

Pour tirer un résumé rapide de la question du pouvoir qui est intervenue à plusieurs reprises dans notre analyse de la catégorie des métatechnologies et de différents exemples, revenons pour un instant au texte de Deleuze sur Foucault. Les deux grandes catégories avancées par Foucault, à des degrés variables, dans la plupart de ses travaux sont le *savoir* [1966 et 1969] et le *pouvoir* [1975]. De façon peu prudente, nous pouvons projeter ce binôme sur l'opposition entre le discursif et le non-discursif, l'épistémè et le dispositif, et, pour replonger dans l'univers des SIC, l'information et l'action. Dans l'interprétation deleuzienne de ce dualisme, un troisième concept se glisse derrière les deux premiers, et joue le rôle d'une « machine abstraite », d'un principe organisateur qui « ignore toute distinction de forme entre [...] une formation discursive et une formation non-discursive » [Deleuze 1986, p. 42] ; l'auteur nomme ce tiers le *diagramme*. Ce terme nous intéresse non seulement parce qu'il permet de penser un dispositif technique en tant que combinaison d'une structure d'action, infrastructure ou agent, et d'une réalité symbolique, qui implique une proposition de sens, mais aussi parce qu'il réunit savoir et pouvoir dans un principe fédérateur où les deux pôles se reflètent l'un dans l'autre. Nous avons dit à plusieurs reprises que les métatechnologies ne se résument pas à une théorie technique, mais qu'elles mobilisent également une théorie du social et / ou de la signification. Le concept du diagramme nous permet de penser ces deux dimensions ensemble, de comprendre que le dispositif est émanation et incitation d'un savoir en même temps qu'il est l'enfant et le géniteur d'un pouvoir. Il fonctionne comme machine de vision (savoir) et comme dispositif d'action (pouvoir) et, en tant qu'objet fait et non devenu, sa genèse est alors marquée par la rencontre entre un discours et une matière en amont de sa naissance. Les métatechnologies pensées en tant que diagrammes ouvrent sur un terrain proprement philosophique que nous ne pouvons pas poursuivre de façon conséquente dans ce travail ce qui ne devrait pas nous empêcher d'indiquer où un tel voyage pourrait amener :

« [Le diagramme] ne fonctionne jamais pour représenter un monde préexistant, il produit un nouveau type de réalité, un nouveau modèle de vérité. » [ibid., p. 43]

Bien que nous ayons donné une définition générale de ce que nous entendons par « métatechnologie » et que nous ayons avancé un certain nombre d'exemples, le côté technique de notre problème reste étrangement flou. Cela s'explique en priorité par notre réticence à définir ces objets en termes purement techniques. Mais une raison plus profonde explique l'absence d'une désignation plus précise des objets que nous visons : les métatechnologies entrent dans l'espace du culturel et du social qui refuse de rentrer dans des tableaux trop rigides. Le concept de Web 2.0 témoigne de la nature dynamique des interactions entre la technique et les usagers ainsi que de la difficulté de fixer ce terrain sur un appareil de catégories rigides. Nous préférons donc traiter la question de la technique d'une façon plus théorique en discutant deux axes d'interrogation qui nous semblent essentiels pour comprendre l'évolution technique telle qu'elle se fait aujourd'hui. Quelles

sont la nature et l'importance des tâches qui peuvent être déléguées ? Quel est le degré d'autonomie possible ?

1.4.1 Que peut-on déléguer ?

Il y a au moins trois façons de répondre à la question de savoir ce qui peut – en principe ou en pratique – être délégué par les êtres humains à la machine. Bien évidemment, une telle question se continue dans le débat sur la possibilité d'une intelligence artificielle, la tâche à déléguer étant limitée par les capacités / performances de la machine censée l'accomplir. Deux de ces trois positions relèvent donc directement des discussions autour de cette vieille idée de l'automate pensant.

