

L'intériorité technologique

Pierre Cassou-Noguès (Univ Paris 8)

Mon but, dans cet article, est d'interroger la façon dont les technologies contemporaines transforment le statut de l'intériorité. Par intériorité, j'entends la sphère du mental, ou de l'expérience en première personne, ces pensées, ces émotions, qui, dans notre forme de vie habituelle, sont supposées se dérouler « dans notre tête ». J'utilise volontairement cette expression métaphorique, non pas pour affirmer que cette intériorité aurait son siège dans le cerveau plutôt que dans l'estomac, ou dans le cerveau plutôt que dans l'esprit, ou dans une âme, mais pour traduire dans la langue quotidienne cette sorte de convention qui invite à considérer que nos pensées et émotions ont lieu dans une sorte d'intérieur (derrière notre front en quelque sorte) auquel nous avons un accès privilégié, immédiat et que nous pouvons masquer à ceux qui nous entourent : je suis censé savoir, en général, plus ou moins, ce que je pense ou ce que j'éprouve, et pouvoir, en général, ou plus ou moins, le cacher aux autres. Or, ce sera mon premier point, les technologies contemporaines transforment cette forme de vie et déplacent cette intériorité, de notre tête sur un écran. Nous acceptons que des machines puissent décrypter mieux que nous le faisons nous-mêmes ce que nous pensons et éprouvons, de sorte que ce serait sur un écran, plutôt que « dans notre tête », que nous découvririons nos pensées et émotions.

Ce déplacement a toute une série de conséquences quant au rapport du sujet à ses pensées et émotions, à la possibilité d'en mesurer objectivement l'existence et l'intensité, ou à la nature de l'inconscient.

J'ajouterai un mot sur le statut de mon propre discours et de ce « nous » dans « nous acceptons que les machines ... ». Il s'agit certes d'un raccourci. Je ne prétends pas faire de la sociologie et analyser nos croyances. Evidemment, ce serait tout à fait légitime et instructif mais c'est une autre tâche que la mienne. A mes yeux, et bien que je ne puisse pas le défendre ici, la philosophie a pour tâche d'analyser le possible tel qu'il s'exprime dans les fictions qui fonctionnent ou auxquelles nous adhérons¹ : récits littéraires, installations artistiques mais aussi articles de presse, « ted talks » de scientifiques. Ces fictions n'expriment pas forcément nos croyances, au sens strict, mais, dans la mesure où elles fonctionnent et nous y adhérons, elles mettent en scène les formes de vie qu'impliquent les technologies contemporaines, et les relations entre les concepts qui s'y instaurent. Elles mettent en évidence des systèmes conceptuels et la façon dont ceux-ci se transforment : ce que deviennent le « mental », « la pensée », « l'émotion ». Systèmes de concepts dans lesquels nous pouvons être nous-mêmes pris, alors que, si l'on nous interrogeait, nous

¹ Cf. notamment, Cassou-Noguès, Pierre. 2016. "A Quasi-Phenomenology with Examples. The Invisible Man, the Posthuman and the Robot." *Iride. Filosofia e discussione pubblica* XXIX/3:299-314.

pourrions ne pas y reconnaître nos croyances.

J'appelle ailleurs « syndrome du thermomètre » ce déplacement de l'intériorité qui la confie à la machine, en référence à ce geste de regarder le thermomètre pour savoir si nous avons nous-mêmes chaud. Le thermomètre mesure la température de l'air mais nous y jetons un œil (parfois) non pas pour savoir s'il fait chaud mais pour affirmer « J'ai chaud » (ou, plus exactement, pousser un soupir « Oh quelle chaleur », qui ne signifie pas « il fait très chaud » mais « j'ai très chaud ».) Avoir chaud est une expérience intérieure, une expérience en première personne, qui dépend seulement à peu près de la température extérieure. Or nous confions à une machine (rudimentaire en l'occurrence, un simple thermomètre) le soin de fixer cette expérience que nous devrions pouvoir seuls saisir dans sa réalité. Ce geste est évidemment anecdotique dans notre forme de vie habituelle mais il prend toute son ampleur dans les neurosciences et l'analyse des données. Je voudrais en donner deux exemples. Le premier est explicitement fictif.

