessaires à cet entraînement n'étant plus des activités soutenables ?**

Je ne crois pas à la décision concertée de suspendre une activité alors qu'on pourrait la prolonger, seulement au prétexte que cette activité ne serait pas raisonnable. Cela supposerait des institutions internationales qui permettraient une coordination – tout ce qui nous fait défaut aujourd'hui. En revanche, l'idée que dans les cinq à dix prochaines années, les composants et matériaux qui nous permettent de calculer, de faire des gros entraînements sur des réseaux de neurones, deviennent vraiment rares et mêmes indisponibles, est juste. Elle risque seulement d'entraîner une concentration capitaliste des ressources et des cartes graphiques qui permettent de calculer en masse au niveau des classes vectorielles[3]. Tout de même, il faut aussi réfléchir au type d'usage que l'on fait de l'intelligence artificielle : la

majorité des scénarios climatiques qui sont synthétisés dans les rapports du GIEC utilisent de l'induction statistique. C'est-à-dire que nous connaissons l'état de notre planète parce qu'il existe des réseaux récurrents de neurones pour le calculer. S'il nous faut ces technologies, ce n'est pas pour entretenir des gadgets comme DALL-E, mais plutôt pour calculer des scénarios climatiques et calculer l'attribution des ressources.

Quel rôle joue le *prompting* – savoir formuler des requêtes à une IA – dans notre manipulation de l'IA ? Cette méthode concentre beaucoup de débats notamment autour de l'interdiction possible de ChatGPT dans certaines universités. Pour les partisans de cette technologie, il faut apprendre le *prompting*, comme une compétence plutôt que d'interdire ChatGPT et les algorithmes les plus puissants fondés sur des réseaux de neurones. Qu'y a-t-il à apprendre ?

Je suis totalement opposé à l'interdiction, qui me semble une erreur complète. L'espace latent va devenir l'un des concepts majeurs au point de vue culturel. C'est un espace abstrait, très difficile à appréhender, d'autant que les tentatives de visualisation me paraissent actuellement inopérantes. Le *prompt*, en revanche, est un organe satisfaisant dont nous disposons pour tâter l'espace latent, un espace synthétique de la culture humaine. Il faut passer du temps en compagnie de cet espace : on ne l'appréhende pas en trente secondes, pas plus qu'on ne commande l'IA en un clin d'œil. Il n'est de toute façon pas question de commandement ici, il faut s'abstraire de ce rapport de domination au profit d'un rapport d'expérimentation. L'agir, l'acte, doit devenir une expérience, une sensation. Agir amène à faire l'expérience, et à dépasser ce dont on avait l'intention. Je propose, pour synthétiser tout cela, d'aliéner et de s'aliéner à l'IA : en apprenant quelque chose à une IA, on la change ; et tout en en faisant l'expérience on se change aussi. Même, je propose de parler d'infra-aliénation, au même titre que Marcel Duchamp parlait d'infra-mince. Ce processus de double influence m'a particulièrement touché au printemps 2020, pendant l'écriture de *Internes* : à ma connaissance le premier roman en langue française co-écrit avec une IA – GPT2, à l'époque. L'écriture a été assistée par un logiciel de complétion, grâce auquel le réseau de neurones finissait des phrases que je commençais. Parfois, j'assistais à des morceaux de bravoure littéraire incroyables. Le réseau apprenait en fonction de ce que j'écrivais ;

moi-même, ayant commencé à mieux appréhender ce que l'IA allait me suggérer, je l'anticipais. À un certain point, je ne savais plus qui avait écrit quoi. De cette expérience résulte une production hybride, d'un auteur qui est un humain transformé et une machine transformée, qui se sont mutuellement aliénés dans ce processus d'infra-aliénation. Il me semble que l'enseignement supérieur devrait s'engouffrer dans cette voie ouverte, pour honorer une vraie politique d'auteurs d'intelligence artificielle. Il se passe un raz-de-marée qui menace de nous balayer si on ne comprend pas qu'il faut refonder notre rapport au monde et à la culture. Il va falloir former des auteurs capables de naviguer dans l'espace latent de manière singulière, tout en restant des auteurs. On est encore loin du compte.

NDLR : Grégory Chatonsky a publié en août 2022, Internes aux éditions RRose, le premier roman francophone co-écrit de part en part avec une intelligence artificielle.

Benjamin Tainturier

SOCIOLOGUE, DOCTORANT AU MEDIALAB DE SCIENCESPO