La première façon de voir le problème est celle des représentants de l'intelligence artificielle classique [McCulloch / Pitts 1988 ; Turing 1950 ; McCarthy 1959 ; etc.], parfois appelée « IA forte » (*strong AI*)¹³¹, dont la position consiste à regarder la question de l'intelligence, et donc aussi celle de la capacité / les performances de la machine dans un contexte fonctionnel plus précis, comme une question purement technique. Cette transposition se fait par le biais d'une conception de l'intelligence comme mouvement de résolution de problèmes d'ordre symbolique (*problem-solving*) qui avance par la formulation et l'exécution de *plans*, concept qu'on identifie à la notion d'algorithme [cf. Agre 1997]. Cette opération, plus sémantique que conceptuelle, permet par la suite de voir la quête d'intelligence comme l'accumulation de programmes toujours plus performants qui s'approchent davantage à chaque nouvelle étape de l'idéal incarné par l'être humain. Dans cette perspective – dont Agre [1997] a sûrement fait la critique la plus puissante et la mieux informée – la question de ce qui peut être délégué à une machine impose de diviser le champ des tâches à transférer en « déjà possible » et « pas encore possible ». Le passage du deuxième champ au premier, c'est-à-dire la simulation future de comportements encore hors de portée aujourd'hui, est perçu comme une simple question de temps. Nous avons déjà rencontrée cette position dans les articles de Turing [1950] et Licklider [1960] pour qui il ne fait aucun doute qu'un jour la machine sera au moins aussi intelligente que l'être humain. Alors, *tout* pourra lui être délégué et l'être humain sera obsolète. Bien que la conviction que ce but puisse être atteint bientôt soit peu répandue dans le champ de l'IA, des représentants de l'*IA forte* comme Ray Kurzweil [1999] et Marvin Minsky [2005] sont persuadés qu'à un moment donné, tous les problèmes seront résolus.

¹³¹ Afin de nuancer l'argument, il faut ajouter qu'un nombre important d'approches – souvent connexionnistes – ne partagent plus certaines préconceptions de la GOFAI (*Good Old-Fashioned Artificial Intelligence*), c'est-à-dire l'école classique qui conçoit la question de l'intelligence comme une série de problèmes d'ordre purement symbolique.

Une deuxième perspective sur la question a été formulée dans le domaine philosophique et notamment autour du travail de deux auteurs. John Searle [1984] essaya de montrer dans sa fameuse expérience de la « chambre chinoise » qu'il n'est pas possible de réduire la sémantique à la syntaxe et qu'il faut un cerveau humain pour « produire » un esprit (*mind*).¹³² Hubert Dreyfus [1993] présenta une critique analogue de l'IA, mais en se concentrant sur la mise en cause du paradigme symbolico-informationnel sur lequel repose l'approche informatique de la question de l'intelligence. Ces deux positions philosophiques affirment que les ordinateurs sont fondamentalement différents des êtres humains et ne pourront, de ce fait, jamais simuler une intelligence qui recourt aux fonctions mentales les plus avancées. Selon ces deux auteurs, même les expériences qui suggèrent qu'une intelligence existe dans la machine ne peuvent pas invalider le fait qu'une machine ne produit jamais de « vraie » intelligence mais uniquement des *comportements* prédéfinis par de malins programmeurs. Dans une telle optique, il semble possible de déléguer certaines tâches simples, mais dès qu'une analyse un peu plus « sémantique » devient nécessaire, l'être humain paraît définitivement irremplaçable.

La différence entre ces deux positions parfaitement antithétiques se fonde sur l'opposition directe et complète de leurs conceptions respectives de l'être humain. La première prétend que toute la pensée est réductible à une question de traitement symbolique tandis que la deuxième souligne que le paradigme « cognitif » est naïf et réducteur, qu'il n'aboutira jamais à ce qu'il recherche car il se base sur une méconnaissance complète de la nature humaine. Cette nature ne se résumerait pas en un traitement par processus mais comprendrait des phénomènes psychiques, sociaux, culturels et émotionnels complexes dont la simulation ne pourrait qu'échouer. Les deux positions se rejoignent pourtant par leur posture essentialiste. Elles ont chacune une idée très forte de l'être humain et, par conséquent, du phénomène de l'intelligence. Chacune avance donc des affirmations musclées et dogmatiques sur un terrain qui reste cependant en grande partie en dehors de la connaissance humaine.

Nous ne souhaitons pas nous positionner de l'un ou de l'autre côté de ce clivage, mais plutôt présenter une troisième perspective qui nous semble non seulement plus adaptée à notre propos, mais aussi épistémologiquement plus prudente : la théorie de l'action de Harry Collins et Martin Kusch [1998].

