Ἐπέκεινα
International Journal of Ontology
History and Critics

Jim Schrub

Ces machines qui *nous* surprennent
Une défense simondonienne de la différence entre la créativité des vivants et la créativité induite des machines depuis la notion de métabolisme

Epekeina, vol. 16, n. 1-2 (2023), pp. 1-43
Miscellanea

ISSN: 2281-3209
DOI: 10.7408/epkn.

Published on-line by:
CRF – Centro Internazionale per la Ricerca Filosofica
Palermo (Italy)
www.ricercafilosofica.it/epekeina

This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivs 3.0 Unported License.

Ces machines qui *nous* surprennent
Une défense simondonienne de la différence entre la créativité des vivants et la créativité induite des machines depuis la notion de métabolisme

Jim Schrub

Paratexte

Je tiens d'abord à remercier tou·te·s les participant·e·s de la 7e édition des passages philosophiques, ainsi qu'Elie During, Séverine Guillet et Cléo Collomb qui ont su nourrir les réflexions développées dans cet article.

Alors que je m'apprete a proposer un developpement relatif au rap- port entre IA, societe et creativite, je tiens a mentionner que Chat-GPT coute sept-cent mille dollars par jour, et que son entrainement aurait requis l'emission de 552 tonnes de CO2 . Son bilan carbone de fonctionnement reste, a ma connaissance, non-evalue. Entrainer Chat-GPT, le jouet de la compagnie d'Elon Musk, Sam Altman et Peter Thiel, a donc emis autant de CO2 que les 22 millions d'habitants du Mali en ont emis en un an. Le 28e rapport du GIEC est pourtant sans appel: chaque tonne de CO2 non-emise compte. Le deploiement de Chat-GPT aura, par ailleurs, constitue une enorme campagne communication favorisant l'acceptabilite sociale du des IAs sur le web, tout en developpant la conscience de la difficulte a venir de distinguer entre un contenu produit par un humain d'un contenu produit par une IA. Ceci permet a Sam Altman, un an plus tard, de presenter sa nouvelle entreprise, Worldcoin, qui declare qu'a l'heure ou l'IA et l'humain deviennent indiscernables en ligne, il est necessaire de consentir a l'emploi massif de donnees biometriques d'identification sur le web. Deja 450 000 iris de personnes vivant pour la plupart dans les Suds ont ete scannees contre de l'argent. Une partie des mots de passes et acces au service ont deja ete hackees. L'article qui suit a ete difficile a ecrire pour exactement cette raison; je prends ici position en faveur de l'impossibilite totale de faire l'economie d'une politisation de ses questionnements.

Introduction

Dès les premières lignes du chapitre que Simondon consacre à l'invention dans un texte de cours, *Imagination et invention*, il s'emploie à définir l'invention depuis la notion de problème, telle qu'elle viendrait qualifier une situation:

> Est problématique la situation qui dualise l'action, la tronçonne en la séparant en segments, soit parce qu'il manque un moyen terme, soit parce que la réalisation d'une partie de l'action détruit une autre partie également nécessaire; hiatus et incompatibilité sont les deux modes problématiques fondamentaux (Simondon 2014, 139).

L'invention suppose une situation problématique, qu'elle permettra de résoudre. On peut donc identifier la situation problématique comme le point de départ de l'inventivité et de la création. Ce point de départ nous permet de poser un cadre théorique depuis lequel traiter la question de la créativité chez les IAs, de trois manières. D'abord, en interrogeant la possibilité pour un modèle de traitement du langage naturel (multimodal si l'on intègre des programmes tels que Midjourney par exemple) d'être concerné par la notion de problème ainsi définie. Ensuite, en interrogeant la manière dont l'introduction de ces nouveaux outils dans la culture matérielle humaine produit à son tour de nouvelles lignes de partage dans nos conceptions de la créativité et de l'expression, lignes de partage qu'il s'agirait donc de penser. Enfin, dans la mesure où Simondon propose deux types de problèmes, celui procédant d'un hiatus, et l'autre d'une incompatibilité, on peut se demander dans les deux cas s'il l'on aurait affaire à un hiatus ou à une incompatibilité.

Nous proposerons, à partir de cette notion de problème, de montrer quelles sont les limites théoriques qui à mon sens empêchent que l'on détermine les IAs comme étant créatives, y compris lorsqu'elles sont utilisées à des fins de productions esthétiques et culturelles. Nous nous interrogerons également sur la nature de cet effort par lequel on essaie d'aménager notre concept de créativité pour qu'il devienne compatible avec la computation. Ceci permettra d'interroger, enfin, ce que l'on peut appeler le milieu associé des IAs, un milieu humain pour lequel il y a culture, production symbolique et production de valeur. Nous mettrons au centre de notre réflexion les notions de virtuel,

de milieu et de valeur. Le virtuel comme dimension nécessaire de l'inventivité, dont sont privées les IAs, nous conduit à refuser qu'elles puissent être déclarées créatives au même titre que les êtres vivants, quoique la métaphore organiciste et l'anthropomorphisme que l'on mobilise pour parler des IAs nous induisent parfois à projeter ces propriétés sur elles. Nous détaillerons les propriétés du milieu associé chez Simondon, ce qui permettra d'insister sur le fait que le milieu dit associé des IAs, ce sont les humains, et que ce sont eux qui restent bien les dernières instances régulatrices des programmes. Enfin, en insistant sur le rapport entre la créativité dans les activités humaines et la valeur (car une situation problématique est une situation qu'il importe de résoudre), et sachant que cette valorisation n'est pas réductible à une récompense quelconque programmée pour favoriser un entraînement, nous conclurons qu'en l'état actuel des choses, il n'est pas possible de qualifier une IA de créative. Précisons toutefois d'emblée que cela ne signifie absolument pas que l'IA n'a pas de valeur ni qu'il est absurde de vouloir créer des objets culturels avec elle. Simplement, nous tenons à insister sur le fait que la valeur de ses productions ne peut pas dériver de l'outil utilisé.

Pour ce faire, nous resterons dans un corpus restreint: nous n'entendons pas produire un discours général sur les IAs qui, en tant que *task-specific*, sont aussi peu homogénéisables que la technique en général. Nous nous concentrerons sur les IAs qui produisent des contenus susceptibles d'avoir une valeur esthétique et culturelle, et qui fonctionnent sur des LLM, *large language models*. Il ne sera par conséquent pas question ici des IAs symboliques (telles que le célèbre AARON d'Harold Cohen), qui procèdent par instructions déterminées. Il s'agira plutôt des IAs connexionnistes les plus avancées, celles dont la lignée technique a commencé avec McCulloch et le modèle d'activation du neurone en 1943, avec le *perceptron* de Rosenblatt dans les années 1960, et dont l'étude a repris spectaculairement en 2012 avec Hinton et LeCun (Cardon *et al.* 2018, 181). Nous aborderons donc les IAs connexionnistes telles que les LLMs de GPT-3 et son agent conversationnel, ainsi que les LLMs multimodaux tels que Midjourney.

1. Qui invente?

Au cours de la première moitié du vingtième siècle, plusieurs entomologistes et biologistes rencontrent et étudient plusieurs espèces d'animaux non humains (la guêpe *Ammophilia urnaria*, les fourmis *Oecophylla smaragdina* ou *Camponotus senex*) qui ont recours à des médiations instrumentales. Voilà que la définition de l'être humain comme *homo faber*, attribuée à Benjamin Franklin puis Bergson, est mise en cause dans son exclusivité. Qu'est-ce qui, dans la médiation instrumentale humaine, permettrait de continuer à distinguer entre la technique hominisante, et une technique toujours dérivative de l'instinct, sans voie d'accès au virtuel ?

On peut ne pas se satisfaire de la manière dont la question est posée, au nom de la distinction qui y est posée entre réponse instinctive et réponse «intelligente». Or, Simondon développe à ce sujet une voie intermédiaire: il fait de l'inventivité une fonction de l'intelligence non seulement individuelle, mais aussi collective et qui ne soit effectivement pas réservée qu'aux humains. Il essaie de penser l'inventivité des autres espèces selon un ensemble de variables plus riche que la simple intelligence individuelle.Il propose :

(1) d'intégrer une échelle collective à l'inventivité, qui n'est alors plus solidaire d'un génie individuel extraordinaire

(2) de penser la conjugaison des échelles individuelles et collectives de l'intelligence

(3) de se garder de faire de l'équidistribution d'une médiation instrumentale à travers l'espèce comme signe de la nature instinctive de cette conduite, en envisageant (a) que plusieurs individus aient pu découvrir la même solution à un problème qui s'était posé à chacun d'eux, ou (b) qu'elle s'est répandue au présent dans une colonie

(4) d'intégrer à l'inventivité des espèces les dimensions situationnelles des problèmes que ces dernières ont dû résoudre, et qui permettent de rappeler que l'invention possède une genèse dans des problèmes posés *hic et nunc* à des êtres.

Ces propositions rendent possible, grâce à la multiplication des échelles et l'intégration d'une dimension situationnelle précise, un cadre d'analyse qui ne soit pas anthropomorphiste ni individualiste, cadre qui avait été adopté majoritairement jusqu'alors (l'intelligence est d'abord la propriété d'un *individu* intelligent, et d'un individu *humain* en particulier). S'éloigner de l'anthropomorphisme et de l'individualisme permet de s'éloigner des modèles que l'on pourrait appeler «héroïques et miraculeux» de l'inventivité, tel qu'on peut les retrouver par exemple dans les théories scientifiques. On retrouve par exemple ce modèle dans un texte de Poincaré en 1908, dont les reprises ont été retracées par Judith Schangler dans la généalogie qu'elle propose, de Poincaré à Koestler, des modèles discontinuistes de l'invention de concepts scientifiques (Stengers & Schlanger 1991, 72). Dans un geste critique similaire, nous nous attaquerons aux limites de l'anthropomorphisme et de la métaphore organiciste souvent mobilisée à titre heuristique ou pédagogique au sujet des IAs et de leur «évolution». L'anthropomorphisme a déjà été traité par Watson dans un article récent (Watson 2019), il se concentrait sur les conséquences du vocable anthropomorphique sur notre promptitude collective à transférer aux IAs des décisions qui nous concernent. Ce problème, actuel, concerne toujours les chercheur·e·s et les communiquant·e·s.[1] En ce qui nous concerne, c'est de l'association de propriétés d'un organisme à une machine computationnelle que nous allons traiter.

2. Les IAs sont-elles des zombies? Sur les limites de la métaphore organiciste

Si les IAs sont effectivement similaires à des organismes, puisqu'elles sont interprétables depuis les mêmes schèmes conceptuels, à quels organismes sont-elles similaires? La plupart du temps, c'est aux humains qu'il s'agit de les assimiler, en particulier pour favoriser l'acceptation de ces dernières au sein de l'espace social le plus large possible (Yang *et al.* 2022). C'est la raison pour laquelle, par exemple, c'est Chat-GPT

1. *Why Does AI Lie, and What Can We Do About It?*, 2022, [en ligne], Url: https://www.youtube.com/watch?v=w65p_IIp6JY. Il est évident que Miles sait de quoi il parle. Néanmoins, l'option pédagogique retenue est discutable car elle repose aussi sur une visualisation fallacieuse des opérations du système.

plutôt que GPT-3 qui parvient à atteindre le grand public. Son anthropomorphisme, trompeur, est aussi ce qui la rend acceptable (Rantala 2020[2]).

Mais partons de cette image d'IA anthropomorphe. Que se passe-t-il lorsque les objets techniques IAs sont mis en circulation? Lorsqu'une IA est déployée, comme pour l'agent conversationnel d'OpenAI, elle n'est plus en cours d'entraînement et n'évolue plus. La nature dynamique et perfectible du programme n'est plus. Dès lors, il faut ôter la vie à cet objet technique anthropomorphe: les IAs sont alors plus proches des zombies que des humains, en ceci qu'elles ne peuvent faire qu'avec ce avec qu'elles sont depuis l'arrêt de l'entraînement. Lorsqu'on les sollicite par l'envoi d'un prompt, elles ne sont plus dans un état dans lequel elles peuvent se transformer. Derrière l'anthropomorphisme et la personnalisation des IAs, c'est le rapport de l'objet technique avec le vivant qui me semble poser le problème le plus conséquent ici.

