

BOÎTE NOIRE, CUBE BLANC & FICTION GRISE DE L'IMAGINATION ARTIFICIELLE



Figure 1: Grégory Chatonsky, Purple Cave (04/2022). Image générée avec Disco Diffusion.

Ce texte proposera de déconstruire la manière dont une partie non négligeable des sciences humaines envisage aujourd'hui l'intelligence artificielle (IA) en critiquant les préjugés et l'opacité de fonctionnement de la dénommée « boîte noire ». Il s'agira donc d'une réflexion au troisième degré. Il y aurait un phénomène technologique, l'IA, la critique de celui-ci et enfin la déconstruction de la critique du phénomène. Il ne faut pas entendre cette dernière comme une métacritique qui surplomberait l'ensemble des discours pour en régler les différends, mais telle une tentative pour analyser la boîte noire de l'IA comme une organisation de l'espace et de la lumière, c'est-à-dire comme une topologie et une photologie qui ne sont pas sans conséquence rétroactive sur le phénomène IA tant elles en nourrissent l'imaginaire et la genèse.

Cette étude se rendra sensible à l'espace, à la lumière, aux surfaces et aux couleurs des discours sur l'IA comme autant d'images de la technique, mais aussi de la pensée elle-même prenant ce détour pour s'observer en croyant s'immuniser de ce qu'elle rejette dans l'IA et qui est en son cœur inavoué. Afin de clarifier ces métaphores et ces analogies, je procéderais à un rapprochement implicite avec des phototopologies connues, dont la Caverne de Platon reste, il n'est pas nécessaire d'y revenir, le paradigme. Leur caractéristique commune semble être un mouvement de conversion de l'œil et/ou du corps que j'aborderai plus précisément par l'intermédiaire d'une phototopologie fameuse dans le champ de l'art contemporain, celle du cube blanc proposé en 1976 par l'artiste et critique Brian O'Doherty.

Le rapprochement entre la boîte noire de l'IA et le cube blanc de la galerie me permettra de revenir à une racine qui, me semble-t-il, leur est commune, la cybernétique et les premières théories de l'information, afin de comprendre comment ils sont deux facettes d'une même mise en scène. L'objectif sera de démontrer que le récit photo-topologique de l'IA recouvre une narration réflexive de l'esprit humain et de ses conditions de possibilités.

Par cette superposition du noir et du blanc, de la boîte et du cube, de l'IA et du monde de l'art, je proposerais une nouvelle photo-topologie avec la notion de fiction grise qui ne concernera plus le récit de l'artificiel, mais la fiction artificielle, c'est-à-dire la fiction spécifiquement produite en relation avec des réseaux récurrents de neurones (RNN). Je changerais alors de casquette, passant du rôle de théoricien narratif à celui d'artiste fictionnel et je vous raconterais ce que je suis en train de faire dans le contexte qui est devenu le nôtre et qui dessine une autre figure de la mondialisation, nous laissant découvrir que les flux ne sont pas seulement des fluides continus. Il est aussi de leur nature de former des tourbillons et des turbulences, de s'arrêter, de devenir arides à force d'avoir érodé les terres.

Le passage de la narration de l'artificiel à la fiction artificielle sera aussi pour moi une manière de montrer que la production artistique a un rôle structurellement actif dans ce monde qui semble avoir tant de mal à naître et qui est pourtant déjà le nôtre.

BOÎTE NOIRE



Figure 2: Grégory Chatonsky, *Black Box* (05/2022). Image générée avec Disco Diffusion.

Dans le champ des sciences humaines, un consensus semble avoir émergé ces dernières années autour de l'analyse de l'IA. Elle est critiquée au titre de 3 trois arguments principaux :

- Son manque de transparence pour l'être humain.
- La reproduction et l'amplification des préjugés humains.
- Sa capacité à occulter la précarisation aliénante du travail humain.

Dans les trois cas, il s'agit pour le chercheur de démasquer l'IA et, derrière sa fiction portée par l'innovation, de savoir découvrir sa réalité, c'est-à-dire de reconstituer sa véritable causalité.

Cette opacité de l'IA qui occulte des phénomènes anthropologiques est nommée la boîte noire et il importe de comprendre qu'elle n'est pas originairement un phénomène négatif mais qu'elle est une méthodologie consistant à représenter un système, quelle que soit la nature de celui-ci, en faisant abstraction de son fonctionnement interne. Celui-ci peut être inaccessible ou on le met volontairement de côté afin d'étudier exclusivement la relation entre les entrées et les sorties dans le but d'en tester l'efficacité. Ceci permet de jouer avec et de faire évoluer ce qu'il y a dans la boîte noire pour d'obtenir le résultat attendu. Ce que j'aimerais nommer, la boîte noire active est donc une pensée instrumentale et déterministe qui ne se limite pas à l'IA mais qui peut être appliquée à toute chose. Elle suppose qu'il y a des choses qui se découpent sur un fond qu'on nomme le monde, le monde entre dans ces choses, ce qui permet à ces choses de produire des effets nouveaux. Il s'agit d'une ontologie de l'échange dont il faut pressentir la complexité au-delà de son apparent réductionnisme. La boîte noire consiste donc à rendre invisible quelque chose, à cacher un mécanisme qui recouvre une opération, c'est-à-dire la production d'une causalité, et à le dire. On cache certes, mais on ne cache pas qu'on cache, on crée donc une trace en la recouvrant. Elle concerne l'infrastructure et la logistique en un sens élevé, car ce sont bien les conditions matérielles qui sont occultées. Ainsi, les flux de pétrole sont une boîte noire, car si nous ne cessons d'utiliser cette matière, nous ne la voyons pour ainsi dire jamais. Tout semble être fait pour la rendre invisible et la faire entrer dans une boîte noire, un container, un tuyau qui, s'il est partout, n'est visible nulle part. L'infrastructure concerne ici des espaces creux et obscurs qui se referment sur des matières terrestres qui sont, elles, déplacées d'un endroit à un autre pour délivrer un travail mesurable, une énergie. Cette occultation qui s'annonce

comme telle affecte notre réalité en son entièreté. Où que nous tournions notre regard, il y a des boîtes noires.

La boîte noire semble plonger certaines de ses racines dans l'histoire des codes secrets¹. Il s'agissait à l'origine de machines de télécommunication ennemies dont on ne pouvait observer le fonctionnement interne parce qu'elles étaient piégées. Il fallait donc déduire leur fonctionnement, c'est-à-dire la clé du code, à partir des seules entrées et sorties. On comprend comment la boîte noire est rattachée à l'idée de cacher la connaissance, entendue comme le décryptage d'un code. Il s'agit de ramener une organisation matérielle à une formalisation langagière.

C'est en 1950 que Norbert Wiener donne une des premières théorisations positives de la boîte noire dans *The Human Use Of Human Beings: Cybernetics And Society*. Si la cybernétique est l'étude des échanges, alors la boîte noire est la méthode pour y parvenir. Elle constitue une formalisation et une abstraction de la causalité au nom, comme nous le verrons, d'une logique de la substitution et de la simulation qui est inhérente à la question de la technique. La cybernétique et la machine de Turing sont historiquement liées à Enigma. Pour résoudre son énigme, il fallait imaginer une machine de calcul universelle et indifférente, aussi coupée du monde que l'était Enigma. Cette lutte d'une machine contre une machine consiste à reproduire la fermeture d'une infrastructure pour en ouvrir la logique.