¹³² Descartes [1637] exposait déjà un argument similaire dans sons *Discours de la méthode* : « car on peut bien concevoir qu'une machine soit tellement faite qu'elle profère des paroles, et même qu'elle en profère quelques unes à propos des actions corporelles qui causeront quelque changement en ses organes, comme, si on la touche en quelque endroit, qu'elle demande ce qu'on lui veut dire; si en un autre, qu'elle crie qu'on lui fait mal, et choses semblables; mais non pas qu'elle les arrange diversement pour répondre au sens de tout ce qui se dira en sa présence, ainsi que les hommes les plus hébétés peuvent faire. »

La morphologie des actions

Dans « *The Shape of Actions. What Humans and Machines Can Do.* » (La forme des actions. Ce que les êtres humains et les machines peuvent faire.), Collins et Kusch choisissent une approche très différente des perspectives classiques que nous venons de résumer schématiquement. Le sous-titre de l'ouvrage indique déjà son orientation pragmatique : les auteurs ne se proposent pas de définir ce que c'est l'intelligence, mais de donner un concept puissant mais flexible de « ce que les êtres humains et les machines peuvent faire ». Le point de départ est ainsi marqué et il ne consiste ni en une interrogation sur la cognition, ni en une contemplation de la nature du raisonnement (bien que ces deux dimensions aient joué leur rôle par la suite), mais en l'élaboration d'une *morphologie* des actions intentionnelles. Ce projet ne prend pourtant pas en compte les actions tout court, mais les étudie telles qu'elles se présentent à un observateur externe. Par le biais de la fonction de l'observateur les auteurs distinguent donc deux types fondamentaux : les actions *mimeomorphes* et les actions *polymorphes*. Au premier coup d'œil, la base qui supporte / soutient ces définitions paraît peu intuitive : la séparation se fait par rapport à la nécessité ou l'absence de nécessité d'être membre d'une culture pour reconnaître une action dans un comportement, c'est-à-dire un pur mouvement moteur.

« La distinction entre les actions polymorphes et mimeomorphes peut être vue [...] comme une distinction basée sur la question de savoir qui (en principe) est capable d'y voir de l'ordre ; qui est capable de former des jugements raisonnables à propos de la similitude des comportements qu'il perçoit. Dans le cas des actions polymorphes, seuls les membres d'une culture donnée peuvent voir la similitude correctement ; dans le cas des actions mimeomorphes, des observateurs non membres de cette culture pourraient la voir aussi. » [Collins / Kusch 1998, p. 29]

Une action doit être qualifiée de mimeomorphe quand le comportement répété par quelqu'un qui n'en *connaît pas* la signification culturelle (p.ex. un extraterrestre) donnerait à quelqu'un qui la *connaît* l'impression que la personne est en train de la reproduire. Prenons un exemple : l'imitation purement comportementale de l'action « écrire son nom » (par quelqu'un qui ne sait pas ce qu'il est en train de faire) donnerait à quelqu'un qui sait ce que c'est l'impression qu'il s'agit bien de cette action ; elle est donc mimeomorphe. L'action « divorcer » par contre ne peut pas être reconnue à travers la seule imitation du mouvement corporel impliqué, c'est-à-dire la signature d'une série de documents juridiques ; il s'agit donc d'une action polymorphe. La distinction peut également se révéler à travers un deuxième « test » qui consiste à déterminer qui peut en principe reconnaître que deux actions sont similaires ou identiques. Encore un exemple : deux réalisations d'une action mimeomorphe comme « crier » peuvent être reconnues, pour quelqu'un qui n'est pas membre de la culture local, comme étant le même geste. Dans le cas d'« insulter quelqu'un » c'est tout de suite plus compliqué parce qu'il y a maintes façons de le faire. Mais un membre de la culture occidentale peut tout de suite reconnaître que faire une grimace à quelqu'un ou traiter cette personne d'idiote sont deux émanations de la même action. Pour quelqu'un qui ne connaît pas les codes sociaux impliqués, il n'y a aucune similarité entre les deux et il serait impossible à cette personne de les classer dans le même registre. Voilà pourquoi une action polymorphe peut être réalisée, en principe, par un nombre illimité de *comportements*