Car dans les conditions d'une IA déployée, il n'est pas possible de concevoir les interactions utilisateurice/IA comme des interactions équivalentes à une interaction entre deux êtres «organiques». Ceci, parce qu'une fois que les programmes sont entraînés, ces derniers ne «métabolisent» plus rien. Une IA déployée n'apprend plus, ne modifie plus ses structures – notamment pour des raisons de sécurité. Il n'y a, alors, plus de relation intégrative, ni d'autorégulation.

2. Ce point a notamment été étudié sous un angle simondonien par Rantala. [En ligne], URL: http://ebooks.iospress.nl/doi/10.3233/FAIA200946. Pour une revue de littérature complète, qui pointe l'instabilité du concept d'anthropomorphisme et les conclusions divergentes quant à l'impact de l'anthropomorphisme sur l'acceptabilité sociale des IAs, voir Li M. & Suh A., *Machinelike or Humanlike? A Literature Review of Anthropomorphism in AI-Enabled Technology* [en ligne], dans *Proceedings of the 54th Hawaii International Conference on System Sciences*, présenté à 54th Hawaii International Conference on System Sciences (HICSS 2021), 2021, Url: https://scholars.cityu.edu.hk/en/publications/machinelike-or-humanlike-a-literature-review-of-anthropomorphism-in-aienabled-technology(7c9fa6e3-ab93-4d72-ab37-219f2733dd1e).html; Blut M. *et al.*, *Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI*, «Journal of the Academy of Marketing Science», 49(4), 2021.

3. Le vivant et le technique

Ce qui fait d'une IA un zombie correspond en langage simondonien à ce qui fait d'un objet technique un objet technique plutôt qu'un être vivant. Le parallèle se justifie ici d'autant mieux que le zombie et le non-zombie entretiennent une proximité que l'objet technique et l'être vivant entretiennent également lorsqu'ils sont définis par l'auteur depuis la notion de transduction (Simondon 2012, 198-199). En effet, on peut distinguer deux types de transducteurs: ceux qui relèvent d'une transduction mécanique, et les transducteurs propres au vivant. Simondon se garde toutefois de rejouer dans cette dichotomie la carte de la métaphysique moderne selon laquelle on aurait au niveau du mécanique un déterminisme strict et laplacien, et au niveau du vivant la présence de la liberté sous la forme d'une indétermination chronique. En effet, il précise immédiatementque

> les transducteurs mécaniques sont des systèmes qui comportent une marge d'indétermination; l'information est ce qui apporte la détermination. Mais il faut que cette information soit donnée au transducteur; il ne l'invente pas; elle lui est donnée par un mécanisme analogue à celui de la perception chez le vivant, par exemple par un signal provenant de la manière dont l'effecteur fonctionne [...]. Au contraire, le vivant a la capacité de se donner à lui-même une information, même en l'absence de toute perception, parce qu'il possède la capacité de modifier les formes des problèmes à résoudre; pour la machine, il n'y a pas de problèmes, mais seulement des données modulant des transducteurs.(Simondon 2012, 199)

Autrement dit, la différence n'est pas celle qui séparerait le déterminisme mécaniste classique de la liberté métaphysique de l'être humain. Ce sont les affordances des deux transducteurs quant aux dimensions temporelles du réel auxquelles ces transducteurs sont réceptifs qui les distinguent. Le vivant possède de manière spécifique la possibilité de s'individuer depuis lui-même, de se «donner à lui-même une information». Cette capacité est justifiée dans le texte par la proposition suivante «il possède la capacité de modifier les formes des problèmes à résoudre». C'est la possibilité d'opérer cette modification qui fait la spécificité du vivant par opposition à la machine, qui ne peut encore une fois fonctionner qu'avec ce qu'elle a. Que suppose «pouvoir modifier les formes du problème à résoudre»? On trouve la réponse à cette

question un peu plus loin dans le texte. Si le vivant peut modifier les formes des problèmes qui sont les siens, c'est aussi parce qu'il peut faire avec ce qui n'est pas et n'est pas encore, avec non seulement le présent et le passé (les données enregistrées pour une IA). C'est ce que Simondon appelle le virtuel, et c'est ce qui rend possible la résolution de problème, c'est-à-dire l'invention:

> La résolution des véritables problèmes est une fonction vitale supposant un mode d'action récurrente qui ne peut exister dans une machine: la récurrence de l'avenir sur le présent, du virtuel sur l'actuel. Il n'y a pas de véritable virtuel pour une machine[…] il n'y a pas de modification de formes qui soit orientée par le pressentiment du problème à résoudre; le virtuel ne réagit pas sur l'actuel, parce que le virtuel ne peut jouer aucun rôle en tant que virtuel pour la machine.(Simondon 2012, 200)

L'intégration de ce qui n'est pas (encore), c'est-à-dire du virtuel nous permet alors de distinguer entre deux types d'indétermination, celle relative à la machine et celle relative au vivant. Il existerait en effet une indétermination déterminée depuis l'actuel seulement, et l'indétermination apportée par une relation à ce qui n'est pas encore, mais qui peut être intégré à ce qui est comme l'une de ses dimensions: c'est le virtuel. Soulignons toutefois que l'«actuel» ne correspond pas à l'«intérieur»: il ne s'agit pas de dire que la machine est fermée sur elle-même, ni qu'elle serait privée de rapport avec l'extérieur et ce qui n'est pas elle-même. Il est même décisif d'insister sur le fait que ce rapport existe: l'envoi d'un prompt, d'une ou plusieurs images, l'alimentation de la machine en sont des exemples. Toutefois ce rapport n'implique pas que les IAs lorsqu'elles sont déployées soient susceptibles d'entretenir un rapport quelconque avec le virtuel, car elles ne sont pas capables d'intégrer quoi que ce soit, ni de faire avec ce qui n'est pas. L'existence d'une interaction dynamique entre un objet technique et un environnement n'est pas analytiquement corrélée à une situation où la virtualité serait partagée. Si une IA interagit avec un humain, le sens de la virtualité sera asymétrique.

Par conséquent, les algorithmes dont nous nous servons, qui sont récursifs, ne reposent que sur la récursion possible du présent vers le passé, c'est-à-dire la transduction machinique; tandis que le vivant dispose d'un autre type de boucle de rétroaction, celui de l'avenir sur

le passé. Le sens de l'avenir dont il est question chez les individus biologiques est en ce sens irréductible à un pronostic, à une prédiction basée sur des données présentes. En tant que machine non vivante, l'IA est, toutes choses égales par ailleurs, incapable de contact avec le virtuel et l'avenir, dont on pose qu'il n'est pas réductible à une prédiction statistique basée sur une appréhension du réel par data-ification d'une partie de ce dernier.(Simondon 2012, 200)

Ainsi, l'invention est suscitée par une situation problématique, c'est-à-dire une situation dans laquelle persiste un hiatus ou une incompatibilité entre deux ou plusieurs choses; et la résolution d'un problème requiert, pour qu'il y ait invention, réforme des formes initiales des données du problème ainsi que rapport avec ce qui n'est pas et avec un avenir possible, c'est-à-dire avec du virtuel. Que devient, dès lors la créativité computationnelle dans ce cadre de référence? Pas grand-chose, puisque les IAs n'ont rien de vivant. Si l'on pose une équivalence entre la création et l'invention, de sorte que la créativité devient la condition de possibilité de l'invention, alors la créativité n'est pas un concept qui a un sens pour les IAs déployées.[3] Mais alors, comment comprendre la dimension dynamique des échanges que l'on peut avoir avec les agents conversationnels ou avec les robots émotionnels (Devillers 2020, chap. 6.3), comment comprendre à quel point ces interactions peuvent nous être *appropriées*, inédites sans pour autant recourir au virtuel? Comment expliquer que l'absence de partage du sens du virtuel semble ne pas (immédiatement) poser problème?

4. De l'organisme au personnage de *Huis-Clos*

Il semble que pour répondre à cette question, il nous faille retravailler l'image de l'IA anthropomorphe et sans vie. En effet, après avoir proposé de visualiser les IAs comme des zombies, saisissons les agents conversationnels – et par extension, les IAs – en leur donnant des noms qui seront peut-être reconnus: Estelle, Garcin, et Inès.

3. Déployées, parce que nous n'excluons pas que la question de la créativité puisse être posée durant la phase d'entraînement, quoique l'ajustement de paramètres pour l'accomplissement d'une fonction donnée (l'accomplissement d'un but, fut-il aligné ou non avec l'objectif humain initial) puisse être distingué de la genèse d'une fonction: il nous faut distinguer entre une IA «qui apprend» et une IA «qui a appris».

Chat-GPT est un personnage de *Huis-Clos* (Sartre 1947): déjà mort, il peut être mis en scène dans une situation donnée. En l'occurrence, l'enfer. Dans cet enfer, l'agent conversationnel pourrait interagir avec d'autres personnages, et ce de manière appropriée à la situation dans laquelle ils se trouvent. Les personnages parviennent à avoir un dialogue convenu, malgré la nouveauté de la situation dans laquelle la pièce les plonge. On peut même aller jusqu'à dire que les personnages semblent avoir un petit rayon d'action, quoique leurs projets respectifs soient voués à l'échec, puisque la pièce est structurée autour d'un *fait* auquel il est impossible d'échapper: la mort a figé leur être dans une identité de laquelle chaque personnage aimerait se détourner, alors que cela n'est plus possible étant donné qu'il et elles sont mort·es. Il ne leur est plus possible, quoiqu'iels en pensent, de se transformer et autre chose que ce que ce qu'iels ont été, et ce de que leur vie aura fait d'elleux. Plus de réforme possible de soi, pas de transformation: seule persiste l'itération de ce qui est.

C'est cela que nous entendons par absence de métabolisation, et ce qui nous semble occulté par un langage qui a tendance à traiter les IAs comme des organismes vivants. L'absence de métabolisme implique une double incapacité: d'une part, l'incapacité à réformer ses propres formes depuis une participation extérieure, en l'occurrence depuis l'avenir comme virtuel. D'autre part, l'incapacité de réformer ses propres formes depuis soi seul, en tant que le vivant est capable de se donner à lui-même une information, c'est-à-dire quelque chose qui produit une transformation du et dans le système original. Cette absence paraît tout à fait décisive pour penser la différence fondamentale qu'il existe entre le résultat d'un processus computationnel d'un côté, et le résultat d'un processus d'individuation de l'autre. Dans le rapport humain-programme, y compris lorsque ce rapport est dynamique, les fonctions de régulation et d'individuation sont bien encore assurées par les humains (Simondon 2012, 96), et non par les êtres techniques. La régulation dont il est question ici concerne aussi bien les normes d'usages que l'apprentissage de la communication avec les IAs, comme on affute ses gestes à un outil pour produire un geste technique de qualité. Voyons pour preuve la prolifération des guides pour bien écrire des prompts et bien obtenir ce que l'on attend de Midjourney par exemple. On apprend à entrer en relation avec un objet technique pour

obtenir un résultat favorable à nos objectifs: lorsqu'on opère via ces objets techniques, on se rapproche d'eux dans nos modes opératoires.

5. Penser la création en computationnaliste?

Imaginons que Sartre décide d'aménager un peu sa pièce. Il fournit à ses personnages une machine à écrire, et de quoi l'utiliser comme bon leur semble. Que se passerait-il? Pourraient-iels écrire un roman, un poème, un texte quelconque de manière créative? Il y a, en réalité, deux questions différentes ici:

- Quel est le lien entre le créateur et la créature/création? (Est-ce parce qu'Inès peut créer que ce qu'elle produit est une création?)
- Qui est le juge adéquat et quels sont les bons critères d'évaluation de la présence de «créativité»? (Inès pourra-t-elle décréter sa production comme créative, cela dépendra-t-il des autres personnages, ou d'un réel humain vivant, mais absent de la pièce?)