Dans le cadre d'une exposition au Louvre en 2012, Jean-Philippe Toussaint a en effet installé « une machine à lire dans les pensées ». Celle-ci ressemblait un peu à une cabine de douche dans lequel le sujet s'installait pour lire un roman, portant un casque pour électroencéphalogramme, comme un bonnet de bain d'où partiraient une multitude de fils électriques. Ceux-ci cheminaient jusqu'à un écran, fixé sur le mur, sur lequel devaient s'afficher les images qui passeraient dans la tête du sujet pendant sa lecture.

Quand on visitait le musée, la cabine était fermée, par une sorte de cordon, comme si elle était temporairement hors service. Un lecteur, la veille, avait-il souffert d'une poussée d'imagination, qui aurait fait sauter le dispositif ? Il ne restait qu'à visionner sur une tablette un peu plus loin quelques vidéos réalisées d'après les pensées enregistrées dans la tête de différents personnages : les pensées de Toussaint lui-même, celles de Pierre Gagnaire (cuisinier étoilé), Tom Novembre (chanteur pop), Pierre Bayard (écrivain), etc. La morale de l'affaire, on le comprenait, était que de tels personnages, un peu exceptionnels mais pour des raisons différentes, auraient à la lecture du même texte des images différentes. Le cuisinier verrait-il un mélange d'épices quand le chanteur imaginerait une histoire d'amour ? Ou l'inverse. Il faut entendre « images » en un sens strict. Les vidéos montraient des images mobiles (par exemple, dans l'esprit de Toussaint, une femme courant dans un couloir) et non pas des pensées avec des fragments de discours.

Toussaint n'a pas beaucoup discuté de sa machine. Elle faisait partie d'une exposition plus importante. Le catalogue montre des photographies de l'écrivain enfourné dans un scanner pour une IRM fonctionnelle mais ne dit rien de la machine elle-même. Dans un entretien, Toussaint remarque néanmoins :

« L'installation de la salle 23 montre comment la réflexion sur la lecture peut s'enrichir de ces

moyens technologiques. C'est une 'machine à lire dans les pensées' installée dans une cabine de douche. Dans cette cabine, il y a un siège, et un casque d'électrodes relié à un amplificateur, lui-même relié à un écran sur lequel les gens pourront voir ce qui se passe dans leurs pensées lorsqu'ils sont en train de lire. Ce n'est pas tout à fait réel, c'est un leurre scientifique qui rappelle que la lecture est une activité cérébrale intense. On ne l'exprime jamais clairement, et je voulais profiter des avancées technologiques pour le mettre explicitement en scène. »²

Il peut y avoir dans ce passage une certaine ironie. Dans plusieurs de ses romans (*La télévision* notamment³) Toussaint décrit des appareils technologiques dans une distance critique. Quoi qu'il en soit, et pour le dire d'emblée, la machine est un « leurre ». Elle ne fonctionne pas. Une telle tentative, pour décrypter les images que nous entretenons en lisant un roman, exigerait un scanner d'IRM fonctionnelle, et non un simple casque pour électro-encéphalogramme. Et, d'autre part, elle exigerait que ce décodeur soit entraîné, réglé, que l'on ajuste ses paramètres, pour chaque sujet. Il serait donc impossible de laisser la machine à la disposition de visiteurs qui s'y installent à tour de rôle pour voir leur pensées intérieures affichées sur l'écran. Néanmoins, l'installation de Toussaint rejoint (dans la fiction donc) une multitude d'expériences de neuroscience, qui visent à capter dans le cerveau des images : les images que le sujet observe sur un écran, les images qu'il visualise intérieurement, les images qui entrent dans un rêve, etc.⁴