(mouvements ou gestes). Et inversement, le même comportement (p. ex. « signer ») peut faire partie d'un grand nombre d'actions polymorphes : écrire un chèque, dédier un livre, divorcer, etc. Dans le cas d'une action mimeomorphe, le rapport entre la dimension d'action et celle de comportement est plus étroit et la répétition exacte du mouvement musculaire peut même être recherchée : tout le secret du golf consiste pour le joueur à reproduire exactement la volée et tout artisan cherche à maîtriser ses gestes avec le plus de précision possible. Un constat s'impose et les auteurs le font explicitement : les ordinateurs peuvent évidemment très bien fonctionner dans le mimeomorphe mais sont incapables du polymorphe.

En quoi cette perspective est-elle donc différente de celle proposée par Searle et Dreyfus ? Ne revient-elle également à installer une ligne de partage dure entre les êtres humains et les machines, une autre frontière essentialiste qui conteste à tout jamais la prétention des ordinateurs à l'intelligence ? Il est certainement vrai que Collins et Kusch ne font pas partie de ceux qui ramènent tout agir dans le registre de la planification et toute pensée dans celui du traitement d'information. Pour eux, une action se fait toujours dans un contexte culturel et sémantique. Il est aussi vrai que leur approche installe une frontière importante entre l'humain et le non-humain. Mais toute une série de subtilités intervient pour la rendre bien plus souple qu'elle ne paraît au premier abord.

La séparation des deux genres d'actions est conceptuelle et non pas empirique. En « réalité », chaque action, chaque pratique, est composée de ces deux genres et Collins et Kusch conçoivent ces actions comme des arborescences dont la partie supérieure relève toujours d'un niveau « culturel » ou polymorphe, et la partie inférieure d'un niveau « moteur » ou mimeomorphe ; au milieu, il existe une épaisse zone intermédiaire où se mélangent les deux formes. Or, il y a dans chaque action polymorphe des actions mimeomorphes et chaque action mimeomorphe s'insère dans le contexte d'une vie sociale et culturelle, c'est-à-dire dans du polymorphe.

Les auteurs ajoutent qu'il est souvent possible d'effectuer une même action de façon plutôt mimeomorphe ou plutôt polymorphe. On peut dire bonjour à quelqu'un en déroulant un programme social (ça va ? – et toi ?) aussi bien qu'en lui posant une question originale et inattendue. Ce choix ne concerne pas seulement les individus, mais aussi les organisations (hiérarchie ou *workflow* libre ?) et même les sociétés (totalitaires ou ouvertes ?), qui peuvent tendre plus dans une direction que dans l'autre. La morphologie des actions n'est pas une ontologie ; elle ne dit pas que le monde des êtres humains est polymorphe et celui des machines mimeomorphe. Selon elle, nos pratiques se déroulent dans un mélange de ces deux dimensions et chaque action connaît une mobilité sur une échelle elle-même variable parce que basée dans des *réseaux de significations* qui changent et évoluent sans cesse.

La frontière entre les deux types ne reproduit finalement pas la différence entre *simple* et *difficile*. Il existe des actions polymorphes apparemment très faciles comme « bavarder » et des actions mimeomorphes extrêmement compliquées comme « gérer un réacteur nucléaire ». La question serait

plutôt de savoir « simple pour qui ? » – question qui ne reflète pas seulement les capacités de l'individu en question mais surtout sa familiarité avec l'arrière-plan culturel qui encadre la tâche.

Actions et délégations

Comment cette approche peut-elle élucider la question que nous nous posons dans ce chapitre ? Quelles sont les tâches que nous pouvons déléguer à l'ordinateur ? Pour répondre, il faut d'abord rappeler que chez Collins et Kusch les machines ne peuvent pas prétendre au domaine des actions polymorphes. Mais en reconsidérant certains des éléments que nous venons de présenter, il faut tout de suite ouvrir une parenthèse :