Pour construire nos réponses à ces différentes questions, on peut se référer au de M. Boden. La chercheure a en effet proposé de distinguer entre l'h-créativité et la p-créativité, c'est-à-dire l'évaluation historique de la créativité et l'évaluation psychologique de la créativité. Il s'agit en effet de deux critères d'évaluation qui instaurent deux échelles possibles de l'avènement du nouveau, et permettent à la fois de considérer que l'enfant qui découvre sans personne la solution déjà connue par d'autres à un problème découvre bien quelque chose, et qu'il aura fallu que l'enfant invente effectivement une solution à un problème qui s'est réellement posé *hic et nunc*. L'évaluation historique de la créativité est, par contraste, une condition maximale, depuis laquelle on peut penser les apports des individus ou collectifs d'artistes et de créatif·ve·s en général à «une grande aventure humaine» que l'on est bien obligé·e de réifier un peu pour que ce concept ait un sens. Pour ce qui concerne le domaine du computationnel, c'est le concept d'évaluation psychologique de la créativité qui est le plus souvent retenu (Gervas 2009, 51). Or, la proposition «*unless a computer program is given access to historical data (and generally provided with means for social interactions with other creators*)»(Gervas 2009, 51) est réalisée aujourd'hui à 50%

(un programme sur ordinateur a accès à des données historiques, celles qui sont présentes parmi d'autres types de données sur l'Internet). Pourrions-nous, dans ces coordonnées, évaluer la créativité putative d'un programme? Si l'évaluation historique de la créativité repose sur la prise en compte de données historiques, c'est-à-dire s'évalue dans le contexte général de la construction culturelle, alors il faudrait pouvoir poser la question en ces termes, volontairement provocateurs: que répondrions-nous à la question: Garcin, Estelle et Inès peuvent-iels ou non inventer le graff, le *drag*, les *snapewives*(Alderton 2014), le point de fuite, la voûte, *Tell me that I'm dreaming*?[4] Et si la réponse et non, pourquoi serions-nous plus disposé·e à accorder cette capacité à une IA?

Ces questions ont l'air absurdes étant donné que l'on défend qu'il existe une différence de nature entre la créativité humaine d'une part, et la «créativité» algorithmique d'autre part en tant que composition de données qui *nous* surprend. Mais ce questionnement sert moins ici à interroger les programmes grâce auxquels on génère un texte ou une image inédite, qu'à interroger la «définition de travail» de la culture depuis laquelle on réfléchit sur les productions algorithmiques dites créatives. Lorsque l'on prend les pans de la culture humaine que l'on pourrait qualifier de «non sollicités», et les formes de socialité qui y sont associées, alors on retrouve dans la notion de créativité ce qui précisément échappe à toute algorithmique et qui engage des problématiques vitales qui donnent une valeur à toutes ces créations. Pour comprendre ces genèses culturelles, on doit à la fois manipuler les échelles micro (*e.g.* la p-créativité) et macro (*e.g.* la h-créativité), mais on doit également introduire des échelles méso, c'est-à-dire l'échelle des communautés, groupes et sous-groupes sociaux, en incluant les conditions matérielles des créateurs et créatrices, en tant qu'elles contribuent à la construction de la problématique à laquelle la pratique culturelle inventée répond.

C'est en ce sens que l'on peut, pour penser l'individuation culturelle et l'individuation des groupes dans la culture, autoriser l'usage du terme de métabolisation, conformément à ce qui a été développé plus

4. Was (not was), *Tell Me That I'm Dreaming*, [en ligne], «Island Records», 1981, URL: https://www.youtube.com/watch?v=ZaawHz6p1QE

haut. Mais ce terme est aussi important pour une autre raison: dans un ouvrage consacré à la question de la vie artificielle, Boden s'interroge à son tour sur la notion de métabolisme et sur son applicabilité. En effet, la chercheuse en informatique et en sciences cognitives développe une argumentation qui entend réfuter la possibilité d'une «*strong artificial life*»[5] ou vie artificielle robuste depuis le concept de métabolisme. Le métabolisme est alors défini comme l'usage et la budgétisation autonomes de l'énergie requise à la constitution, le maintien et le comportement de la chose vivante (Boden 1999, 236-237).

Ce que Boden ne fait pas, en revanche, c'est faire le lien entre le métabolisme et la créativité. Son travail aura consisté à construire un concept de créativité qui fonctionne essentiellement sur l'évaluation des propriétés d'un résultat, plutôt que sur la description d'un processus. La dimension processuelle une fois exclue de l'analyse, la créativité devient une propriété projetée sur l'être qui a produit quelque chose que l'on aura qualifié de création.

Comment justifier ici la mise en cause d'un lien entre les notions de métabolisme et de créativité? On peut faire l'hypothèse que c'est en vertu des hypothèses de travail qui traversent le champ de la recherche en IA, à savoir l'adoption plus ou moins explicite de théories computationnalistes,[6] et connexionnistes[7] ou fonctionnalistes[8] de la conscience et du mental. C'est cette adoption qui justifie que l'on passe parfois trop vite d'une ressemblance entre un réseau de neurones et un fonctionnement humain à une identification de l'un et de l'autre et de leurs affordances, *modulo* les différences quantitatives entre les

5. Une réplique mécanique d'un être vivant dans un milieu de vie non-simulé. C'est-à-dire un organisme avec qui on partagerait le monde.

6. Le mental est équivalent à une machine computationnelle de Turing. Ce modèle est informé par un article publié par *le Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133 en 1943, donc nous citons ici la republication: McCulloch W. S. & Pitts W., *A logical calculus of the ideas immanent in nervous activity*, «Bulletin of Mathematical Biology», 52(1/2), 1990, p. 113.

7. L'équivalent du mental est le réseau mécanique de neurones, et le mental est compris d'après un modèle computationnel. La conception est également informée par *Ibid.*, p. 108.

8. Le psychique est pensé en termes d'états, et ces états sont décrits dans leurs occurrences comme répondant à une fonction donnée. Cette position est développée par Putnam H., *Minds and Machines* [en ligne], dans Hook S. (éd.), *Dimensions of Minds*, New York University Press, 1960, URL: https://philarchive.org/rec/PUTMAM.

deux systèmes, et ceci parce qu'ils produiraient les mêmes effets. La production d'un analogue fonctionnel technique à un neurone depuis McCulloch et Rosenblatt, puisqu'on met le cerveau seul au centre de la créativité, suffirait à penser la créativité sans l'associer à la totalité d'un être vivant, et donc à un métabolisme.

Le computationnalisme est un cas particulier de fonctionnalisme,[9] selon lequel l'esprit/le cerveau a pour fonction de traiter de l'information, ce qui est accompli par le calcul. On peut envisager deux objections au fonctionnalisme et au computationnalisme. La première, dirigée contre une conception fonctionnaliste du mental, est la plus connue. C'est l'objection de la chambre Chinoise de Searle (Searle 1980). Moyennant la mise à disposition dans une pièce de toutes les combinaisons possibles autorisées par une langue, et le temps pour trouver chaque phrase et écrire une réponse qui corresponde à la phrase envoyée, un homme non-sinophone enfermé dans cette pièce pourrait donner l'illusion de comprendre le chinois, alors qu'il ne ferait que fournir des *outputs* y conformes à des inputs x qu'il ne comprendrait pas par ailleurs. L'identité de résultat ne peut donc être l'indice suffisant de l'identité d'un processus. La deuxième objection est celle du paradoxe des mouches et des chaussures, que l'on trouve développé chez Sober (Sober 1996) dans le cadre d'une argumentation contre le fonctionnalisme, dans lequel il trouve pour sa part la source de la tromperie. Nous le restituons ici d'après un article de Li (Li 2018, 501).[10]:

La population génétique (de mouches, pour l'exemple) évolue selon la loi d'Hardy-Weinberg, selon laquelle la fréquence d'un allèle dans les rencontres randomisée entre membres d'une population ne sera pas changée d'une génération à l'autre s'il n'y a ni sélection, ni migration, et s'il y a nombre égal de partenaires mâles et femelles. Soient les fréquences p et q des allèles A et a, selon ce modèle la fréquence des

9. Les deux positions sont ordinairement associées. Pour une défense de la nécessité de les dissocier, voir Piccinini G., *Functionalism, Computationalism, and Mental Contents*, «Canadian Journal of Philosophy», 34(3), [Taylor & Francis, Ltd., Canadian Journal of Philosophy], 2004; Piccinini G., *The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism*, «Philosophy and Phenomenological Research», 81(2), [International Phenomenological Society, Philosophy and Phenomenological Research, Wiley], 2010.

10. Les deux paragraphes d'exposés qui suivent sont une traduction partielle de la présentation du paradoxe par Li.

paires AA, Aa et aa dans une population donnée sera de p^2, $2pq$ et q^2. On retrouve cela dans l'étude de populations données de mouches, par exemple.

Or, si l'on prend une population de chaussures – l'une marron et l'une noire – et qu'en vertu d'une erreur de paramétrage sur la ligne de production, la couleur des chaussures qui se retrouvent empaquetées ensemble est choisie au hasard, avec p la fréquence de sélection d'une chaussure noire et q la fréquence de sélection d'une chaussure marron, il se trouve que la fréquence attendue de présence de différentes paires (noire-noire, noire-marron et marron-marron) sera respectivement de p^2, $2pq$ et q^2.

Ce n'est donc pas parce que deux choses peuvent répondre de la même loi que ces deux choses sont identiques, et que toutes leurs propriétés associées sont identiques. On ne peut déduire de la subsomption sous une loi la nature d'un phénomène. Ce qui signifie que l'on ne peut se servir de leur proximité avérée pour dériver d'autres propriétés communes – comme on entreprend de dériver de l'expérience d'un *output* qui *nous* surprend chez une machine computationnelle, la présence d'une propriété qui produit aussi la surprise chez les humains, à savoir la créativité. En effet, il n'est évidemment pas question de dire que les chaussures sont en vie, et que l'on peut inférer la vie de celles-ci quoique la loi semble s'appliquer aussi bien à des êtres vivants qu'à des objets techniques inertes. De manière prospective, ce paradoxe permet de faire un pas supplémentaire, pour dire ceci : même dans le cas où la créativité humaine et la créativité computationnelle répondraient des mêmes lois, rien ne garantirait pour autant que les deux soient identiques, ni qu'elles soient équivalentes. Ici, la formule mathématique permet de prédire le développement d'une situation qui, dans le cas des mouches, prend sens du point de vue de l'individuation vitale, tandis que dans l'autre (les chaussures), cette prise de sens n'existe pas. Cela n'ôte rien au potentiel explicatif et prédictif de ladite loi, mais cela montre que l'existence d'une loi prédictive partagée ne permet pas de déduire des propriétés communes.

Dans un sens comme dans l'autre, du résultat vers le processus ou de la loi qui permet de prédire vers la nature de ce qui est prédit, nous rencontrons des limites qui nous permettent de mettre en doute la possibilité de séparer à ce point la créativité du métabolisme. Mais essayons de faire un pas supplémentaire. Admettons qu'il existe un

calque fonctionnel de l'intelligence, ou que nous soyons parvenu·e·s à optimiser une fonction telle qu'elle produise des *outputs* tout à fait inédits. Même dans ce cas, il resterait à statuer sur ceci: «du point de vue de la connaissance, la question n'est pas d'inventer des idées originales, mais d'inventer des idées pertinentes», comme l'écrit Schlanger (Stengers & Schlanger 1991, 124). Cela s'applique aussi à la production symbolique. Dans quelle mesure les productions d'IA authentiquement créatives, si elles l'étaient, auraient un sens pour nous?