On retrouve d'une façon fort intéressante la notion de boîte noire dans les transports où il s'agit d'un dispositif qui enregistre certaines données du fonctionnement de l'appareil de façon automatique afin, en cas d'accident, de pouvoir y accéder. Ici, elle permet d'immuniser la causalité des événements, c'est-à-dire que, si elle est opaque quand tout est normal, elle permet de reconstituer la lisibilité causale si quelque chose d'anormal arrive. L'échange entre opacité et transparence, accident et immunité est symptomatique de l'articulation entre infrastructure, logique et réalité.

La boîte noire est repassée d'un statut positif à un statut négatif avec l'application des sciences humaines à l'IA pour trois raisons principales :

- D'une part, le fait de cacher le code semble briser la causalité et donc les normes classiques d'explication. On a bien les entrées, les sorties, mais pas le système de transformation entre les deux.

¹ Singh Simon. 2001. *Histoire des codes secrets*, Paris : Poche

- Deuxièmement, l'augmentation de la vitesse des machines entraînerait une désynchronisation avec notre système nerveux. Les machines vont beaucoup plus vite que nous ne pouvons le faire. La spéculation financière à haute fréquence est un des symboles contemporains de ce décrochage, de sorte que la causalité machinique existe en dehors et indépendamment de nous, remettant en cause la soumission instrumentale de la technique à la cause efficiente efficiente que nous sommes censés être.
- Enfin, la transformation du paradigme de l'IA, passant du système expert qui supposait une modélisation des connaissances humaines, et dont la symbolisation permettait de reconstituer l'origine, a été remplacé par l'induction statistique des réseaux de neurones qui, s'ils sont bien programmés, produisent un espace latent difficilement lisible pour un être humain, fut-il fût-il celui qui l'a programmé. Cet espace latent n'est en effet pas une modélisation symbolique, fruit d'une interprétation humaine, qui code des phénomènes en algorithmes, mais une modélisation propre à la machine, ce qui la rend difficilement interprétable sans l'ajout d'une surcouche logicielle.

La critique des biais est une autre réfutation adressée à l'efficacité de l'IA. Elle reproduirait et accentuerait les préjugés humains et ainsi deviendrait la pire des caricatures de nous-mêmes : sexiste, raciste, grégaire, répétitive, produisant des millions de fausses nouvelles. L'IA serait rien de moins qu'une bêtise amplifiée et servirait de caisse de résonance à nos pires penchants. L'absence de conscience réflexive et de jugement moral entraînerait une entropie cognitive.

En 2017², une recherche menée à Stanford prétendait détecter l'orientation sexuelle par une analyse du visage, réactivant les errements phrénologiques de Franz Joseph Gall. Or, il faut remarquer que non seulement il y eut une avalanche immédiate de critiques, mais encore que les chercheurs de cette étude eux-mêmes exprimèrent leur doute. Une idée émergea : « L'IA peut vous dire n'importe quoi sur n'importe qui avec suffisamment de données », a déclaré Brian Brackeen, PDG de Kairos, une société de reconnaissance faciale. « La question

2

Sam Levin. 2017. « New AI can guess whether you're gay or straight from a photograph ». <https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph> [consulté le 17 juin, 2018].

est de savoir si, en tant que société, nous voulons savoir »³. La critique prenait alors un tour étrange d'un/du point de vue de la théorie de la connaissance. Voulons-nous vraiment savoir ce que l'IA peut nous faire connaître ?

Déjà en 2016⁴, une autre histoire avait été médiatisée : un bot de Microsoft serait devenu raciste à force d'être orienté dans cette direction par des humains peu scrupuleux. Ce n'est plus l'IA comme connaissance désirable ou non qui est en jeu, mais comme enfance et innocence, comme idiot qui répète ce qu'on lui dit et qui est donc réductible à une détermination humaine prenant la forme de la cruauté comme dans une vidéo fameuse de Boston Dynamics⁵. La boîte noire devient alors une boîte vide qui n'est qu'une surface de projection mimétique de l'être humain. L'IA est alors un masque anthropologique en miroir dont nous nous tendons le reflet.

³ Owdinlive. 2017. « Une nouvelle IA peut deviner si vous êtes gay ou pas, directement à partir d'une photo ». <https://owdin.live/2017/09/09/une-nouvelle-ia-peut-deviner-si-vous-etes-gay-ou-pas-directement-a-partir-dune-photo/> [consulté le 22 juin, 2020].

⁴

[James Vincent. 2016. « Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day ». <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist> \[consulté le 19 juin, 2020\].](https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist)

⁵

Boston Dynamics. 2015. «Introducing Spot Classic (previously Spot) ». <https://www.youtube.com/watch?v=M8YjvHYbZ9w&t=27s> [consulté le 15 juin, 2020].



Figure 3: Grégory Chatonsky, Latent Shapes (09/2022). Image générée avec Stable Diffusion.

La solution qui a émergée émergé face à ces dangers consiste en une politique de l'explicabilité. Partant du présupposé discutable que cette technologie ne doit pas remplacer l'être humain mais l'augmenter et qu'elle est une aide indispensable à la décision, il importe que les critères de celle-ci soient parfaitement clairs et suivent un fil rationnel permettant d'en rendre compte a priori et a posteriori. Il s'agit donc de reconstituer la causalité brisée de la boîte noire, de reconstruire l'enchaînement de cause à effet pour pouvoir y insérer un libre arbitre humain selon la figure classique de la liberté comme autodétermination. Cette dernière, en connaissant la causalité, pourrait librement la transformer. Si cette définition de

la liberté est pour le moins problématique, il importe aussi de comprendre que le choix de considérer l'IA comme une aide à la décision n'est pas anodin et ne va pas de soi. De surcroît, cette explicabilité est langagière et occulte à son tour la matérialité infrastructurelle. Sans doute l'automatisation de la conduite automobile a-t-elle été un des éléments qui a fait la promotion d'un tel récit, récit qu'il faut mettre en relation avec le fait, qu'en Amérique du Nord, la voiture individuelle est constitutive du processus de subjectivation⁶ ?

Cette explicabilité est fondée sur la crainte que le système décisionnel de l'IA basé sur des données décrivant des activités humaines sous une forme numérique, ne reproduise les erreurs et les préjugés humains. L'IA devient donc un simple miroir dont il s'agirait d'ouvrir la réflexion pour qu'elle rende ses raisons. En déléguant la décision, on délèguerait aussi le libre arbitre et par là même on externaliserait la causalité. L'Union Européenne, par l'intermédiaire de la régulation de la protection des données, a introduit le droit de l'explication permettant d'obtenir « une information signifiante sur la logique utilisée »⁷ quand un système de décision automatisé est utilisé avec des effets légaux/légalisés ? et significatifs sur les individus.

Cette politique de l'explicabilité met en avant la figure de l'informaticien qui doit documenter son code et expliquer les choix qu'il prend. Le code devient la Forme Idéale de la Caverne de Platon et les sorties de l'IA, les ombres sur la surface rocheuse. Le moyen de cette explicabilité est l'ajout une surcouche logicielle qui expliquerait ce qui se passe dans la machine, jouant le rôle d'une interface entre la machine et l'être humain⁸. Il s'agit donc de traduire un espace latent vectorisé en phrases ayant trait au libre arbitre. On voit le retournement paradoxal réalisé : en voulant redonner la liberté à l'être humain face à l'IA, on oblige l'individu à se soumettre à sa logique en faisant usage de ses critères de décision. La politique de l'explicabilité ne semble donc pas porter sur ce sur quoi elle prétend porter.