Ce qui m'intéresse surtout est cet « écran sur lequel les gens pourront voir ce qui se passe dans leurs pensées lorsqu'ils sont en train de lire ». Dans le dispositif de Toussaint, l'écran, sur le mur à côté de la cabine dans laquelle est installé le lecteur, est tourné vers les spectateurs qui assistent à l'expérience. Ce sont d'abord qui eux qui voient (ou verraient) les images qui passent dans la tête du lecteur. Celui-ci doit, après avoir terminé sa lecture, sortir de la cabine pour alors visionner le film qu'on lui repasse. Mais, ce film, il devrait l'avoir déjà vu, dans sa tête pendant qu'il lisait. La machine saisit les images qui lui passent par la tête : il ne devrait rien apprendre de nouveau en les revoyant à l'écran. Ou alors, s'il faut prendre les mots de Toussaint à la lettre, si c'est sur cet écran que nous verrons « ce qui se passe dans nos pensées », nous avons accepté que nous ne savons pas, ou ne savons pas précisément, ce qui se passe dans nos pensées, et avons confié à une machine le soin de le déterminer.

Si je ne peux pas ici discuter d'exemples trop nombreux, l'hypothèse que je ferai est que le discours qui accompagne les neurosciences, les articles de journaux mais aussi les vidéos des savants, les introductions plus générales qui ouvrent les articles techniques, donnent corps à l'idée que, en effet, une machine est en principe susceptible de décrypter mes pensées, de sorte qu'elle peut fort bien servir « d'instrument d'introspection » pour fixer des contenus de pensée dont je

2 <http://livre-louvre.arte.tv/entretien/>. Consulté en septembre 2015, a disparu depuis.

3 Jean-Philippe Toussaint, *La télévision*, Paris, Minuit, 1997.

4 Je permets de renvoyer à mon essai, *Lire le cerveau*, Paris, Seuil, 2011.

n'aurais pas moi-même conscience ou dont je n'aurais pas bien, ou précisément conscience. C'est une généralisation de ce geste que j'évoquais au début, celui de regarder le thermomètre avant de dire « ah, ce que j'ai chaud » : la machine a pour fonction de fixer une expérience qui devrait relever de la première personne.

Ce déplacement de l'intériorité qui sort de nos têtes pour se loger sur l'écran (pour reprendre la formule de Toussaint, « un écran sur lequel les gens pourront voir ce qui se passe dans leurs pensées ») a certainement d'abord été engagé dans les neurosciences mais ce ne sont pas celles-ci qui en tirent bénéfice. Leur appareillage est trop coûteux pour sortir des laboratoires, alors que nous portons dans notre poche les machines sur lesquelles s'appuie l'analyse des données.

Le second dispositif que j'évoquerai est une application dans le domaine de la « santé mentale numérique » (en anglais, « *digital mental health* »). Ce titre, tapé sur la barre de recherche de Google Play donne une liste d'environ deux cents applications. L'application « Cogito Companion » est disponible sur Google Play. Son nom même est évocateur. Comme si la machine se faisait à ce point intime qu'elle accompagnait jusqu'à ce noyau de notre vie intérieure, ce « Je pense » dont Descartes déduisait sa propre existence dans l'abstraction du monde environnant.

Il s'agit en réalité d'une application développée par la société Cogito. Celle-ci s'est spécialisée dans l'analyse de la voix pour améliorer les performances des opérateurs téléphoniques : commerces à distance, services après-vente, compagnies d'assurance⁵. L'algorithme analyse les intonations de l'opérateur et du client durant leur conversation, pour évaluer la satisfaction du client et donner des conseils à l'opérateur : faire ressortir de l'émotion, prendre un ton plus ferme, ou plus empathique, ralentir son débit ... La satisfaction du client est mesurée par un chiffre sous les yeux de l'opérateur qui cherche à faire grimper ce score en suivant les conseils qui glissent sur son écran. En même temps, l'algorithme isole des mots clés dans la conversation et ouvre les pages contenant les informations dont l'opérateur peut avoir besoin : « police d'assurance », « ma connexion ne fonctionne plus », etc.