6. Métabolisme, milieu, individuation

La définition que Boden propose du métabolisme, mentionnée plus haut,[11] en fait la propriété abstraite d'un être tout aussi abstrait, et dont le milieu est par conséquent tout aussi abstrait: nulle mention en effet du milieu dans lequel et par lequel s'effectue cette métabolisation. Tout est réduit à une question de gestion assez flottante de l'énergie sur laquelle semble flotter le paradigme classique de *l'homo economicus,* l'homme de la théorie du choix rationnel, qui cherche à maximiser son propre bien conçu comme un ensemble de fonctions d'utilités à optimiser (Pareto 1964; 1981).

Or, la gestion et la budgétisation de l'énergie pour la croissance, le maintien et le comportement d'un être sont toujours non seulement spécifiques, mais aussi fondamentalement relationnelles et situées. Les comportements se font, s'apprennent, s'affinent notamment dans le jeu, mais parfois aussi s'inventent à un endroit puis se répandent sur un territoire (Firth *et al.* 2020). Autrement dit, il n'y a pas de métabolisme sans métabolisation, c'est-à-dire sans milieu dans et avec lequel métaboliser. Avec, parce que la métabolisation n'est pas tant une relation entre milieu constitué et individu constitué, dans les coordonnées simondoniennes, que ce qui est constitué dans une relation individuante. Pour qu'il y ait métabolisation en effet, il ne faut pas en premier lieu qu'il existe un individu et un milieu séparés: il faut qu'il existe une homogénéité et une compatibilité entre un intérieur et un extérieur, de sorte que le passage et l'intégration à un nouveau système d'un élément extérieur et d'un élément intérieur soient ga-

11. Je la rappelle ici, il s'agit de la budgétisation autonome de l'énergie requise à la constitution, le maintien et le comportement de la chose vivante.

rantis. Ce que le métabolisme comme opération paradigmatique du vivant suppose n'est donc pas simplement l'hypothèse bodenienne d'une fonction d'alimentation en énergie à optimiser, mais une relation active et individuante.

C'est la raison pour laquelle la notion de milieu est ici décisive dans notre compréhension de ce qu'est un métabolisme une fois posé le diagnostic de l'insuffisance de la définition qu'en donne Boden. Un petit détour par les conceptions simondoniennes des milieuxs'impose par conséquent. Nous utilisons le pluriel, car on peut en distinguer deux figures, auxquelles nous ajoutons un troisième élément : le milieu associé, le milieu, et l'ensemble.

7. Milieu associé, milieu, ensemble

7.1. Milieu associé[12]

Dans le premier sens, le milieu associé correspond à un concept qui désigne dans les objets techniques concrets les conditions d'existences qui adviennent dans et par l'objet et qui permettent son fonctionnement. Pour illustrer ce concept dans des coordonnées simondoniennes classiques, le milieu associé de l'individu technique «turbine de Guimbal» n'est pas le cours d'eau ni le milieu naturel dans lequel la turbine est insérée, mais le complexe plurifonctionnel huile-eau rendu possible par la réduction de la taille de la génératrice et désormais insérable dans un carter d'huile sous pression, qui graisse et transmet la chaleur à l'eau qui, elle, refroidit et met aussi en mouvement la turbine (Simondon 2012, 70-73). La plurifonctionnalité rend possible l'objet, elle est l'avènement d'une composition nouvelle et d'affordances jusqu'alors inaccessibles. Le milieu associé est ici *intégré*: il est condition *sine*

12. Le concept de «milieuassocié» n'intervient en ces termes que dans les travaux de thèse de Simondon. On le rencontre dans le *Mode d'existence des objets techniques*, et on ne le recroisera plus dans les travaux plus tardifs, y compris ceux pour lesquels le concept de «milieu» tout court est prépondérant, par exemple le texte intitulé «Culture et technique», l'article de 1965, et le cours «Imagination et invention», prononcé en Sorbonne la même année. On peut distinguer deux usages du milieu associé dans le *MEOT*: un usage du terme en un sens restreint, appliqué aux individus techniques; et un usage plus large, et plus souple. Nous ajouterons, pour ce travail, un petit mot sur la notion d'ensemble. Essayons donc de saisir plus précisément ce que métaboliser peut bien vouloir dire.

qua non (Simondon 2012, 75) de l'objet technique, et cette intégration fonctionnera pour l'auteur comme un critère d'individualité des objets. En effet, l'individu technique devient alors ce qui possède son propre régime de causalité réciproque et régulatrice (cela maintient l'intégrité de l'objet), sachant que cette régulation est simultanée: ce n'est pas un événement A qui enclenche une réponse B dans le système, mais le débit d'eau augmentant, il augmente la puissance du processus de refroidissement de la turbine au moment même où il augmente la température produite en accélérant la rotation de la génératrice.

7.2. Milieu

À côté de cette première conception, on trouve une acception plus souple, et plus large de la notion de milieu: il désigne alors ce qui en réalité n'est pas un milieu associé au sens étroit, mais nomme plutôt la portion de réel homogène avec l'être dont le milieu est le milieu. Cette homogénéité n'implique pas nécessairement de réciprocité dans la causalité (un effet sur y entraînant un contre-effet régulateur sur x) (Simondon 2012, 76). C'est ce deuxième usage plus souple du concept de milieu qui restera courant chez Simondon, pour désigner des ensembles qualifiés (milieu naturel, social, professionnel, technique, symbolique…) qui rencontrent causalement les objets techniques, les ensembles sociaux et les règnes de l'inerte, du vivant et du psychique.

7.3. Ensemble

Si nous faisons intervenir le terme d'ensemble ici, c'est parce qu'il permet de présenter un dernier mode d'organisation des objets techniques. Là où le milieu et le milieu associé pouvaient correspondre à deux manières de se rapporter aux individus et aux individus techniques en particulier pour le milieu associé au sens étroit, l'ensemble technique est un terme opérant à un autre niveau. En effet, chez Simondon, l'ensemble technique est une catégorie typologique parmi deux autres, les éléments techniques, et les individus techniques (qui ont, donc, un milieu associé). Un ensemble technique correspond à régime de technicité qui articule différents fonctionnements techniques sans entreprendre de les fondre dans des conditionnements mutuels et simultanés. L'enjeu d'un ensemble est donc d'articuler des résultats de fonctionnement

(Simondon 2012, 79), plutôt que d'organiser techniquement l'interaction des conditions de fonctionnement de sorte que chaque partie en vue de la plurifonctionnalité. Il s'agit plutôt de maintenir les unités fonctionnelles et les fonctionnements indépendants les uns des autres: «l'ensemble se distingue des individus techniques en ce sens que la création d'un unique milieu associé est indésirable; l'ensemble comporte un certain nombre de dispositifs pour lutter contre cette création possible d'un unique milieu associé. Il évite la concrétisation intérieure des objets techniques qu'il contient, et n'utilise que les résultats de leur fonctionnement, sans autoriser l'interaction des conditionnements»(Simondon 2012, 79). Reprenons l'exemple du laboratoire, il faut par exemple y séparer certaines fonctions d'alimentation des instruments de mesure qui sont à alimenter parce que le fonctionnement qui rend possible l'alimentation peut produire du son, des ondes, des rayonnements qui risquent de gêner les instruments de mesure qui seront sensibles à ce bruit. Orchestrer des résultats de fonctionnement (alimentation, effectuation de la mesure) suppose un autre type de synchronisation entre les différents composants et les différentes unités fonctionnelles de l'objet technique.

8. L'IA comme ensemble technique

Le choix du terme de milieu, milieu associé et ensemble technique a des conséquences théoriques. En effet, le milieu associé implique que l'on se place dans le contexte d'un individu (technique) pour lequel il y a causalité récurrente et autorégulation simultanée, et où l'objet devient sa propre condition de possibilité, tandis que le milieu désigne cet espace d'échange plus vaste parce qu'il n'est pas défini par l'exigence de causalité récurrente et de plurifonctionnalité à l'intérieur d'un système, et que l'ensemble désigne un objet technique qui ne doit pas avoir un milieu associé pour fonctionner. Dans la mesure où le vivant a un milieu associé qu'il individue, et que nous avons associé la valeur de la créativité à son lien avec le vital comme problème, alors qualifier l'objet technique IA à l'intérieur des types singularisés permettra d'éclairer la nature des rapports entre le technologique et le vital, rapports qui sont brouillés par la métaphore organiciste.

Si l'individu technique est ce qui a un milieu associé, alors on peut se demander s'il est possible de parler d'individu technique pour une

IA. Or, il semblerait que la réponse soit non: sur le plan strictement technique, un algorithme de production qui fonctionne par apprentissage non-supervisé ne possède pas d'individualité - ni humaine ni technique en ce sens précis, donc -. Il n'a pas, en tant qu'objet, de cohérence interne. Non pas parce qu'il n'est pas cohérent, mais parce que les typologies simondoniennes de l'élément, l'individu et l'ensemble s'appliquent mal à ce type d'objet informatique, comme Y. Hui l'a souligné (Hui 2016, 76). Il écrit qu'«à première vue, on ne peut réutiliser le vocabulaire de Simondon pour comprendre les objets digitaux, parce que l'on n'y trouve pas un tel [celui des objets techniques] mécanisme de causalité réciproque dans les objets digitaux tels qu'ils garantissent leur propre stabilisation»(Hui 2016, 56). Hui va d'ailleurs plus loin: ayant identifié les objets digitaux comme un type d'objet industriel susceptible d'être digitalisé par objectification de data (des données concaténées en vidéo ou en fichier audio) ou par data-ification d'objet (Hui 2016, 50) (le répertoriage des objets grâce à leur insertion dans des réseaux de tags), il se donne les moyens de concevoir les programmes et les algorithmes non pas comme des objets digitaux, mais comme les milieux associés (qui correspond dans notre typologie au «milieu» simple) de ces objets (Hui 2016, 76). L'argument par lequel il défend cette définition est un argument *a fortiori*: puisque Simondon applique aux humains les catégories d'individu technique (et de milieu associé par voie de conséquence) pour nommer la place des humains dans le monde du travail préindustriel, alors il est également possible d'appliquer ces termes à des programmes et des algorithmes même si l'on ne peut pas directement calquer sur eux le modèle de la turbine de Guimbal, c'est-à-dire le modèle de l'individualité technique au sens étroit.

Toutefois, la piste dessinée par Hui nous conduit à appréhender le milieu associé des objets techniques digitaux comme les conditions de leur pertinence dans un espace sémantique donné, c'est-à-dire à partir des conditions de possibilité d'usage plutôt qu'à partir des conditions de possibilité *tout court*. Il écrit, par exemple, que «les objets digitaux assument également la fonction de maintenir des émotions, des atmosphères, du collectif, des souvenirs, etc.»(Hui 2016, 57). Ces fonctions sont des usages, pas des fonctions techniques. Elles déterminent les conditions de pertinence d'un objet *pour nous, humains*, plutôt que les conditions d'intégrité technologique. De sorte que cette condition de

pertinence, en tant que condition de possibilité d'usage, rapproche en réalité la notion de milieu dit associé qu'Hui utilise dans son texte à la notion d'ensemble. Elle implique non pas une autorégulation interne de l'objet, mais une séparation des différents aspects de l'objet technique informatique, pour articuler des résultats de fonctionnement en vue de la satisfaction de nos désirs.

La distinction entre le milieu et l'ensemble nous permet donc d'identifier deux types de liens possibles dans les objets techniques et naturels – sachant que les humains sont inclus dans cette dernière catégorie. La première correspond à un lien organique, celui de la rencontre synchrone et entre des conditions d'existence qui permettent à un objet de ne pas être autodestructif en vertu de son autorégulation, et la deuxième correspond à un lien de solidarité qui suppose l'absence de rencontre entre les conditions de fonctionnement et la rencontre entre les résultats des fonctionnements séparés.