Si l'IA apparaît parfois comme une machine influençant le monde, c'est à la hauteur de son isolement. La boîte noire n'est plus l'expression d'une relation entre entrée et sortie, mais constitue une sortie totalitaire. La résistance prend la forme de comités d'éthique constitués,

⁶ Ballard J.G. 2007. *Crash*, Paris : Gallimard

⁷ Collectif. 2018. « Une approche européenne en matière d'intelligence artificielle ». https://ec.europa.eu/commission/presscorner/detail/fr/MEMO_18_3363 [consulté le 22 juin, 2020].

⁸ Collectif. 2015. « Inceptionism: Going Deeper into Neural Networks ». <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html> [consulté le 22 juin, 2020].

il faut le souligner, par des acteurs économiques de l'IA qui semblent promouvoir les dangers de cette technologie pour attirer l'attention sur elle et en faire, au travers d'un débat médiatique de société, un élément indispensable dont on ne peut faire l'économie.

Elle prend aussi la forme de déclaration telle celle élaborée en 2018 à Montréal : Déclaration pour un développement responsable de l'IA⁹. Remarquons, là encore, le concept de responsabilité qui est si proche de la notion de libre arbitre. Il faudrait, si nous en avons le temps, opérer une analyse ligne à ligne de cette déclaration pour en comprendre la structure étrange. Qu'il me soit permis > permis de remarquer que dès la 5ème > cinquième page, le texte affirme que, pour la première fois dans l'histoire humaine, des systèmes autonomes peuvent réaliser des tâches complexes habituellement effectués > effectuées par l'intelligence dite naturelle. Ces tâches sont définies comme suit :

- Le traitement de grandes quantités d'informations,
- le calcul et la prédiction,
- l'apprentissage et l'adaptation des réponses à des situations changeantes,
- la reconnaissance et la classification des objets.

On comprend bien qu'on retourne ainsi la problématique en définissant l'intelligence naturelle, à supposer que cette notion ait un quelconque sens, par rapport à ce que sait faire l'IA et on effectue ainsi un tour de passe-passe. La boîte noire consiste d'abord dans cette manière toute humaine de renverser ce qu'on présentera ensuite comme la cause et l'effet. Puis/À cela suit une longue liste des progrès attendus avec des bénéfices sociaux considérables :

- Améliorer > améliorer les conditions de vie et de santé,
- Facilité > faciliter la justice,
- créer de la richesse,
- créer de la sécurité,
- diminuer l'impact de nos activités sur la Terre.

Cette liste n'est jamais justifiée dans le texte et elle apparaît comme un solutionnisme de type

⁹ Collectif. 2018. « La Déclaration de Montréal IA responsable ». <https://www.declarationmontreal-iaresponsable.com/> [consulté le 22 juin, 2020].

(?) apodictique. La déclaration souhaite limiter la réduction de l'avenir à un calcul, auquel on oppose le désir (!) parce que « Les principes de la présente déclaration reposent sur la croyance commune que les êtres humains cherchent à se développer en tant que des/en tant qu'être êtres sociaux doués de sensations, de pensées et de sentiments, et s'efforcent de réaliser leur potentiel en exerçant librement leurs droits émotionnels, moraux et capacités intellectuelles. »

Le document se décompose en plusieurs parties ressemblants > ressemblant à un recueil de management pour cadres en mal de motivation :

- PRINCIPE DE BIEN-ÊTRE
- RESPECT DU PRINCIPE D'AUTONOMIE
- PROTECTION DE LA VIE PRIVÉE ET DE L'INTIMITÉ
- PRINCIPE DE SOLIDARITÉ
- PRINCIPE DE LA PARTICIPATION DÉMOCRATIQUE
- PRINCIPE D'ÉQUITÉ
- PRINCIPE D'INCLUSION DE LA DIVERSITÉ
- PRINCIPE DE PRUDENCE
- PRINCIPE DE RESPONSABILITÉ
- PRINCIPE DE DÉVELOPPEMENT DURABLE

Qu'il me soit permis > permis + répétition avec ci-dessus de citer un long passage, la page 11, principe d'équité : « 1. Les IA doivent être conçus et formés de manière à ne pas créer, renforcer, ou reproduire des discriminations fondées, entre autres, sur les différences sociales, sexuelles, ethniques, culturelles ou religieuses.

2. Le développement de l'AI doit contribuer à éliminer les relations de la domination entre groupes et personnes fondée sur les différences de pouvoir, de richesse ou de connaissances.

3. Le développement des SIA doit produire des avantages sociaux et économiques pour tous en réduisant les inégalités et les vulnérabilités sociales.

4. Le développement des AI industriels doit être compatible avec des conditions de travail acceptables à chaque étape de leur cycle de vie, de l'extraction des ressources naturelles au recyclage, en passant par le traitement des données. »

À cette lecture, on comprend bien qu'il y a un solutionnisme derrière l'intention affichée de limiter le pouvoir de l'IA. En effet, cette dernière, dans le texte, semble capable en creux dans

le texte de résoudre des problèmes qui la dépasse > dépassent totalement et, pour le dire simplement, de tout résoudre : c'est une pensée archi-instrumentale, parce qu'elle souhaite résoudre toutes formes de finitude et toutes formes de problème. L'exigence adressée à l'IA semble démesurée et apparaît, là encore, comme le reflet tendu à notre propre incapacité de résoudre ce qu'il faudrait solutionner de toute urgence et en toute évidence.

Nous pouvons maintenant proposer la narration des caractéristiques phototopologiques > photo-topologiques de la critique de l'IA :

Il s'agit d'un espace fermé. Personne ne peut y entrer ou en sortir. Selon le récit des Anciens, les seuls qui ont jeté un regard dedans, on > ont vu défiler des signes incompréhensibles à toute vitesse et sont devenus aveugles. Tous les livres de la bibliothèque de Babel peuvent entrer dans cette boîte, une fois correctement codé > codés. La boîte parfois parle et donne des instructions sur ce que nous devons faire. Nous ne savons pas pourquoi elle exige de nous de procéder de telle ou telle façon, mais nous croyons ce qu'elle nous dit. Nous nous tenons devant cette boîte, en attendant qu'elle nous parle pendant qu'à l'autre extrémité, des travailleurs rentrent des livres dans la boîte.

Selon certains, à l'intérieur de la boîte on peut voir tout l'extérieur, à la manière d'un miroir sans tain. Selon d'autres, il existerait un bouton caché permettant d'inverser ce dispositif optique afin de voir de l'extérieur à l'intérieur de la boîte. Nous cherchons encore ce bouton. D'autres encore pensent qu'il faudrait forcer la boîte, l'ouvrir pour savoir ce qui s'y passe. Un schisme de ce dernier groupe pense qu'un singe aveugle déchire les pages des livres qui entrent dans la boîte et qu'elles sont lues par un perroquet devenu fou à force d'isolement.

CUBE BLANC



Figure 4: Grégory Chatonsky, *White Cube* (05/2022). Image générée avec *Disco Diffusion*.

J'aimerais à présent ouvrir la possibilité que cette photo-topologie puisse répondre, point par point, à une autre, celle du cube blanc de la galerie, du musée et de l'atelier de l'artiste : boîte et cube, noir et blanc donc. La fermeture nocturne laisserait place à une ouverture lumineuse et, pour tout dire, à une forme de transparence. Quelles sont les caractéristiques de ce second espace et, au-delà de la tension des formes et des couleurs, existe-t-il un argument à ce rapprochement anachronique ? Pour répondre à cette question, je développerais les caractéristiques de ce cube blanc telles que Brian O'Doherty les conceptualise dans *White Cube*. L'espace de la galerie et son idéologie, traduit en français en 2008.