L'application Companion utilise également un algorithme d'analyse vocale. Il ne s'agit plus alors d'évaluer la satisfaction du client mais l'humeur de l'utilisateur lorsqu'il parle dans son téléphone. Celle-ci est représentée par un ruban orange qui s'étend à travers l'écran (selon le principe : plus c'est long, plus c'est bon) complété par un petit bonhomme dont le sourire peut s'inverser en une moue boudeuse. L'utilisateur peut observer l'évolution de son humeur et de son activité physique, son sommeil, sa sociabilité, sur une semaine, un mois, un an, sous la forme d'un graphique qui lui rappelle ses hauts et ses bas.

« Cogito Companion est un outil qui vous permet de mieux comprendre votre santé. Avec Companion, vous avez une idée précise de votre santé, dans votre poche. Companion fournit un

5 <http://www.cogitocorp.com/product/>

feedback significatif sur votre santé. Vous n'avez besoin que de votre smartphone. »⁶

D'autres applications, dans le domaine de la santé mentale numérique, utilisent différentes données, récupérées dans le téléphone pour mesurer l'activité physique de l'utilisateur (grâce à l'accéléromètre), sa sociabilité (en captant les sons environnants dans le microphone), son temps de sommeil (grâce au détecteur de lumière ambiante), et sont capables de compiler ces données pour évaluer, par exemple, le degré de stress de l'utilisateur⁷.

Des applications fonctionnant sur les mêmes principes ont également été testées pour déceler les fluctuations d'humeur chez des sujets bipolaires ou observer des sujets suicidaires. Évaluer leur mérite reste problématique⁸.

Cependant, je ne veux pas nier en général l'utilité de ces applications. Je ne prends aucunement position quant à leur capacité à déceler en effet nos fluctuations d'humeur. Je veux seulement souligner que leur usage correspond à une transformation de notre forme de vie. Lorsqu'un ami me dit « tu es triste ? », et que je lui réponds « non pas du tout, je rêvassais », il peut croire que je lui mens, et veux cacher ma mélancolie mais, s'il accorde que je suis de bonne foi, il reconnaîtra s'être trompé : je n'étais pas triste, j'étais seulement absent. Parce que je suis censé savoir comment je me sens : j'ai moi-même accès, et je suis le seul à avoir immédiatement accès à mes pensées et mes émotions. Or c'est cette forme de vie que ces applications transforment. En utilisant Cogito Companion, j'accepte que je ne sais pas moi-même comment je me sens, ou que je peux tromper quant à mon humeur et que la machine évalue mon humeur mieux que je ne le fais moi-même. C'est comme si mon ami ajoutait : « Tu te trompes, tu es triste, je te le dis, crois-moi ». Or ce serait aberrant de la part d'un ami mais cela ne l'est nullement venant d'une machine. Ce n'est pas que nous n'avons plus d'émotions, ou que cette sphère de pensées et d'émotions que nous pouvons cacher à nos proches se soit effacée, ce n'est pas que soyons devenus des humains « sans intériorité » pour reprendre un titre de Philippe Breton⁹. C'est seulement que notre intériorité, nos pensées et nos émotions, se déplacent pour se faire reconnaître non plus dans notre tête mais sur un écran.

Dans cette perspective, il faut noter que, une fois sorti du laboratoire, le bon fonctionnement d'une telle application ne se mesure pas par l'adéquation de son diagnostic avec un état mental qui serait bien défini. Si j'utilise Cogito Companion, c'est parce que je ne reconnais pas moi-même mon humeur. Il ne s'agit pas donc pas de savoir si Cogito Companion me donne un score d'humeur 6 quand je me sens content à un degré de 6/10. La question est de savoir si je peux m'adapter au

6 « Cogito Companion is a tool that helps you better understand your health. Companion allows you to: *Have health insights in your pocket* Companion provides meaningful feedback on your health. All you need is your smartphone! »

7 Cf. en particulier, l'application pionnière StudentLife : <https://studentlife.cs.dartmouth.edu>

8 Or, F., Torous, J., & Onnela, J.-P. (2017). « High potential but limited evidence: Using voice data from smartphones to monitor and diagnose mood disorders ». *Psychiatric Rehabilitation Journal*, 40, 320–324.