Cela permet de saisir que la définition opératoire du métabolisme que propose Boden doit être enrichie d'une définition structurale des conditions d'opérativité de ce dernier, en particulier dans la mesure où les différences structurales permettent de saisir différents régimes d'opération métaboliques. Par exemple, les différences du milieu et du milieu associé qui impliquent une certaine réciprocité des fonctions, et celles de l'ensemble qui en impliquent l'hétérogénéité. Puisque nous enquêtons sur le sens de la créativité computationnelle en tant que processus concret, et par comparaison avec la créativité humaine, il est nécessaire d'identifier le type d'opération et le type de structure auxquels correspondent les «résultats de fonctionnement» des IAs dites créatives.

Il apparaît donc d'importance de rappeler la mise en garde de Simondon quant aux effets de l'anthropomorphisme et de la métaphore organiciste lorsqu'il s'agit des objets techniques. Comme le disait Parrochia, tâchons de ne pas confondre les abeilles et les abeilles de verre, le vivant et ses images (Parrochia 1994). Ce que Simondon apporte, c'est la possibilité de libérer la pensée d'un enfermement dans cette métaphore, qui rend difficile la construction d'autres hypothèses plus fécondes et à d'autres échelles.

S'interroger la possession ou non, pour un objet digital, d'un milieu ou d'un milieu associé, c'est mettre encore en question l'anthropomorphisation des programmes, et leur rapport avec la notion d'organisme.

Il a été établi plus haut que dans la relation de l'individuation d'un être vivant à son milieu, il existe une dimension spécifique de cette individuation, celle du virtuel. Ce virtuel procède de l'existence psychosociale des individus biologiques – c'est la raison pour laquelle il faut inviter à conjuguer les échelles de réflexions et à les pluraliser davantage que ne l'a fait Boden jusqu'ici, entre une échelle strictement individuelle et une échelle strictement collective – en rapport avec «l'histoire», ce qui est un peu vague. Le dernier pas que nous proposons d'effectuer à ce sujet consiste dans la réaffirmation de la proposition suivante: il semble nécessaire de ne pas séparer la question de la vie (artificielle ou non) de la question de la créativité. Si une vie artificielle robuste n'est (pour lors[13]) pas envisageable, alors la créativité telle que nous la connaissons ne l'est pas non plus. Ceci, parce que contrairement à ce que dit Boden, nous maintenons une conception de la créativité n'est pas uniquement l'affaire de propriétés d'un résultat: elle est toujours prise dans une situation *hic et nunc* qui rend non-superposable l'inédit et le créatif, quoi que ces derniers puissent se ressembler. La créativité a toujours un sens, fût-il inchoatif, pour cellui qui s'engage dans une activité créative – à titre d'exemple extrait de la typographie de cette phrase même, les recherches contemporaines en écriture inclusive (quelle typographie, quelle prononciation, quelles règles) sont créatives. Elles sont associées à des recherches en design qui supposent l'identification et la fabrication d'un problème. Ce sens n'existe pas pour les IAs, et la capacité de ces dernières à satisfaire nos désirs et requêtes (ou à refuser de le faire) procède, elle, d'un travail humain en amont dans la préparation des données, et en aval lorsque les programmes sont entraînés et leurs performances affinées par l'humain dans la boucle.

9. Là où se trouvent l'art et la culture

Si le feedback humain est nécessaire, ce n'est donc pas uniquement parce que cela accélère notoirement l'apprentissage des programmes et leurs alignements. Le feedback humain est utile pour accélérer l'ap-

13. Il ne s'agit à aucun moment pour nous de condamner dans l'impossibilité telle ou telle propriété – il s'agit plutôt de contribuer à recalibrer nos descriptions et nos attentes à hauteur de ce que nous sommes capables d'accomplir pour le moment.

prentissage et faire en sorte que les programmes pour qui c'est une propriété importante, comme les agents conversationnels, fassent illusion (Gratien 2023[14]; Delouya 2023[15]; Cole 2023[16]). Et le feedback humain est utile, par conséquent, parce que l'illusion finit par suffire (Zador *et al.* 2023, 2); les humains génèrent déjà des sentiments à l'égard de certains agents conversationnels tels que ceux proposés par la désormais célèbre entreprise Replika. C'est précisément parce que l'illusion suffit que les mises en garde prolifèrent, et que les projections anthropomorphiques fonctionnent si bien, alors même que lorsque nous échangeons avec quelqu'un, nous n'avons pas pour fonction première de prédire le prochain mot le plus probable pour poursuivre la phrase que nous avons commencée, ni de produire un résultat maximalement conforme à ce que notre interlocuteurice attend de nous.

Par conséquent, le mathématicien Du Sautoy énonce quelque chose d'important dans son ouvrage consacré à la créativité, *The creative code* (Du sautoy 2019, chap. 7), lorsqu'il écrit «*art is ultimately an expression of human free will, and until computers have their own version of this, art created by a computer will always be traceable back to a human desire to create*». La première proposition doit néanmoins être amendée, puisque nous avons entrepris de montrer que l'invention, et avec elle l'invention de contenu culturel et symbolique, répondaient moins d'une liberté métaphysique occidentale traditionnelle dont il faudrait prouver l'existence, que de l'effort par lequel on pose et répond à une problématique vitale qui peut advenir à l'échelle individuelle, groupale, collective voire universelle. Toutefois, tant que les artefacts à valeur culturelle seront produits pour les humains, ces artefacts trouveront toujours leur source dans le désir humain de créer. Ce sur quoi cela permet d'insister, c'est que l'invention humaine, les pratiques créatives et culturelles ont de la valeur parce qu'elles permettent de prendre en charge des problématiques vitales.

14. URL: https://www.slate.fr/story/243005/replika-fonctionnalite-erotique-roleplay-amants-virtuels-reconfort-amour-rupture.

15. URL: https://www.businessinsider.com/replika-chatbot-users-dont-like-nsfw-sexual-content-bans-2023-2.

16. URL: https://www.vice.com/en/article/z34d43/my-ai-is-sexually-harassing-me-replika-chatbot-nudes.

Rattachée comme un outil à cellui qui s'en sert, l'IA acquiert une autre valeur culturelle. Par exemple, Hans Ulrich Obrist cité par Du Sautoy défend l'idée qu'il n'est pas impossible que l'art que nous cherchons absolument à identifier dans les productions des IAs et des programmes informatiques soit en fait à trouver dans ces IAs et programmes informatiques eux-mêmes (Du sautoy 2019, chap. 8). Pour le dire en utilisant deux programmes célèbres: l'art est dans AARON[17] et dans The Painting Fool,[18] elle n'est pas dans leurs productions respectives. De la même manière, la créativité se manifeste dans le travail de recherche et d'expérimentation individuel et collectif, et dans toute l'externalisation du travail de feedback humain, en particulier depuis la mise à disposition de certaines IAs auprès du grand public. C'est cela qui permet de repérer les limites actuelles des programmes pour rendre possibles leurs améliorations. Cela suppose que des humains exercent leur créativité «malveillante» (Cropley *et al.* 2008), pour identifier les failles de sécurité ou de probité du système, ainsi que les quelques endroits où l'on parvient encore à pointer les limites de l'illusion.

L'invention, l'art et la culture sont dans les IAs elles-mêmes et dans l'écosystème humain qui les entoure. Il s'agit d'étudier ces IAs, et d'essayer d'instaurer de nouvelles relations entre elles et nous. Ces relations sont en acte, par le truchement d'un fonctionnement orienté par la nécessité de satisfaire une requête. C'est la formulation de cette requête qui devient l'enjeu culturel collectif véritable. Le prompt, la commande et tout ce que la formulation d'une requête esthétique auprès d'un programme à qui on en aura délégué la production réclame des personnes une inflexion nouvelle du langage du désir et des intentions humaines. C'est une inflexion par laquelle on apprend à déterminer son désir sous la forme d'un texte qui suscitera ensuite un résultat qui lui sera aussi adéquat que possible, et ce indépendamment d'un geste d'exécution dont on se déleste de la charge.

Deux options se présentent alors: la première, faire du prompt lui-même le contenu créatif, dont la valeur culturelle résiderait tou-

17. Programme d'IA symbolique (qui applique des règles strictes et déterminées de la forme si x alors y) développé durant plusieurs décennies par Harold Cohen.

18. "Artiste digital" de Simon Colton, dont l'agentivité a été interrogée, car il se serait une fois mis en route en dehors de ses temps de programmation pour une raison inconnue.

jours aussi dans l'effort humain de spécification et de communication d'un désir, aussi trivial soit-il. La deuxième serait de faire du résultat sélectionné une œuvre, soit parce que l'alignement entre le prompt et le résultat est parfaitement accompli, soit parce qu'il ne l'est pas et que cet écart est susceptible de devenir un nouvel objet d'études, scientifique ou esthétique.

10. La créativité sans la problématique vitale? Sur la créativité induite

À quoi correspond la résolution esthétique d'une problématique vitale? C'est-à-dire, comment au juste penser ce à quoi correspond l'esthétique chez Simondon? Entre le cours de 1965-66 *Imagination et invention* et la troisième partie de sa thèse secondaire *Du mode d'existence des objets techniques*, nous pouvons définir un ensemble de choses.

L'esthétique est le mode de résolution d'une tension dans le rapport entre les humains et le monde modalisée par un écart entre le sens du singulier (le technique) et le sens du tout (le religieux), sur un mode pré-noétique, antérieur donc à la pensée (Simondon 2012, 218). La catégorie d'esthétique est très générale dans son effectivité, et ne fait pas référence à l'art tel qu'il est institutionnalisé dans les espaces essentiellement occidentaux dans lesquels telle ou telle pratique acquiert ses lettres de noblesse. Simondon précise plutôt que cette forme institutionnelle, aussi construite qu'elle soit, suppose pour exister et pour se maintenir dans l'existence quelque chose comme sa rencontre avec une «tendance fondamentale de l'être humain», celle de chercher «le complément par rapport à la totalité» (Simondon 2012, 248). Ce sens de la totalité permet de concrétiser la rencontre entre le général et le particulier, le théorique et le pratique, l'universel et l'*hic et nunc*, dans une mise en relation des contraires qui fait de ce qui suscite une «impression esthétique», terme que les études simondoniennes doivent encore éclairer, un «point remarquable de la réalité vécue» (Simondon 2012, 249). L'invention dans le domaine de l'esthétique correspondrait à la résolution d'un problème qui serait celui de l'épreuve d'un écart (c'est le hiatus) entre ce sens du total, de l'ici et du maintenant, dont la résolution serait l'introduction d'une nouvelle réticulation possible de ce rapport au monde, qui prend sens pour un humain à l'intérieur de la situation vécue. C'est l'épreuve de cet écart qui serait, pour quel-

qu'un qui entreprendrait quelque chose avec Midjourney par exemple, à l'origine de ses prompts.

Sans nécessairement admettre l'ensemble de l'édifice de l'esthétique simondonienne, on peut alors comprendre que, si le sens de l'écart entre le tout et l'organisme concret qui vit ici et maintenant est éprouvé *par un organisme* pour lequel il y a épreuve de ce type d'écart, alors c'est l'épreuve de cet écart qui justifie l'existence même des pratiques de création dans le domaine de l'esthétique et qui leur confère une valeur. De sorte qu'encore une fois, non seulement l'exercice, mais aussi le sens de la créativité supposent l'existence d'un organisme qui est une relation individuante entre soi et son milieu – ce qui vaut d'ailleurs à l'impression esthétique de devenir un peu plus loin dans le texte la «réalité esthétique» (Simondon 2012, 252), jamais vraiment séparée du monde ou de l'humain, toujours insérée, toujours située. Or, comme l'écrit Simondon, «l'objet technique se distingue de l'être naturel en ce qu'il ne fait pas partie du monde» (Simondon 2012, 235). Nous pointons leur absence d'insertion dans un monde pour mettre en cause la pertinence de la proposition qui consiste à attribuer à des IAs quelque chose de l'ordre de la créativité.