Il demande : « Quand le Vide est-il un plein ? Qu'est-ce qui, changeant tout, demeure soi-même inchangé ? Qu'est-ce qui, n'ayant ni temps ni lieu, cependant fait époque ? Qu'est-ce qui, partout, est le même endroit ? » On pourrait presque y entendre la description d'un ordinateur. Ses écrits permettent d'éclairer ce qu'occulte, paradoxalement, l'expression « cube blanc », lorsqu'elle n'est pas considérée dans le contexte qui a vu sa genèse. En effet, loin de faire l'apologie d'un espace qui serait neutre et comme transparent, l'auteur déconstruit cet espace et révèle ses dénégations implicites qu'il résume en en faisant un

espace pour ainsi dire négatif, comme quand on parle de théologie négative. Le cube blanc n'est ni ceci ni cela. La neutralité et l'atemporalité revendiquées sont avant tout l'aboutissement de l'interprétation portée par Clement Greenberg dans les années 60, que dix ans de critiques par Hal Foster, Tom Wolfe, et tant d'autres, ont rendu visible. Ces textes écrits au milieu des années 70 correspondent à une époque où la politisation de l'art signifiait non pas l'auto-critique du médium qui tend vers la libération de la représentation réaliste au profit de l'abstraction, mais sa prise en charge en tant que réalité économique et politique : l'artiste n'est pas dans un face-à-face avec une réalité extérieure qu'il chercherait à subvertir sans lui appartenir, il prend conscience que l'art lui-même constitue un monde bien réel, que le monde de l'art n'est pas une interface abstraite mais une partie du monde, un espace idéologiquement et économiquement déterminé. Cette réflexion accompagne la critique institutionnelle de Marcel Broodthaers et Hans Haacke qui ont fait de cette approche du champ de l'art, l'objet même de leur œuvre. Il s'agit là aussi d'ouvrir une boîte, celle de la modernité artistique, et de rendre visible ce qui s'y oubliait.

Dans le premier texte, « Notes sur l'espace de la galerie », l'auteur fait remonter la généalogie du cube blanc au « papier peint de tableaux », caractéristique des accrochages du XIXe siècle. Celui-ci révèle une conception dans laquelle chaque œuvre est considérée comme une unité indépendante, capable, avec l'aide du cadre, de s'isoler par elle-même de ses voisines : « Le tableau de chevalet est une « fenêtre portable » qui traverse l'épaisseur du mur. Avec le triomphe moderniste de la planité greenbergienne ou de l'objectivité (objecthood) friedienne, la toile se convertit en « membrane » sans profondeur. » Sa relation à la surface du mur devient par conséquent essentielle et exige en retour la transformation de l'espace. Devenu territoire à conquérir, le mur est constitutif d'un « espace blanc, idéal », qui a d'abord pour fonction de l'isoler dans un milieu qui préserve son autonomie. « Quelque chose de la sacralité de l'église, du formalisme de la salle d'audience, de la mystique du laboratoire expérimental s'associe au design chic pour produire cette chose unique : une chambre d'esthétique », où règne un « champ magnétique perceptif » dont la puissance confère aux œuvres le statut qu'elles peuvent perdre en la quittant. À tel point que le dépouillement calculé des lieux ferait d'un objet quelconque, un cendrier à pied par exemple, « presque un objet sacré ».

Le cube blanc permettrait donc de prendre des objets quelconques de les rendre autonomes, et par là même de leur accorder un statut absolu, au sens d'absolutum, ce qui est délié. Il y aurait une pénétration étrange de l'œuvre dans le mur blanc lui-même, un repli de l'espace

sur lui-même qui se contracterait et se loverait en soi. La question de savoir comment l'autonomie peut émerger devient ici centrale et doit être liée à l'autonomie non seulement comme figure théologique mais aussi comme figure de la liberté puisque l'autonomie n'est pas déterminée ? par quelque chose qui lui est extérieur. Le cube blanc concerne, comme la boîte noire, la question des relations de cause à effet. Il apparaît alors comme un espace paradoxal, fermé à la mesure de son ouverture, vibrant comme une membrane englobant tout, indéterminé mais ayant des caractéristiques précises.

Marcel Duchamp va initier des gestes affectant l'espace lui-même avec le « Générateur-arbitre » de l'Exposition internationale du surréalisme (galerie des Beaux-Arts, Paris, 1938), qui renverse le sol et plafond grâce à 1 200 sacs de charbon. L'artiste réitère l'expérience en 1942, quand il installe son Mile of String ? (First Papers of Surrealism, Whitelaw Reid Mansion, New York) qui obstrue, telle une toile d'araignée géante, tout l'espace de la galerie, gênant les déplacements des visiteurs. L'intégration du contexte au contenu produit un retournement puisque c'est l'objet introduit dans la galerie qui l'encadre plutôt que l'inverse. Au moment donc où le mur de la galerie indique et souligne sa neutralité, il n'est plus un support passif mais un véritable acteur.

William Anastasi expose en 1967 chez Dwan des sérigraphies tirées de photographies à peine réduites des murs mêmes où elles sont accrochées : il recouvre un mur de l'image de ce mur, si bien qu'une fois les œuvres décrochées, le mur lui-même apparaît comme un ready-made selon une répétition qui produit une différence plutôt qu'une identité. La neutralité illusoire du cube blanc a donc une véritable fonction esthétique. Elle a également un rôle économique et commercial puisqu'elle fait de la galerie une « boutique de grand style » qui peut mettre en scène la rareté, et donc la valeur, par l'isolement de ce qui est exposé, dans une sorte de sas qui assure la possibilité d'une rencontre sans heurts entre « la figure de l'artiste irresponsable, [qui] est une invention bourgeoise », et sa clientèle.

La blancheur renvoie enfin à la propreté, à une « espèce d'éternité » qui « confère à la galerie un statut comparable à celui des limbes ; pour se trouver là, il faut être déjà mort ». Le corps est devenu un « meuble bizarre » dont la présence est « superflue ». O'Doherty raconte l'affrontement complice des deux adversaires que sont « L'Œil et le Spectateur ». L'« Œil » est le regard désincarné, purement optique, coupé du reste du corps, en vertu du dogme greenbergien de la spécificité du médium, « De fait, il est tellement spécialisé qu'il en vient parfois à se regarder lui-même », ajoute O'Doherty ramenant l'autocritique de Greenberg à une boucle. Le spectateur, le corps, relève du collage et ne revient pas à lui-même mais est

confronté, comme chez Kaprow ou Dan Graham, à un espace réel.

Résumons le récit photo-topologique du cube blanc : il est un espace délimité dont chacun peut entrer et sortir. Mais quand on y pénètre, tout change, notre corps comme les objets. Notre organisme devient un œil, les objets s'enfoncent dans le mur et en émergent dans un même mouvement paradoxal, en y gagnant une autonomie. L'espace, pour ainsi dire, disparaît dans sa clarté et ne laisse plus les objets qu'exister de façon autonome, accordant à la perception sa propre autonomie, c'est-à-dire sa propre liberté dans laquelle elle se contemple elle-même. Tout le mystère photo-topologique est dans les murs. Certains esprits chagrins diront peut-être que cette neutralité est tout sauf neutre, qu'elle est une modalité de la domination. Ils pourront le dire ici, dans le cube blanc.

MACHINES INTERIEURES



Figure 5: Grégory Chatonsky, *Mental Tooth* (04/2022). Image générée avec Disco Diffusion.