9 Philippe Breton, *L'homme sans intériorité*, \$\$\$

diagnostic de l'application et me sentir relativement content quand l'application me donne un tel score. Cela suppose évidemment que l'application ne commette pas trop souvent d'erreurs trop grossières, pour nous déclarer en grande forme quand nous sommes vraiment déprimés.

Je n'affirme pas que, en réalité, nous savons toujours et parfaitement comment nous nous sentons. Je dis seulement (en suivant Wittgenstein) que c'est une sorte de convention, une forme de vie, que en règle générale nous sommes censés connaître nos pensées et nos émotions. Et, justement parce que c'est une sorte de convention, nous pouvons la modifier et poser collectivement, sans forcément le rendre explicite, que certaines machines connaissent nos pensées et nos émotions mieux que nous-mêmes. Il suffit que ce nouvel arrangement fonctionne dans la vie et ne nous conduise pas trop souvent à des aberrations trop flagrantes.

Ce déplacement ne concerne pas seulement les applications dans le domaine de la santé mentale digitale mais l'ensemble des procédures par lesquelles des algorithmes sont censés déterminer nos pensées, nos émotions, nos intentions. Il peut recouper certaines problématiques de la sociologie ou de la psychanalyse en les orientant alors entièrement. Un sociologue peut déterminer (en fonction de nos revenus, notre niveau d'étude, notre lieu d'habitation, etc.) pour qui statistiquement nous allons voter alors même que nous ne le savons pas encore, mais, dans la perspective de ce déplacement de l'intériorité, il s'agirait de déterminer pour qui nous *voulons*, ou *avons l'intention de*, voter, et cela alors même possiblement nous ne le savons pas nous-mêmes. De même, la psychanalyse met certes en évidence une dimension inconsciente de l'esprit. Cependant, dans sa version habituelle, elle s'appuie sur l'ambiguïté des contenus de pensée et ne prétend pas déceler dans notre esprit de nouveaux contenus de pensée. Pour revenir à la machine de Toussaint (on pourrait aussi évoquer les machines à décrypter les rêves¹⁰), un psychanalyste pourrait dire que, sous ce visage barbu qui me vient à l'esprit en rêve, ou lorsque je lis le roman, je pense à mon père autant qu'à l'un de mes collègues. Mais le psychanalyste ne dira pas que j'ai rêvé d'un visage barbu alors que je ne m'en souviens pas moi-même ou alors que je croyais avoir rêvé d'une plage au soleil. La psychanalyse situe l'inconscient dans l'ambiguïté des contenus de pensée, dont « la façade », pour reprendre l'expression de Freud, cache une multitude d'autres sens. Avec les machines à lire les pensées, l'inconscient devient un autre contenu de pensée que nous ne savions pas entretenir.

De façon très pragmatique, les applications, dans le domaine de la santé mentale digitale, pour poursuivre sur le deuxième exemple évoqué, permettent de récolter des données sur les habitudes et la psychologie du sujet qui, piratées ou revendues, sont susceptibles de rendre un message publicitaire, commercial ou politique, particulièrement efficace. Plus profondément, dès lors que l'on accepte que les pensées, émotions, intentions, ne relèvent plus forcément d'une

10 T. Horikawa, M. Tamaki, Y. Miyawaki, Y. Kamitani, « Neural Decoding of Visual Imagery During Sleep », 340/6132, 2013, pp. 639-642 ; F. Siclari, B. Baird, L. Perogamvros, G. Bernardi, J.J LaRocque, B. Riedner, M. Boly, B. R Postle, G. Tononi, « The neural correlates of dreaming », Nature Neuroscience 20, 872-878 (2017).

expérience en première personne, à laquelle le sujet a seul immédiatement accès, mais peuvent être décryptées, avec plus de précision peut-être, par une machine, c'est toute la question du statut du sujet et de ses états mentaux qui est remise en question.