Selon ce chemin, on aboutit à ce qu'écrit Du Sautoy: «*At the moment all the creativity in machines is being initiated and driven by the human code. We are not seeing machines compelled to express themselves. They don't really seem to have anything to say beyond what we are getting them to do. They are the ventriloquist's dummy providing the mouthpiece for our urge still to express ourselves. And that creative urge is an expression of our belief in free will. […] creative activity was meant to capture humans' attempts to understand being in the world* (Du sautoy 2019, chap. 16)». Nous avons proposé au cours de cet article d'amender un peu cette proposition en substituant la volonté libre au concept simondonien de métabolisme dans lequel on pouvait donner un sens et une explication à l'esthétique et à la créativité, qui permette non seulement d'écarter l'hypothèse que les LLM puissent produire du créatif d'une part, et de ne pas fonder la créativité dans une sorte de privilège humain qui viendrait justifier par ailleurs une coupure anthropologique qui nous réinstallerait dans les coordonnées surannées de la métaphysique moderne – puisque la créativité est une propriété

du vivant en général (Simondon 2005, 334).[19] L'image du ventriloque me semble être le complément adéquat à celle des personnages de *Huis-Clos*, en ce qu'elle permet de rappeler que les intentions que le désir de création est bien celui de l'être humain qui se sert des IAs plus que les IAs elles-mêmes.

Ce qui signifie qu'il ne s'agit pas de nier que les IAs puissent contribuer à la culture humaine y compris la production de symboles. Elles le pourront sur un mode combinatoire.[20] Or la combinatoire ne paraît pas suffisante pour définir la conduite créative: la combinatoire pose le problème à un niveau trop général, car: l'«intelligence créative [...] ne se borne pas à un constat de la situation avec les propriétés données, mais ajoute une médiation fournie par des propriétés latentes»(Simondon 2005, 335).[21] Entendons par là que l'échelle d'apparition de conduites créatives est l'échelle du particulier, *hic et nunc*, échelle à laquelle les données du problème et les normes de l'action peuvent être transformées. C'est moins la combinaison en elle-même qui est importante, que la norme de laquelle la combinaison procède: une médiation entre propriétés latentes et aspects d'une situation. En ce sens du terme, la

19. Simondon de manque pas d'y insister: dans un petit cours donné sur la créativité et l'invention, c'est depuis l'analyse de la capacité ou non de faire des détours que l'auteur statue sur la possibilité pour un organisme d'être ou non créatif. Il est conduit, pas là, à dire que les enfants ne sont pas immédiatement des êtres créatifs, c'est-à-dire capables d'effectuer un détour dans l'accomplissement d'une tâche. Autrement dit, la créativité n'est pas une propriété substantielle attribuée aux humains depuis les cieux: c'est une propriété qui peut être développée qui suppose d'être engagé dans le monde.

20. Auteur du *Manifeste pour l'Art Permutationnel* en 1952, le professeur de psychosociologie et cybernéticien Abraham A. Moles écrivait d'ailleurs dans *Art et ordinateur* que ses méthodes ne s'appliquaient qu'à des symboles mis dans un répertoire fini et établi. *In* Moles A., *Art et Ordinateur*, Blusson, Paris, 1990, pp. 155,158.

21. Simondon déduit cette proposition de l'analyse de la conduite créative suivante: un enseignant cherche à ouvrir un paquet de polycopiés tenus par une ficelle, demande aux étudiant·e·s si ils et elles auraient un couteau à prêter, et une étudiante répond qu'elle a un briquet. Il y a deux éléments ici: la personne a proposé une bonne manière de résoudre le problème depuis une donnée qui ne faisait pas partie de celles dans lesquelles le problème était posé. Cela suppose une compréhension de la question qui la reformule complètement, de possédez-vous tel objet à avez-vous un moyen d'accomplir la libération du paquet de polycopiés du joug de la ficelle? L'étudiante s'est également détournée du protocole normatif d'usage du briquet pour allumer des cigarettes, pour le réorienter vers la découpe de ficelle.

contribution des IAs à la vie culturelle et symbolique ne sera pas une contribution créative.

Elle pourra l'être en un autre sens affaibli, proposé par Abraham A. Moles dans *Art et Ordinateur*. Traitant de l'art permutationnel, qui est un art combinatoire entre variables d'un répertoire donné, l'auteur propose de nommer l'œuvre permutationnelle comme un jeu entre deux domaines, le choix des éléments et l'algorithme combinatoire. La rencontre entre ces deux éléments instaure un champ des possibles (déterminé par le répertoire et ses règles de composition) qu'il ne «reste plus qu'à explorer», écrit Moles, «dans une *créativité induite*»(Moles 1990, 119). C'est cette créativité induite, instaurée par le haut depuis la détermination humainement supervisée du répertoire et des règles de composition (en tant qu'elles sont vérifiées et modifiées par des humains), peut éventuellement concerner les IAs, à qui on déléguerait une fonction d'exploration des combinaisons possibles d'un répertoire donné selon certaines règles. Cette créativité est applicable parce qu'elle est une opération déterminée d'exploration, qui n'implique pas la rencontre d'un problème ni d'une situation dont la résolution nécessiterait un détour.

La dimension du sens reste, dans ce cadre de la créativité induite, arrimée aux humains qui créent, ou promptent les programmes. Pour ce qui nous concerne, toutes choses étant égales par ailleurs, nous maintiendrons pour l'instant que la véritable pratique créative se trouvera dans les prompts, la genèse technique des modèles et la recherche de leurs failles en vue de leur amélioration, plutôt que dans les résultats produits par les prompts. Nous maintiendrons également que les IAs ont bien une valeur culturelle, parce qu'elles ont la vertu heuristique de permettre d'explorer le langage naturel et nos représentations sur un mode qui n'est pas le nôtre parce que nous ne sommes pas des *transformers* de séquence et que le monde n'est pas un vecteur (Cardon *et al.* 2018, part. 4.3). Nous apprendrons par-là plus de choses sur la manière dont nous ne comprenons pas le monde et dont ne nous faisons pas sens des choses, que sur la manière dont nous faisons sens de nous-mêmes et du monde.[22] Mais cette exploration et ces combi-

22. Ce que les plus grands chantres de l'artificialisation et la digitalisation capitalistique du monde comme Kai-Fu Lee défendent dans des ouvrages grand public par

natoires, pour évoquer deux termes bodeniens, ne sont et ne restent signifiantes que pour nous.

Par conséquent, il est fallacieux de défendre la valeur et l'intérêt des technologies qui rendent possibles les agents conversationnels, la génération de texte, d'image ou de musique depuis le concept de créativité. Cette stratégie argumentative repose à mon sens sur des représentations collectives. L'une d'entre elles renvoie au lien cartésien établi entre l'intelligence et la créativité. Il est celui sur lequel capitalisent celles et ceux qui entendent financer la recherche en IA et éviter un hiver comme la prise de parole de James Lighthill en 1973 (Lighthill 1973[23]) en avait suscité. Ce que l'on appelle hiver ici, c'est la chute drastique des financements dans la recherche en intelligence artificielle (connexionniste), à la suite de la publication du rapport de Lighthill. Le fonctionnement de cette rhétorique est le suivant: on défend la valeur du développement des IAs et l'intérêt de financer le développement des LLM et LLM multimodaux en faisant valoir leur intelligence. Or, l'intelligence est associée à la créativité et l'usage créatif du langage (Descartes 1999, 73). Par conséquent, la créativité peut être considérée comme la preuve de l'intelligence, et l'intelligence comme ce qui doit avoir pour conséquence la créativité. C'est ce lien entre intelligence, créativité et valorisation préalable de l'intelligence comme le point à atteindre des objets techniques informatiques fabriqués qui donne son sens à toute cette quête, en plus de la possibilité d'automatiser certaines tâches qui requièrent des compétences artistiques de conception telles que le design de logos, d'applications et d'interfaces, et peut-être de contribuer à alimenter la pyramide de Ponzi des NFT et du web3 (Nikkidale 2023[24]). Grâce à notre détour simondonien, nous avons entrepris de sectionner ce lien entre intelligence et créativité pour subordonner, depuis les concepts d'invention et de problème, la créativité non plus à l'intelligence, mais à la métabolisation comme

ailleurs. Qiufan C. *et al.*, *I.A 2042 – Dix scénarios pour notre futur*, Paris, Les Arènes, 2022, p. 128, 132.

23. URL: https://rodsmith.nz/wp-content/uploads/Lighthill_1973_Report.pdf.

24. URL: https://startistsociety.com/aiart/. Dans ce podcast, les deux personnes discutent de la manière dont l'une d'entre elles a généré des animaux steampunks avec Midjourney et pensait les vendre en NFT. L'enjeu ici étant moins la création elle-même qu'un design «cool» susceptible de s'insérer sur un marché dont le «cool» l'une des valeurs motrices

résolution inventive d'un problème, celui d'une relation individuante avec un milieu.

11. Le coût humain des stratégies de développement de l'acceptabilité sociale d'un objet technique

Là où Boden proposait de définir la créativité comme la propriété d'un résultat, nous avons proposé une approche plus processuelle et holistique du concept. Pour montrer que seule la créativité induite était applicable aux IAs, nous avons pu insister, grâce au concept de datafication des objets de Yuk Hui, sur la manière dont le sens était saisi et produit en étroite dépendance avec un travail humain préalable de classification. Le tri et la labellisation des données d'entraînements sont des opérations centrales à l'efficacité des IAs, elles impliquent encore du travail humain en dépit des progrès accomplis dans le domaine. Mais surtout, c'est encore sur décision humaine que l'entraînement est arrêté, que l'IA est jugée déployable et déployée. Plus encore, c'est sur du travail humain que l'inscription d'un régime normatif des usages et des affordances de ces IAs repose. Encore une fois, c'est pour l'effort créatif humain que les outputs des machines prennent un sens, et c'est selon des valeurs humaines que ces outputs sont évalués.

Ce que cela signifie, c'est que la valeur des «co-créations» est attribuée et régulée par les humains et repose, puisque les données de travail sont des productions humaines, sur les biais et les valeurs humaines exprimées dans les données d'entraînement. On retrouve la normativité humaine dans l'objet technique, non seulement dans les conditions de possibilité d'usage qui sont les siennes, telles qu'on les a définies avec Y. Hui, mais aussi au sein du système technique industriel qui constitue la condition de possibilité matérielle de tels objets techniques. Il convient par conséquent de se faire politique ici.

D'abord, tous les produits d'Open AI, de Google Bard, ou de Midjourney sont dépendants d'une armée de modérateurices. Pour Google et OpenAI au moins, elle est constituée par des travailleurs et des travailleuses du tiers monde recruté·e·s par l'entreprise Sama, et ce jusqu'au début de cette année. Sama, qui revendique une intelligence artificielle «éthique» a à peine 1.5 dollar de l'heure comme l'a révélé

le journal TIME,[25] tout en traumatisant des travailleur·se·s, qui lisent des horreurs ou les voient jusqu'à 9h par jour afin de les classer, un «processus nécessaire pour rendre l'IA plus sûre», selon OpenAI. Plus sûr pour les occidentaux qui utilisent ces technologies; leur confort aura été garanti par transfert de risque sur une population que l'on aura eu aucune vergogne à mettre en danger à notre place. C'est une condition de possibilité d'usage qui ne peut pas faire l'objet de négligence: le coût humain de nos questionnements philosophiques relatif aux objets techniques est un sujet qui nous concerne au premier chef.

Deuxième élément, la normativité humaine est toujours requise au bon fonctionnement des IAs, et c'est par feedback humain que nous garantissons ce bon fonctionnement le mieux pour l'instant. De sorte que l'espace conceptuel, c'est-à-dire l'espace des possibilités ouvertes selon des règles données, pour reprendre un terme de Boden, est un espace toujours déjà humainement normé, mais où le mode d'existence de la norme en question est un mode mort: le biais est intégré, le mot-clef interdit, et c'est trop tard – comme il est trop tard pour Inès, Estelle ou Garcin. L'espace conceptuel est ici déterminé en extension par les régularités statistiques des associations de *tokens*, qui contribuent à renforcer les associations que nous produisons déjà, comme le souligne Jenny Nicholson dans son article consacré aux biais sexistes de Chat GPT (Nicholson 2022[26]).