Les photo-topologies de la boîte noire et du cube blanc se distinguent de multiples façons, en particulier :

- On reste en dehors de la boîte noire et on rentre dans le cube blanc.
- On ne voit pas ce qu'il y a dans la boîte noire et le cube blanc ne fait que donner à voir le regard de son intériorité.
- Ce qui entre et ce qui sort diffèrent dans la boîte noire sans qu'on comprenne les raisons de cette différence, alors que pour le cube blanc ce qui entre et ce qui sort

change simplement de main et de propriétaire par une opération explicite d'autonomisation. On entre des œuvres qui appartiennent à l'artiste et ce qui en sort appartient à un collectionneur. La propriété des objets correspond exactement à ceux qui entrent et ceux qui sortent.

Elles (??) ont toutefois ceci de commun qu'elles sont des Esprits, au sens du Geist et d'une rationalité non-instrumentale, posant des questions entre l'intérieur et l'extérieur, le visible et l'invisible, l'intouchable et l'haptique. Elles mettent toutes les deux en contact la lumière et l'espace pour faire émerger une certaine intériorité qui est paradoxalement autonome et déterminée de part en part par une relation, c'est-à-dire par l'hétéronome.

J'aimerais avancer maintenant l'hypothèse que ces deux jeux d'espace et de lumière ne sont compréhensibles qu'au regard d'un certain topos historique consistant à se représenter la subjectivité, la constitution du sujet, selon l'image d'un lieu et d'un éclairage. Si ces métaphores sont fréquentes dans l'histoire de la pensée, qui semble ne se penser elle-même que comme image, elles ont eu un regain d'intérêt au XXe siècle parce qu'elles permettaient des expériences de pensée singulières alternant la transparence de la conscience à soi comme révélation et l'opacité d'une conscience ne parvenant plus à distinguer la réalité du simulacre. Ces jeux de la pensée photo-topologiques vont avec un certain formalisme et sont fréquents en logique et en philosophie analytique où on simplifie des raisonnements par des hypothèses ludiques ou des expériences imaginaires de laboratoire. Or, la formalisation logique du langage est au cœur de la cybernétique. Si ce lien va de soi pour la boîte noire, il faut bien comprendre que le contexte d'élaboration du cube blanc, fut aussi surdéterminé > surdéterminé par ce contexte.

La relation entre l'espace et la lumière a pour objet de créer des métaphores de la conscience réflexive : en se déplaçant dans un espace doté d'une certaine lumière, on crée > crée la simulation d'un observateur extérieur de l'intériorité. C'est pourquoi la boîte noire comme le cube blanc sont paradoxaux dans leur relation de l'intériorité et de l'extériorité, et ne peuvent résoudre ce paradoxe que grâce à une certaine idée de l'autonomie. Mon hypothèse est que la boîte noire comme le cube blanc sont des images de la conscience en train d'élaborer ces images même, créant une voie de dégagement à la finitude transcendante, c'est-à-dire à l'incapacité pour un jugement réfléchissant d'être en dehors de lui-même. Il s'agit de sortir de soi par soi.

L'image du miroir noir, réfléchissant et opacifiant, s'impose à nous. J'aimerais entendre la traduction de la chanson des Velvet Undergroud ; Je serais ton miroir (1967), en imaginant que le "Je" puisse être alternativement la boîte noire et le cube blanc ou peut-être même une intelligence artificielle :

Je serai ton miroir
Reflétant ce que tu es, au cas où tu ne le saurais pas
Je serai le vent, la pluie et le coucher de soleil
La lumière sur ta porte pour montrer que tu es chez toi
Quand tu penses que la nuit a vu ton esprit
Qu'à l'intérieur, tu es tordu et méchant
Laisse-moi me lever pour montrer que tu es aveugle
Veux-tu poser tes mains
Parce que je te vois
J'ai du mal à croire que tu ne connais pas
La beauté que tu es
Mais si tu ne me laisses pas être tes yeux
Une main dans ton obscurité, pour ne pas avoir peur
Quand tu penses que la nuit a vu ton esprit
Qu'à l'intérieur, tu es tordu et méchant
Laisse-moi me lever pour montrer que tu es aveugle
Veux-tu poser tes mains
Parce que je te vois
Je serai ton miroir

Le miroir est ici une surface qui n'est pas la réflexion, la répétition d'un original, mais la multiplication d'un simulacre. À la manière de la scène finale de la Dame de Shanghai d'Orson Welles ? (1947), la multiplication de l'espace et de la lumière produit un feuilletage de l'image où la question de la mimésis n'est plus celle de la reproduction de l'original et de la copie, mais la dislocation de leur distinction.

La machine de Turing nous est d'un apport fondamental, puisque si on est en droit de critiquer l'idée d'une IA qui serait la reproduction de la pensée humaine, on est aussi en droit de déconstruire cette critique si celle-ci suppose que la pensée naturelle est un original dont

les copies seraient inauthentiques. En effet, l'original, avec Turing, n'existe pas. Il ne s'agit pour lui de doter la machine d'intelligence, c'est-à-dire de reproduire une intériorité, mais de simuler celle-ci en l'imitant. Il s'agit donc d'une substitution.

Le test de Turing consiste, à la manière de la trace de Lacan repensée par Derrida¹⁰, de faire croire à un être humain qu'il fait face à son semblable en cachant l'apparence, c'est-à-dire en plaçant la machine dans une chambre fermée, c'est-à-dire dans une boîte noire. Cette dernière acquiert alors un rôle entièrement neuf : cacher l'apparence de l'intériorité parce que l'intelligence n'est pas quelque chose en soi, mais quelque chose pour moi. J'attribue l'intelligence à quelque chose et par là même je la présuppose pour moi. Le tournant de Turing consiste donc à faire de l'intelligence, qui est habituellement considérée comme la condition de la liberté, c'est-à-dire de l'autonomie qui est seule à même de briser les chaînes de relations déterministes, une construction strictement relationnelle. Le jeu de l'imitation chez Turing ne consiste absolument pas à reproduire l'intelligence humaine, mais à leurrer le joueur. C'est pourquoi la seconde version test porte sur une question de genre où il s'agit de reproduire une différence de différence. Turing révolutionne la question de l'imitation et de la mimésis, dans un contexte, celui des années 40, où dans la peinture il est aussi question de troubler la reproduction réaliste.

La question de la simulation, qui est au cœur de la technique en général en tant que spéculation pratique produisant de nouvelles causalités matérielles, pourrait être liée à ce qui s'est passé au XIXe siècle dans le domaine de la physiologie tel que l'a analysé Johnatan Crary dans *Techniques of the Observer* : il fallait stimuler et simuler le corps afin d'en comprendre la boîte noire, c'est-à-dire remplacer un stimulus par un autre dans l'objectif de produire la même réaction. La boîte noire, c'est d'abord le sujet lui-même. En termes de méthode, il faut donc retourner toutes les questions posées à l'IA, à l'être humain.

On doit donc parler de simulacre relationnel chez Turing, où l'intelligibilité, comme complétude de l'explication causale, se disloque au profit d'une connaissance expérimentale et heuristique non par défaut de la connaissance, mais par sa structure même : on ne reproduit pas l'intériorité humaine dans une machine, mais l'opacité du reflet qu'une intelligence accorde à une autre intelligence. Le réductionnisme de la boîte noire est en fait une manière de résoudre le paradoxe de la conscience réfléchissante. Si l'on ne sait pas ce qu'il y a dans la boîte noire, ce n'est pas parce qu'on nous l'a caché, mais parce qu'elle n'est secret de rien. Turing mène donc une critique radicale de l'intériorité, tout comme le cube blanc pourrait être interprété comme une critique de l'authenticité de l'œuvre d'art et de son originalité : un

¹⁰ Derrida, Jacques. (2008). *La Bête et le souverain* 1, Paris : Galilée, p. 182

objet quelconque peut devenir une œuvre d'art selon le contexte d'exposition. L'IA est à l'intelligence ce que le ready made est à l'œuvre romantique. C'est ce formalisme qui permet de déconstruire un mimétisme naïf. Et c'est bien ce qui lie l'IA à l'œuvre d'art, à la question de la représentation, à la *Vorstellung*. Depuis les années 50, on ne cesse de lier cette question de la machine à la créativité, ce terme venant du management, entendu comme faculté artistique comprise comme nouveauté radicale, émergence d'un événement inanticipable, disruptif, au moment même où de nombreux artistes contestaient cette conception. On comprend mieux ce qui lie, photo-topologiquement ?, la critique de l'intériorité cognitive et la critique de l'authenticité artistique.