D'après un article du *Scientific American*, un chauffeur de poids lourd, Carl Koch, blessé au poignet un an auparavant par de l'asphalte en fusion et se plaignant de douleurs chroniques, a attaqué son ancien employeur, Tucson, en Arizona, en 2005. Joy Hirsch, neurologue dans un centre de recherche à l'université de Columbia à New York, l'a soumis à une expérience de neuro-imagerie, au moyen d'une IRM fonctionnelle, et établi qu'un léger contact sur le poignet blessé produisait un signal dans des régions cérébrales associées à la douleur, alors qu'un même contact sur l'autre poignet n'en produisait pas. Le juge a accepté cette mesure objective de la douleur et accordé au plaignant une indemnité de 800 000 dollars¹¹.

Au procès de Carl Koch, l'employeur, Tucson, fait aussi témoigner un neurologue, Sean C. Mackey, de l'université de Stanford. Celui-ci ne met pas en question que les expériences imaginées par son collègue prouvent que le patient ressent de la douleur. Il argue que ces expériences ne prouvent pas que cette douleur soit due à l'accident subi. Parce que le patient pourrait déclencher lui-même dans son cerveau, par un mouvement quasi-volontaire, le signal de la douleur, ou alors la crainte de la douleur dans le cadre de cette expérience pourrait déclencher le signal qui ne se produirait pas dans la vie courante.

La même année, en 2005, Mackey a en effet participé à une curieuse étude. Il s'agit encore de brûlure. Allongé dans le scanner, le sujet tient à la main un objet métallique que l'expérimentateur chauffe pour produire une douleur. Le scanner repère l'intensité de l'activité dans les régions associées à la douleur. Cette intensité est représentée sur un écran, en temps réel, sous le regard du sujet, par des flammes plus ou moins hautes. Le but de l'expérience est d'apprendre aux sujets à contrôler leur cerveau et leur douleur : penser à autre chose par exemple, trouver une façon d'activer d'autres régions du cerveau pour diminuer le signal dans les régions associées à la douleur. Et l'écran sous leurs yeux leur montre leurs progrès : les flammes diminuent quand le signal dans ces zones de douleur diminue d'intensité. Et, en effet, les sujets affirment que la douleur alors baisse en intensité.

La même expérience est tentée sur des patients souffrant de douleurs chroniques. Et eux aussi apprennent à contrôler leur cerveau et, par là, à diminuer la douleur par ces exercices mentaux. « Nos résultats montrent que des sujets peuvent, après un apprentissage approprié, obtenir un contrôle volontaire de l'activation d'une région spécifique de leur cerveau [...] et que l'effet est assez puissant pour influencer sur de sévères douleurs chroniques. »¹²

11 S. Reardon, « Pain Cases May Usher Brain Scans into Court Rooms », *Scientific American*, 2 mars 2015.

12 R. Christopher deCharms, « Control over brain activation and pain learned by using real-time functional MRI », *Proc Natl Acad Sci U S A*. 20 décembre 2005, 102(51): 18626–18631.

Ce déplacement de l'intériorité ne signifie donc pas forcément une perte de contrôle du sujet. Mais il met certainement en question l'autonomie du sujet et le rapport du sujet à ces propres états mentaux, sensations, émotions, intentions, images, rêves, pensées ... Ceux-ci ne restent plus « dans nos tête », dans une intériorité à laquelle nous aurions seuls un accès immédiat, mais se dévoilent sur l'écran de machines introspectives, ce qui engage de nouvelles procédures de mesure, comme de nouvelles possibilités de surveillance et d'influence et, très largement, une autre sorte de rapport à soi.