Troisième élément, Midjourney fonctionne sur une appropriation des travaux d'artistes qui n'ont pas consenti à voir leurs œuvres intégrées aux données d'entraînement de cette IA. Cette disruption de la propriété intellectuelle par noyade à l'intérieur de quantités colossales de données reste une disruption de la propriété intellectuelle. Cela n'est pas anodinsi l'on suit Boden, la créativité est jugée depuis les paramètres psychologiques et historiques qui font de la créativité un jugement relatif et subjectif. Relatif à un état de connaissance individuelle, ou relatif à un état général de la connaissance. Si cette connaissance est empêchée par l'inaccessibilité des données d'entraînement ou du pro-

25. «Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer» [en ligne], *Time*, 2023, URL: https://time.com/6247678/openai-chatgpt-kenya-workers/, consulté le 28 mars 2023.

26. URL: https://medium.com/madebymckinney/the-gender-bias-inside-gpt-3-748404a3a96c.

gramme lui-même, pour des raisons relatives à la gestion de l'entreprise ou pour des raisons de complexité des programmes, alors il devient difficile de produire la comparaison qui permet de rendre un jugement sur la créativité d'un résultat. En effet, si l'on met à disposition ces données, on peut au moins objectiver une situation depuis laquelle établir une échelle selon laquelle évaluer la nouveauté, la surprise, voire le détour.

On peut alors conclure deux choses. Premièrement, l'omniprésence des normes de vie des groupes humains telles qu'elles ont pu se cristalliser dans les discours publiés sur le *world wide web*, et la matérialisation de ces normes de vie dans des pratiques de modération de contenu implique l'existence assumée d'une valorisation extrinsèque des productions des IAs. Seules certaines d'entre elles passeront le seuil de la publicité. C'est la raison pour laquelle le rapport entre invention et valeur est un rapport que, contrairement à Boden, nous n'abandonnons pas. Deuxièmement, la position d'un problème et l'identification de solutions possibles impliquent de l'axiologie. C'est sur ce point que nous terminerons, en ouvrant quelques pistes de recherches depuis l'hypothèse de la nécessité de mettre en regard du développement d'IAs susceptibles de co-création avec des individus ou groupes humains avec les efforts parallèles d'automatiser l'évaluation esthétique.

12. À production automatisée, évaluation automatisée? La double délégation de la production et de l'évaluation des objets d'expérience esthétique

Les IAs par lesquelles on travaille à favoriser l'acceptabilité sociale de ces nouveaux objets techniques à l'intérieur de notre écosystème digital sont les IAs qui produisent des artefacts à dimension culturelle symbolique: un texte – fut-ce une lettre de motivation –, un logo, une image. Or, qualifier une chose de créative est en soi une évaluation (Boden 2004, 76[27]). C'est l'évaluation des propriétés de surprise et de nouveauté. La première chose qu'il faut faire valoir à ce niveau,

27. «*Creativity, whether in children or adults, involves evaluation. The new idea must be compared to some pre-existing mental structure, and judged to be 'interesting' by the relevant criteria. People who can evaluate their own novel ideas will accept them or (sometimes) correct them, but will often be unable to explain in just what way they are*

c'est que sans accès aux données de bases ni au régime normal de fonctionnement de l'objet technique, il paraît impossible d'évaluer la créativité d'une production depuis sont producteur plutôt que depuis nous-mêmes, à la réception du produit. En effet, si l'on cherche à identifier la capacité créative d'un individu, humain, animal ou mécanique, il est nécessaire de savoir quelle est la condition initiale, quel problème on pose et d'identifier les stratégies de résolution. Par exemple, Simondon dans un petit cours sur la créativité (Simondon 2005, 332-343) se réfère à un exemple développé par Osborn d'un ingénieur à qui l'on avait demandé de construire une machine à emballer des pneumatiques, si cela était bien nécessaire. En d'autres termes, il met en cause la norme même de l'action, l'objectif que ce corps professionnel s'est donné. Il y a nouveauté et surprise, créativité dans cette proposition parce que la personne résout le problème initial en détruisant l'un des paramètres dans lesquels le problème était donné, à savoir la nécessité de concevoir et produire l'objet technique servant à emballer des pneumatiques. Par conséquent, le raisonnement que je propose est le suivant: s'il faut, pour évaluer la créativité, avoir accès à de quoi établir un protocole d'évaluation de cette créativité, qui ne peut pas reposer sur notre seule appréhension d'un *output*, sachant que nous ne connaissons ni les données de départ, ni le programme.[28] Pour mettre l'évaluation à niveau, on peut supposer qu'il faille par conséquent trouver un moyen d'automatiser l'évaluation de la créativité de ces productions, afin que la puissance de calcul qui permette d'évaluer la capacité créative d'une IA donnée atteigne un degré pertinent pour l'analyse. Comment faudrait-il procéder?

13. Les critères de Birkhoff: conception objectiviste

Une première piste de réflexion serait celle de l'objectivation de critères esthétiques mesurables hérités des théoriciens de l'information, à savoir un rapport riche entre l'ordre et la complexité tels que ce rapport a été théorisé par Birkhoff en 1933.

---

interesting. We need to understand why this is, and how the ability to map and explore aspects of one's own mind develops in the first place»

28. Sauf à admettre une hypothèse behavioriste à l'égard des IAs, ce qui supposerait qu'on leur accorde, en premier lieu, un comportement.

Dans cette mesure rationaliste, la référence à un contenu sémantique donné est absente. Il ne s'agit pas de mesurer le sens, mais de produire une mesure synthétique de la qualité depuis deux paramètres objectivement mesurables et déjà préalablement valorisés: l'effort attentionnel requis pour saisir l'objet de la perception (la complexité), et la saisie de l'ordre d'un objet sont responsables d'une évaluation intuitive du sens de la valeur de l'objet de la perception selon Birkhoff (Birkhoff 1933, 3-4). Dans la mesure où l'ordre et la complexité sont définis sur le plan formel depuis les notions de redondance, d'ambiguïté, la similitude, le contraste, et sur le nombre de notes qui compose un morceau et fait varier la perception et l'activation cérébrale (Birkhoff 1933, 6,9), qui ont des mesures objectives, on appellera objectiviste cette théorie de la valeur esthétique.

Or, si on l'applique aujourd'hui aux productions générées grâce à GPT-3 depuis un prompt envoyé à Chat GPT, on constate rapidement certaines limites structurales à l'application de ce critère. D'abord, Chat GPT a une fenêtre de contexte (c'est-à-dire une mémoire) de 1500 *tokens*, qui empêche de produire des structurations complexes qui peuvent être celles d'un long roman par exemple. Raison pour laquelle on a essayé d'étudier la valeur de ses propositions sur des haïkus (Hitsuwari *et al.* 2023[29]) ou des pièces de théâtre. Ce n'est pas la limite en elle-même qui pose problème: il est évident qu'il n'y a aucun intérêt à évaluer le succès de l'accomplissement d'une tâche (comme écrire un roman) que l'IA n'est pas équipée pour accomplir. Mais cette proposition a un revers sur lequel il faut encore insister: prétendre que l'IA est ce qu'elle n'est pas ou peut être interprétée avec des catégories auxquelles ses opérations ne correspondent pas n'a pas d'intérêt non plus. Ce qui importe, c'est de dire que si l'on adopte la posture objectiviste met en lumière la limitation de la créativité combinatoire potentielle en vertu de la fenêtre de contexte dont, en l'occurrence, l'agent conversationnel dispose. Ensuite, le processus de production *token* après *token* et prédiction après prédiction empêche la complexité créative associée au travail d'écriture constitué d'allers-retours, par lesquels on laisse la suite du texte influer sur le début du

29. URL: https://www.sciencedirect.com/science/article/pii/S0747563222003223.

texte. Autrement dit, la production textuelle ici empêche un usage non linéaire et non limité en *tokens* de la mémoire.

C'est un régime d'évaluation qui trouve ici une limite, parce qu'il évacue encore la question du sens des œuvres. S'il est intéressant de mesurer le rapport entre ordre et complexité dans une œuvre, un courant artistique, une manière de peindre, d'orner, de sculpter, c'est moins pour pouvoir produire ces mesures seules que pour essayer de donner un sens à ces mesures, de les insérer dans l'individuation d'un rapport au monde. Cette dynamique, vitale, est absente de cet objectivisme birkhoffien, et elle oblige à externaliser partiellement la valeur des œuvres et des productions et pratiques culturelles en général, en la décorrélant également du plaisir que nous prendrions, simplement, à saisir empiriquement un ratio ordre/complexité.

Que faudrait-il faire, alors, pour saisir cette valeur? Serait-il possible d'adopter une position plus subjectiviste et essayer de modéliser l'être humain en position d'évaluation?

14. Conception subjectiviste de la valeur avec objectivation du sujet: simulations de spectateurices

C'est notre deuxième piste, et la voie qu'adopte depuis quelques années les «subjectivistes» par Fan et collaborateurs dans *Measuring and Evaluating the Visual Complexity of Chinese Ink Paintings* (Fan *et al.* 2022.[30] Les auteurs entreprennent de reproduire les mouvements de l'œil qui traque les formes et leur complexité dans le domaine de la perception visuelle. Cette fois, il s'agit donc de data-ifier quelque chose que l'on objective dans le même geste, à savoir les trajectoires du regard d'un observateur ou d'une observatrice en train de regarder une image – sachant que l'intérêt des chercheurs s'est porté sur le rôle du blanc dans le tracé du parcours du regard, afin d'établir un rapport entre la trajectoire du regard et la complexité perçue. La mesure esthétique est ici objectivée comme mesure de la complexité depuis le regard d'un individu moyen et non plus depuis une mesure objective de quantité d'information issue de la théorie de l'information.

30. Une publication datée de mars 2021.

Dans le domaine du spectacle vivant, c'est aussi cette technique qui est utilisée pour essayer de fournir à un robot un sens esthétique de ses propres mouvements lorsqu'il les produit sur de la musique. Si le *machine learning* intervient évidemment dans le processus, l'apprentissage est supervisé, et ce qui est numérisé correspond aux appréciations esthétiques fournies par des professionnels de sorte qu'elles puissent être intégrées par le programme (Huapeng *et al.* 2019, 6). L'évaluation est construite depuis l'évaluation d'une danse en particulier, le hip-hop, qui repose sur un certain nombre de poses et sur une certaine discontinuité qui permet par ailleurs aux chercheurs d'isoler des états, les poses, sur lesquelles travailler.

L'objectif est par conséquent de reproduire les normes esthétiques humaines et de les intégrer au programme, afin de pouvoir prédire/attester de la qualité esthétique d'une image, que ce soit sur le plan de son adéquation avec ce à quoi elle est censée faire référence (Ding *et al.* 2022, 15[31]) ou sur le plan de son intérêt perceptif *tout court*. Cela permettrait au programme de développer une compétence esthétique qui aurait un sens pour les humains et qui par conséquent resterait complètement anthropocentrée.

Conclusion

Cet anthropocentrisme auquel on n'échappe pas dans la recherche en créativité computationnelle nous conduit à achever ce texte en ouvrant sur un questionnement relatif au sens même de ces pratiques, sans toutefois nier leur intérêt scientifique et heuristique. Imaginons un instant que nous parvenions effectivement à automatiser à la fois la production et l'évaluation des contenus culturels et symboliques à un temps *t*, et ce de manière parfaite. À quoi aura servi toute cette agitation, sachant que les normes instaurées par feedback humain et labellisation dans les programmes, ou que les programmes auraient saisies depuis leurs données d'entraînement, seraient vite caduques et qu'il nous faudrait régulièrement mettre à jour (sauf à soutenir que les normes de vie sont prédictibles)? Pour revenir aux coordonnées dans lesquelles nous avons proposé de penser la créativité computationnelle,

31. URL:http://arxiv.org/abs/2204.14217.

que reste-t-il des normes de vies sans la vie qu'elles permettent de métaboliser et que nous avons placée au fondement de l'intérêt de la créativité, si elles sont détachées du sens de leur invention? Le milieu humain est et reste le milieu associé (au sens étroit) de cet ensemble technique que sont les IAs, même lorsqu'elles créent des objets appréciables esthétiquement, c'est-à-dire avec nos sens. C'est encore l'humain qui assure la fonction de régulation.