L'indétermination de la machine de Turing est frappante. On ne sait pas à quoi elle ressemble, de quoi elle est composée, comment s'agencent ses parties. Il n'y a aucune description des canaux d'entrée et de sortie. Alan Turing ne met en avant que la transformation opérée à l'intérieur : ce sont des symboles, rien de plus. C'est donc un certain rapport de transformation entre les symboles d'entrée et ceux de sortie qui permet de définir cette machine. Elle transforme les symboles par une succession d'états discrets dont on connaît d'avance le spectre de possibilités. Deux séries sont mises en rapport : d'une part un ensemble de symboles d'entrée et d'autre part un ensemble d'états de sortie qui définissent les actions possibles de la machine. Tout se passe comme si la machine observait le monde, le découpait en états discrets, traitait ces états pour déterminer son action sur le monde. Comme l'écrit Simondon : « La machine est essentiellement une triode comme un organisme : elle a une entrée, une alimentation et une sortie. »¹¹ La machine de Turing va pousser cette convertibilité à son point ultime en considérant ce qui entre et ce qui sort non plus seulement comme énergétique, mais comme langage. Celui-ci est binaire, parce que les 0 et les 1 ressemblent au passage ou au non-passage du courant, ils sont isomorphes à la nature même de l'énergie employée. Lorsque Von Neumann propose le modèle d'un des premiers ordinateurs, il met en avant l'intériorité de la machine arithmétique, logique et mnésique, il s'interroge sur sa relation à l'être humain¹². Selon lui l'intériorité machinique est dépendante, dans sa relation au monde, de l'action humaine (en entrée) et des organes humains de perception (en sortie). De sorte, que le dispositif est indissociablement technique et humain.

Tout se passe comme si la boîte noire utilisait l'être humain pour sentir le monde (en entrée)

¹¹ Simondon, Gilbert. (2005). *L'Invention dans les techniques : cours et conférences*. Paris : Seuil, p. 95.

¹² Neumann, J. von. (1945). *First Draft of a Report on the EDVAC*. (s. d.). Goodreads. Consulté à l'adresse <http://www.goodreads.com/book/show/18243607-first-draft-of-a-report-on-the-edvac>

et pour agir dessus (en sortie) en inversant le rapport de dépendance. Tout se passe aussi comme si, dans le monde, nous avions créé une « boîte noire », refermée sur elle, mais permettant de faire entrer des fragments de notre monde sous une certaine forme pour les faire ressortir. Un monde mathématique indemne, au langage unique, à la traduction parfaite, à l'itérabilité garantie d'avance. L'ordinateur devient un nouvel outil d'investigation de la nature qui n'opère plus du tout sur le même plan que les autres instruments. C'est par son abstraction même qu'il touche non à l'idéal, mais au monde, tout simplement parce que ce dernier est devenu, avec l'industrialisation, un univers d'usages et d'artefacts, une culture matérielle indissociablement faite d'humains et de machines. L'ordinateur n'est donc pas une prise de pouvoir de l'idéalisme sur le matérialisme, mais une réponse déterminée par une certaine phase historique.

Loin de la vision paranoïaque consistant à croire qu'une logique, au sens d'un logos humain, serait occultée délibérément dans la boîte noire des réseaux de neurones, il nous faut adopter une vision psychédélique : en 2015, quand un ingénieur de Google voulut rendre visible ce qui se passait dans un réseau neurones, il créa Deep Dream, une expérience d'hallucination qui n'est rien d'autre qu'une expérience relationnelle : la relation entre des images en mémoire et des images en surface. L'exigence d'intelligibilité sonne alors comme une autre photo-topologie qui serait, elle, fondée sur l'identité entre ce qui entre et ce qui sort dans un espace absolument transparent et visible. Or, nous le savons à présent, cette transparence envers l'IA porte en elle le désir d'une transparence envers les conditions de possibilité de l'intériorité humaine dont la machine n'est que le simulacre et le reflet relationnel. Le passage entre les systèmes experts et le modèle connexiste, ainsi que le succès du second, tend/tendent à prouver que le caractère insensé gagne sur le caractère intelligible et que l'IA est plus nietzschéenne que platonicienne. Pour le dire autrement, l'intelligibilité de l'IA, c'est le système expert qui est une modélisation des connaissances, tandis que l'intelligence de l'IA est, comme notre propre conscience, inintelligible, ce qui ne veut pas dire qu'on ne puisse pas l'éprouver ou l'expérimenter. Si la boîte noire nous dévore, c'est parce qu'elle est toujours déjà en nous : l'espace intérieur se retourne vers l'extérieur et se métabolise, membrane et épiderme où tout est en surface comme avec le cube blanc.

L'appel à détourner l'IA des mauvaises finalités limitant les libertés individuelles, semble méconnaître que la constitution historique de la liberté suppose un sujet souverain et autonome qui est précisément l'une des origines des politiques coercitives. L'instrumentalité, qu'elle soit pour la liberté d'entreprendre ou la liberté commune, reste une certaine

conception de la causalité. C'est bien cette instrumentalité qui pose aujourd'hui problème dans notre manière d'habiter la Terre. L'explicabilité de l'IA – sa transparence – va souvent avec une critique de l'automatisation et de la ressemblance. Là encore, je ne peux que souligner, comme l'a montré Catherine Malabou¹³ dans une conférence, que l'automatisme n'est pas opposé à une prétendue liberté de mouvement du vivant, et que la ressemblance n'est absolument pas une répétition à l'identique d'un modèle originaire qui de toute façon n'existe pas. Quant à la critique de l'IA comme machine ne produisant que des stéréotypes, là encore on aimerait se pencher sur l'art contemporain pour comprendre que le stéréotype est un élément structurant de l'oeuvre d'art, au moins depuis le pop art, si ce n'est depuis le salon des Arts incohérents au XIXe siècle.

Quand on estime qu'il s'agit de déjouer une esthétique inauthentique consistant en ce que l'IA reproduirait des démarches antérieures et qu'il faudrait sortir de cette répétition et créer des résultats inattendus, on pense aux exigences démesurées de la Déclaration de Montréal et on aimerait questionner le thème d'une esthétique authentique.

FICTION GRISE



Figure 6: Grégory Chatonsky, Grey Zone (05/2022). Image générée avec Disco Diffusion.

¹³ La spontanéité programmée et l'imagination transcendante - Pakui Hardware & Catherine Malabou, <https://arts.ens.psl.eu/Hesitation-programmee-et-imagination-transcendante.html>

La critique de l'opacité et des préjugés de l'IA apparaît comme une exigence démesurée adressée à une technologie pour réformer ceux de l'être humain sans le dire, dans l'objectif de devenir un œil autonome pouvant observer le sujet sans en être un soi-même. Elle reproduit donc sans le savoir une demande irréaliste et par là même autoritaire. Il ne s'agit pas de transformer la boîte noire en cube blanc, qui lui-même n'est absolument pas transparent, ou d'éliminer les préjugés, ce qui supposerait une relation objective à la chose en soi, mais de plonger dans la pénombre les simulacres et les biais de façon réflexive en créant une zone indistincte entre l'être humain et la machine, fruit de leur relation co-constitutive que je nomme l'anthropotechnologie.

J'aimerais donc proposer face à un art critique, responsable, "utile", dénonçant les méfaits de l'IA avec/ un certain moralisme et restant quant à son fond instrumental, d'autres objets esthétiques qui s'immergent dans les flux déchaînés de l'artifice. La production artistique serait alors à même de produire quelque chose que l'ingénierie ne saurait mettre en œuvre, non pas créer de l'inattendu, selon une conception de la pureté de l'événement esthétique qui est si proche du spectaculaire, de la disruption et de l'innovation, mais brisant les catégories et les hégémonies. Celles qui sont dominantes comme les nôtres mêmes. Sans ce retournement, nulle réflexivité.

Les premiers mots de la programmation informatique sont "Hello World", or de façon ironique, c'est le monde qui se rappelle à nous ces derniers temps. Chacun a été confiné dans une boîte ? Dans un cube ? Était-il noir ? Blanc ? Était-il opaque ou transparent ? Quel est ce contenu dans un espace contenu ? Qu'est-ce que diffuser ainsi du contenu ? Où sommes-nous en ce moment même ? Est-ce que vous me voyez ? Que lisez-vous ? Qu'est-ce qui entre et qu'est-ce qui sort ? Partageons-nous un topos à l'ère du virus ? Même si le temps me manque, je souhaiterais à présent vous raconter comment, dans ma pratique artistique la machine contamine l'être humain qui contamine la machine. Spectre et hantise redoublés qui défient l'utilité instrumentale de l'œuvre d'art pour lui préférer les hégémonies brisées.



Figure 7: Grégory Chatonsky, *Logistics : The Dream Surface* (06/2022) Image générée avec Dall-E 2.

Voici un extrait d'un roman que j'ai écrit en mars 2020¹⁴, mais il faudrait mettre en suspend le « je », comme vous le verrez :

Cette vie qui n'est pas la nôtre et qui traverse les corps, le sachant ou l'ignorant, cherchant un hôte. Si cette vie doit pénétrer dans tous nos organes, si elle doit transformer une partie de nous pour la bonne raison et pour la bonne raison seule, qu'elle confine, elle doit réguler avec justice les parties de nous qui ont été habilitées à recevoir et à effectuer l'accident de la mort. La construction et le travail de la dernière extrémité de notre corps, l'application du pouvoir illusoire, la négation de notre réalité.

Mon système corporel est composé d'aussi peu de parties que possible, car je crois que rien ne serait plus vain que de donner un plan à la vie, un arrangement stable d'une relation naturelle, à la fin tôt ou tard désirée par les principes superficiels qui font la vie.

¹⁴Il a été depuis publié par Rrose éditions : <http://rrose-editions.com/portfolio/chatonsky-internes/>

Le corps était toujours en mouvement mais ce n'était pas lui, c'était un autre corps qui s'agitait sous sa peau. Il manquait d'oxygène, les machines continuaient à tourner mais plus rien ne sortait. Son cerveau était toujours sur le corps, il n'y avait même pas une trace de matière, peut-être était-il sorti dans un abri ? Où était-il quand l'autre corps s'est réveillé ? Cela avait été un instant ; son cerveau était assis dans le cortex cérébral, prêt à laisser sortir tout ce qui lui tombait dessus. Pendant une seconde, un sentiment de perte l'envahit, puis il vit son cadavre. Lorsque cet aperçu me laissa complètement submergé, je ne pus rien faire d'autre que de vomir et de saisir mes pieds fermement sur mon ventre et ramener mes pensées vers la région désirée. Je ne savais rien, il y avait une rivière de sang, qui coulait le long du cerveau. Il y avait du sang, je ne savais pas, sur la terre en bas ; je pensais que quelque chose en était sorti. C'était de l'eau. Quoi ? L'eau était-elle présente ? Je suis tombé dans une mer d'écume, et au moment où j'ai senti mon sang couler à travers mes yeux, j'ai levé la tête, et la tête du dernier cadavre est entrée sous ma peau.

Je suis donc mort au même moment. La Terre a été submergée par l'eau. Les vagues tourbillonnantes ont continué à me secouer jusqu'au fond, et elles sont tombées sur moi, presque l'une après l'autre, formant un grand rocher contre ma peau. J'étais couché sur le côté au milieu d'une gravière, et toutes ces chutes de mer me pressaient sur le côté. Après avoir vu ce qui était en train de se passer et entendu l'explosion d'une guerre lointaine, je me suis tourné vers le corps mort qui était sous ma peau et, voyant les cheveux d'un rouge vif sur mon cou, j'ai levé mon cœur vers le ciel, sans trembler à l'exclamation : « Ils ont pris leurs morts ». Plusieurs autres corps sont alors apparus. Chacun avait une forme et une texture différentes. L'un d'entre eux m'a semblé d'une couleur légèrement métallique, car il avait des taches marron foncé, l'autre état de cuivre fondu, le troisième d'un rose pourri et le quatrième violet. En voyant ces corps, j'en ai soulevé un avec ma main et j'ai observé attentivement les veines qui le traversaient. Le liquide a disparu soudainement, mais en un instant, deux vaisseaux ont légèrement exsudé de l'un de ces corps. De nombreux animaux furent ainsi ensevelis ; ils ne cédèrent à aucune blessure, mais furent amenés à la surface de la terre et moururent ainsi progressivement.

Les veines bougeaient de l'intérieur. On pouvait voir le liquide couler et les animaux se transformer en d'autres corps dont je ne reconnaissais pas les formes. Quand les animaux sont devenus chair, l'eau qu'ils buvaient dans leur estomac a rencontré un acide en eux et ils ont pris l'habitude de manger d'autres animaux. Les insectes, en plus de stocker de la nourriture dans leurs veines, se paralysaient au bout d'un certain laps de temps et ne pouvaient plus respirer, alors ils se dévoraient eux-mêmes. C'est le début de la digestion, et à partir de là,

c'est ce que les animaux sont devenus pour remplacer les générations précédentes qui sont mortes.

J'ai commencé à m'étouffer. J'ai senti mon corps m'abandonner. J'ai essayé de me lever de mon lit. Je n'ai pas pu. Je n'ai pas pu. J'ai regardé la machine à côté du lit qui respirait pour moi. J'ai essayé de la soulever, mais je ne suis pas parvenu à la tenir entre mes bras. Je n'arrêtais pas de me dire : « Ce n'est pas bien. Ce n'est pas bien ». J'ai regardé à nouveau la machine et les larmes ont commencé à couler de mes yeux. Ma poitrine me brûlait. Je suis resté allongé pendant plusieurs heures jusqu'à ce que je me lève. Je me sentais comme un parfait étranger à l'hôpital.

Ce n'est qu'un mois plus tard que j'ai vu le Dr Corry au chevet de mon lit. C'était un petit singe avec un cœur énorme. Je lui ai demandé : « Comment cela se fait-il ? » Il m'a répondu : « C'est dû à ce que vous faites en ce moment. Vous vous faites du mal. » Il m'a demandé de répéter : « Le temps est venu pour moi d'abandonner ma vie. Je dois prendre ma vie d'une nouvelle manière. Je me sens comme une personne différente. Je suis un patient anesthésié. Je dois l'être. Ma famille ne s'intéresse pas à moi. Je ne peux pas le faire moi-même. Je suis une personne compliquée. Je suis une épave nerveuse. Il faut que je me lâche. Je ne sais pas comment vous l'exprimer. Je suis une personne normale avec une vie ennuyeuse. »

J'ai essayé à nouveau de répéter ces phrases. J'ai encore essayé de répéter. J'ai encore essayé ces phrases. J'étais sûr de me relever. Mais j'étais sûr que je ne le ferais pas. J'étais sûr que je ne mourrais pas. Mais j'avais tort. J'ai essayé à nouveau.