La créativité induite et combinatoire des IAs, est la créativité du personnage de *Huis-Clos*: elle est combinatoire, et exploratoire. Elle n'offre en ce sens aucune possibilité de détour ni de réforme des données de bases d'un problème depuis un cas particulier, comme on a pu en trouver la description chez Simondon. Il est nécessaire de pointer cette restriction de la créativité computationnelle à la créativité induite, pour ne pas se laisser prendre par des narratifs anthropomorphiques. Ces narratifs répondent avant tout d'une rhétorique commerciale qui parie qu'il y a là une voie pour augmenter l'acceptabilité sociale de ces programmes et de leurs contreparties si l'on écoute les discours émis par les entreprises qui «offrent», vendent ou louent ces IAs sur le web. Dans ces conditions, il nous aura importé de ne pas contribuer à entretenir un enthousiasme à l'égard d'objets techniques dont le coût humain de production et de fonctionnement nous paraît démesurément élevé.

De surcroît, même en admettant que l'on parvienne à automatiser la production et l'évaluation de contenus culturels et symboliques, il nous semble que si cela était le cas, l'intérêt que l'on porterait à cet ensemble technique *pour ses productions esthétiques* s'éteindrait immédiatement, et à bon droit. Car l'intérêt des IAs semble moins résider dans l'hypothèse de pouvoir les rendre créatives que dans ce que l'écriture informatique construit un rapport au monde, à la langue et au percept différent du nôtre, dès lors que nous ne sommes pas computationnalistes. Pour l'heure, le discours de légitimation des IAs servent à défendre auprès du grand public la production à peu de frais (sinon un peu de cobalt, sa pollution et ses morts associés) (Kara 2023) de logos, de chartes graphiques, de *memes*, de publicités ou encore de peuplement du *metaverse*. Rien de bien intéressant pour qui n'est pas convaincu par le monde tel qu'il est.

Bibliographie

Alderton Z. 2014, *'Snapewives' and 'Snapeism': A Fiction-Based Religion within the Harry Potter Fandom*, "Religions", 5(1), Multidisciplinary Digital Publishing Institute, pp. 219-267.

Birkhoff G. D. 1933, *Aesthetic Measure*, Harvard University Press, Cambridge (Mass.).

Blut M., Wang C., Wünderlich N. V., *et al.* 2021, *Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI*, "Journal of the Academy of Marketing Science", 49(4), pp. 632-658.

Boden M.A. 1999, *Is Metabolism Necessary?*, 'The British Journal for the Philosophy of Science", 50(2), The University of Chicago Press, pp. 231-248.

Boden M. A. 2004, *The creative mind: myths and mechanisms*, Routledge, London/New York.

Cardon D., Cointet J.P., Mazières A. 2018, *La revanche des neurones. L'invention des machines inductives et la controverse de l'intelligence artificielle*, "Réseaux", 211(5), La Découverte, pp. 173-220.

Cole S. 2023, *'My AI Is Sexually Harassing Me': R*eplika Users Say the Chatbot Has Gotten Way Too Horny, [en ligne], URL: https://www.vice.com/en/article/z34d43/my-ai-is-sexually-harassing-me-replika-chatbot-nudes.

Cropley D. H., Kaufman J. C., Cropley A. J. 2008, *Malevolent creativity: A functional model of creativity in terrorism and crime*,"Creativity Research Journal", 20, Taylor & Francis, pp. 105-115.

Delouya S. 2023, *Replika users say they fell in love with their AI chatbots, until a software update made them seem less human*, "Business Insider", [en ligne], URL: https://www.businessinsider.com/replika-chatbot-users-dont-like-nsfw-sexual-content-bans-2023-2.

Descartes R. 1999, *Le Discours de la Méthode*, Numilog, coll. «Les Classiques».

Devillers L. 2020, *Les robots émotionnels*, Editions de l'Observatoire, Paris.

Ding M., Zheng W., Hong W., *et al.* 2022, *CogView2: Faster and Better Text-to-Image Generation via Hierarchical Transformers*, [en ligne], "arXiv", URL : http://arxiv.org/abs/2204.14217.

Du sautoy M. 2019, *The Creativity Code, How AI is Learning to Write, Paint and Think*, 4th Estate, London.

Fan Z., Li Y., Zhang K., *et al.* 2022, *Measuring and Evaluating the Visual Complexity Of Chinese Ink Paintings*, "The Computer Journal", 65(8), pp. 1964-1976.

Firth J., Albery G., Beck K., *et al.* 2020, *Analysing the Social Spread of Behaviour: Integrating Complex Contagions into Network Based Diffusions*, [Pre-print].

Gervas P. 2009, *Computational Approaches to Storytelling and Creativity*, "AI Magazine", 30(3), p. 49.

Gratien A. 2023, *Privés d'amants numériques, les usagers de Replika ont le cœur brisé*, [en ligne], "Slate.fr", URL :https://www.slate.fr/story/243005/replika-fonctionnalite-erotique-roleplay-amants-virtuels-reconfort-amour-rupture.

Hitsuwari J., Ueda Y., Yun W., *et al.* 2023, *Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry*, [en ligne], "Computers in Human Behavior", 139, URL : https://www.sciencedirect.com/science/article/pii/S0747563222003223.

Hua peng L., Jing H., *et al.* 2019, *Feature fusion based automatic aesthetics evaluation of robotic dance poses*, "Robotics and Autonomous

Systems", 111, pp. 99-109.

Hui Y. 2016, *On the Existence of Digital Objects*, University of Minnesota Press, Minnesota.

Kara S. 2023, *Cobalt red: how the blood of the Congo powers our lives*, St. Martin's Press, New York.

Li J. 2018, *On the Possibility of Strong Artificial Life*, "Open Journal of Philosophy", 8, pp. 495-505.

Li M., Suh A. 2021, *Machinelike or Humanlike? A Literature Review of Anthropomorphism in AI-Enabled Technology*, [en ligne], dans "Proceedings of the 54th Hawaii International Conference on System Sciences", présenté à 54th Hawaii International Conference on System Sciences (HICSS 2021), pp. 4053-4062, URL : https://scholars.cityu.edu.hk/en/publications/machinelike-or-humanlike-a-literature-review-of-anthropomorphism-in-aienabled-technology(7c9fa6e3-ab93-4d72-ab37-219f2733dd1e).html.

Lighthill J. 1973, *Artificial Intelligence: A General Survey* [en ligne], "Science Research Council", URL : https://rodsmith.nz/wp-content/uploads/Lighthill_1973_Report.pdf.

McCulloch W. S., Pitts W. 1990, *A logical calculus of the ideas immanent in nervous activity*, "Bulletin of Mathematical Biology", 52(1/2), pp. 99-115.

Moles A. A. 1990, *Art et Ordinateur*, Blusson, Paris.

Nicholson J. 2022, *The Gender Bias Inside GPT-3*, [en ligne], URL: https://medium.com/madebymckinney/the-gender-bias-inside-gpt-3-748404a3a96c.

Nikkidale 2023, *88 – Will AI Art Replace Artists?*, [en ligne], URL: https://startistsociety.com/aiart/.

Page C. 2023, *Hackers stole passwords of Worldcoin Orb operators*, [en

ligne], URL: https://techcrunch.com/2023/05/12/hackers-stole-passwords-of-worldcoin-orb-operators/.

Pareto V. 1981, *Principes généraux*, "Manuel d'économie politique", Librairie Droz, coll. «Travaux de Sciences Sociales», Genève, pp. 1-39.

Id. 1964, *Principes d'économie politique pure*, "Cours d'économie politique", Librairie Droz, coll. «Travaux de Sciences Sociales», Genève, pp. 1-73.

Parrochia D. 1994, *Le statut épistémologique de la "vie artificielle"*, "Ordre biologique ordre technologique", Champ Vallon, coll. «Collection milieux».

Patterson D., Gonzalez J., Le Q., *et al.*, *Carbon Emissions and Large Neural Network Training*.

Piccinini G. 2010, *The Mind as Neural Software? Understand ing Functionalism, Computationalism, and Computational Functionalism*, "Philosophy and Phenomenological Research", 81(2), [International Phenomenological Society, Philosophy and Phenomenological Research, Wiley], pp. 269-311.

Id. 2004, *Functionalism, Computationalism, and Mental Contents*, "Canadian Journal of Philosophy", 34(3), [Taylor & Francis, Ltd., Canadian Journal of Philosophy], pp. 375-410.

Putnam H. 1960, *Minds and Machines*, [en ligne], "Dimensions of Minds", New York University Press, pp. 138-164, URL: https://philarchive.org/rec/PUTMAM.

Qiufan C., Lee K.F., Guez S., *et al.* 2022, *I.A 2042 - Dix scénarios pour notre futur*, Les Arènes, Paris.

Rantala J. 2020, *Anthropomorphism in Social Robotics: Simondon and the Human in Technology*, [en ligne], dans Nørskov M., Seibt J. et Quick S. O.(éds.), *Frontiers in Artificial Intelligence and Applications*, IOS Press, URL: http://ebooks.iospress.nl/doi/10.3233/FAIA200946.

Sartre J. P. 1947, *Huis clos: suivi de: les mouches*, Gallimard, Paris.

Searle J. R. 1980, *Minds, Brains and Programs*, "Behavioral and Brain Sciences", 3(3), pp. 417-457.

Simondon G. 2014, *Imagination et invention (1965-1966)*, PUF, Paris.

Simondon G. 2012, *Du mode d'existence des objets techniques*, rééd., Aubier, coll. «Philosophie»., Paris

Simondon G., Chateau J.Y. 2005, *L'invention dans les techniques: cours et conférences*, Seuil, coll. «Traces écrites», Paris.

Stengers I., Schlanger J. E. 1991, *Les concepts scientifiques: invention et pouvoir*, Gallimard, coll. «Folio Essais» 161, Paris.

Was (not was) 1981, *Tell Me That I'm Dreaming*, [en ligne], "Island Records", URL: https://www.youtube.com/watch?v=ZaawHz6p1QE.

Watson D. 2019, *The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence*, "Minds and Machines", 29(3), pp. 417-440.

Yang Y., Liu Y., Lv X., *et al.* 2022, *Anthropomorphism and customers' willingness to use artificial intelligence service agents*, "Journal of Hospitality Marketing & Management", 31(1), Routledge, pp. 1-23.

Zador A., Escola S., Richards B., *et al.* 2023, *Catalyzing next-generation Artificial Intelligence through NeuroAI*, "Nature Communications", 14(1), Nature Publishing Group, p. 1597.

«Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer» [en ligne], *Time*, le 18 janvier 2023, URL: https://time.com/6247678/openai-chatgpt-kenya-workers/, consulté le 28 mars 2023.

Why Does AI Lie, and What Can We Do About It?, [en ligne], 2022., URL:https://www.youtube.com/watch?v=w65p_IIp6JY.

«Worldcoin Whitepaper» [en ligne], *Worldcoin Whitepaper*, URL: https://whitepaper.worldcoin.org/, consulté le 25 juillet 2023.

«Deception, exploited workers, and cash handouts: How Worldcoin recruited its first half a million test users» [en ligne], *MIT Technology Review*, URL: https://www.technologyreview.com/2022/04/06/1048981/worldcoin-cryptocurrency-biometrics-web3/.

Jim Schrub
Université Paris Nanterre
jim.schrub@gmail.com