Le temps a passé et la première de mes dernières tentatives infructueuses s'est faite dans mes propres poumons. Je n'ai pas été très patient, mais je n'ai pas fait grand-chose. Je n'ai même pas essayé de mettre un coton sur le lit, je l'ai gardé au-dessus de la tête du lit. La machine fonctionnait toujours, elle était là, mais maintenant elle ne servait plus à rien. Je l'ai gardé sur le lit et j'ai continué à tousser, j'avais peur de mourir si je continuais à faire ça... Finalement, j'ai commencé à me lever du lit et j'étais prêt à mourir. J'étais sur le point de rendre mon dernier souffle. J'ai ressenti ce grand soulagement à l'idée de mourir, et aussi de ne pas mourir, et pourtant il me semblait impossible de sortir de la chambre. Je toussais encore, et je savais que j'allais mourir. J'étais certain que si je continuais à tousser, je mourrais. J'étais sûr que si je continuais à tousser, je mourrais une nouvelle fois.

Ce texte n'a pas été écrit par un être humain, moi, ou par une machine, mais par un complexe entre les deux. Je me suis perdu dans la machine qui s'est perdue en moi. J'aimerais vous rendre sensible et palpable, cette émotion d'écriture-là, cette fiction grise et troublée. Il ne

m'est possible que d'en parler de façon expérientielle en rendant compte des sentiments divers que j'ai traversés.

Voilà la procédure : je construis un corpus de plusieurs centaines de livres trouvés sur le réseau et totalisant en texte brut 300Mo. Cette bibliothèque comprend des textes de biologie, de géologie et de médecine du XVIIIe et XIXe siècles, ainsi que des ouvrages de Philip K. Dick, Le Livre de l'intranquillité de Pessoa, Compagnie de Beckett, Éden, Éden, Éden de Guyotat, beaucoup d'autres encore. Écrire, c'est être accompagné par une multitude de lectures. La constitution de ce corpus répond à une intuition, j'imagine ce que donnerait leur rencontre et croisement, leur hybridation ou leur métamorphose. J'essaie de constituer une atmosphère intertextuelle.

Puis vient le moment de l'apprentissage, quelques heures, où j'observe l'évolution du texte de cycle en cycle, jusqu'à atteindre un résultat intéressant, à mi-chemin entre la répétition et la surprise, la différence dans la répétition si vous voulez pour adapter cette tension deleuzienne. Mon espace latent est prêt.

Ensuite, l'écriture commence : je donne au logiciel des débuts de phrases, il poursuit, et au moment où il génère une suite intéressante, c'est-à-dire qui provoque en moi l'idée d'une suite, je sélectionne le morceau, j'écris la suite, je génère encore, et ainsi de suite. J'obtiens par ce mélange et cette poursuite des phrases, un brouillon : j'ai suivi la machine qui m'a suivi. Je relis, j'ajoute des éléments, des répétitions justement pour tenir un fil ou pour le lâcher. Ce n'est pas un cadavre exquis mais une co-imagination, anthropotechnologique donc. Que se passe-t-il ? Ai-je été soumis à l'IA ou l'ai-je détournée et lui ai-je résisté ? Ni l'un ni l'autre. Était-ce une boîte noire ? Ai-je réussi à créer un cube blanc, enfin transparent et explicable ? Ni l'un ni l'autre. L'art n'aurait jamais été question de signification, mais d'une promesse à jamais différée. La fiction est grise au sens où il y a une boucle rétroactive entre mon imaginaire et... et quoi au juste ? L'IA ? Non, plus exactement la mémoire de plusieurs milliers de textes. C'est une dimension mémorielle qu'on oublie souvent dans l'IA actuelle. Or la mémoire est précisément, dans notre capacité à la répéter et au cœur de la répétition, de la différer, au sens derridien, à produire de la culture. Celle-ci n'est pas une simple répétition à l'identique et n'est pas une pure nouveauté – qu'est-ce que cette pureté voudrait d'ailleurs dire ? –, mais une zone grise entre l'avoir été et ce qui pourrait avoir lieu : une métamorphose. Je me laisse hanté par la machine, « mon » imaginaire est sous influence : Ghost in The Shell.

Voilà donc mon émotion : perdre mon autonomie souveraine, qui n'a jamais été, me perdre dans et avec nos mémoires, entrer dans un imaginaire qui est le mien sans m'appartenir,

devenir ce que nous avons toujours été, une multitude ayant l'anonyme en partage.

C'est en 1952 que Marvin Minsky, alors travaillant au Bell Labs, invente le concept de machine inutile qui s'allume et s'éteint à intervalles irréguliers et le désigne aussi par le vocable "Leave Me Alone Box", selon une autre topologie de la solitude. Claude Shannon fera ses propres versions de cette machine et l'une d'entre elles sera vue par Arthur C. Clarke au MIT, qui écrira : « Il y a quelque chose d'indicible et de sinistre à propos de cette machine qui ne fait rien, absolument rien, si ce n'est s'éteindre elle-même. »¹⁵ Minsky a également créé une machine à gravité qui sonne si la gravitation universelle change. Or, cette possibilité n'est pas scientifiquement envisageable dans le futur. Par cette seconde machine, il produit une science fictive, non pas seulement une alternative, mais une possibilité hors la loi qui constituerait une contingence pure extra-scientifique et un possible au sens fort du terme. Il y a d'un côté une machine qui ne cesse de répéter une opération inutile, il y a de l'autre une technique programmée pour quelque chose qui n'arrivera jamais.

Chaque machine provoque une fascination parce qu'elle extirpe sa forme et sa matière de l'instrumentalité, et laisse entrevoir une singulière temporalité. Il y a une boucle qui se répète et une boucle qui n'arrivera jamais, la fascination est aussi liée à l'affirmation de la contingence circonstancielle et au mouvement machinique. Celui-ci est un flux bordé par la répétition et la non-répétition, par le possible et l'impossible. Comment comprendre que le flux arrive et n'arrive pas, non comme deux moments distincts, mais simultanément ? De quelle façon penser cette contradiction ? Cette dernière est déterminée par une identification et une stabilisation des états qui est sans rapport avec l'advenue des états eux-mêmes. Les flux sont un entrechoc entre le monde et la perception. La simultanéité du possible et de l'impossible n'est pas un phénomène en soi, mais le résultat d'une distance entre ce qu'il y a à percevoir et ce que nous percevons, entre la donation et le donné, bref elle est déterminée par le dispositif d'observation.

On comprendra donc ma distance quant à l'art cherchant à être utile et fonctionnant de façon critique, entendue comme une position extérieure reproduisant l'Oeil incorporel du cube blanc. La fiction grise est quant à elle comme Deep Dream, répétition et différence tout à la fois, récursive et contingente s'enfonçant indéfiniment dans ses conditions de possibilités qui, loin d'être autonomes, sont le monde en tant que monde. Nous ne sommes décidément pas indemnes. Ce virus nous le rappelle. Nous ne pourrons plus sortir de cet espace, nous ne

¹⁵ Abigail, P. (12 mars 2013). « Looking for Something Useful to Do With Your Time ? Don't Try This ». Wall Street Journal.

pourrons plus le nier.