Du fait du caractère hybride des arborescences d'actions, toujours composées de polymorphe et de mimeomorphe, les éléments mimeomorphes peuvent être délégués à la machine, et les aspects polymorphes conservés du côté des êtres humains. Dans l'un des nos exemples, *Slashdot*, cette distribution se montre très bien : les usagers jugent les interventions selon des critères sémantiques et culturels et le système se charge de générer les niveaux de visibilité à partir de l'ensemble des jugements et d'organiser la distribution des droits de modération. Il en va de même dans le cas de *Google* : les créateurs des sites « votent » l'un pour l'autre en plaçant des liens et le moteur d'indexation les traduit en *PageRank*, mesure de centralité par rapport aux autres sites. Les métatechnologies, qu'elles se trouvent davantage du côté de l'outil ou de celui de l'infrastructure, se trouvent toujours impliquées dans ce mouvement d'hybridation où leurs capacités d'agir sur le plan mimeomorphe s'associent intimement au contexte plus large des pratiques polymorphes. Une bonne partie de la discipline « intelligence artificielle » s'intéresse aujourd'hui à la création de logiciels ou de robots qui s'insèrent de cette façon dans le quotidien des êtres humains.

La deuxième parenthèse concerne plus directement le débat autour du projet de créer une « vraie » intelligence artificielle, à l'image de l'intelligence humaine. Dans ce contexte, la question ne consiste plus à savoir comment déléguer les éléments mimeomorphes d'une arborescence, mais plutôt comment simuler et remplacer les actions polymorphes par des processus algorithmiques. Bien que Collins et Kusch réservent le polymorphe aux êtres humains, ils ne nient pas la possibilité que beaucoup de choses puissent être accomplies en produisant des séries toujours plus fines d'enchaînements d'actions mimeomorphes. Il existe donc une frontière, certes, mais cette frontière est assiégée, repoussée et pratiquement subvertie par l'avancement des savoir-faire. Des applications comme *SpamBayes* et *Google News* montrent que des opérations qui paraissent finalement assez fortement ancrées dans le domaine culturel comme le filtrage d'une boîte électronique où la compilation d'une revue de presse peuvent être déléguées à des programmes, avec des résultats plutôt

remarquables. Cela confirme la prudence des auteurs, qui leur a fait choisir d'installer une large zone intermédiaire (hybride) entre les capacités des êtres humains et celles des machines, tout en gardant l'idée d'une séparation entre les deux. Une telle position pragmatique¹³³ permet d'éviter l'essentialisme des deux perspectives classiques et de garder l'esprit ouvert aux différentes appréciations du problème.

À la fin de leur livre [ibid., p. 198], Collins et Kusch évoquent un détail intéressant et perturbant : selon eux, la réussite du programme de l'IA ne dépend pas seulement des résultats techniques et scientifiques, mais aussi du développement des êtres humains et des sociétés. Si nous prenons le chemin d'un fordisme intellectuel, d'une rationalisation du quotidien dans le sens de la logique mathématique, nous évoluerons dans la direction du mimeomorphe et serons, au bout du compte, bien plus facile à simuler que nous ne le sommes à présent. Dans un esprit provocateur on pourrait dire qu'un soldat qui ne fait qu'obéir à des ordres se prête bien plus à la simulation qu'un artiste.

Le sens des machines

Les métatechnologies convergent parfaitement avec la perspective de Collins et Kusch parce qu'elles se sont développées dans cette zone frontalière marquée par l'interaction et l'hybridation du mimeomorphe et du polymorphe. Au lieu de vouloir simuler une véritable intelligence, cette « IA faible » (*weak AI*) tente d'utiliser le traitement algorithmique pour retrouver et réorganiser les traces significatives laissées par l'intelligence humaine dans la masse de données, pour traiter les structures qu'elles contiennent et les rendre, par la suite, exploitables dans le contexte de l'accomplissement d'une tâche. Ce projet recoupe donc la remarque de Lucy Suchman :

« Le projet de créer des machines interactives a plus à gagner par la compréhension des différences entre les interactions humaines et le fonctionnement de la machine, que par la simple supposition de leur similitude. » [Suchman 1987, p. 189]

En suivant ce conseil, nous pensons que les capacités des êtres humains et des machines ne sont pas les mêmes et conformément au raisonnement de Douglas Engelbart [1962], nous sommes convaincus que le vecteur central de nos recherches techniques devrait être la combinaison intelligente des différents talents plutôt que l'imitation de l'un par l'autre, peu importe dans quel sens. Les métatechnologies et toutes ces applications qui surgissent avec le Web 2.0 peuvent certainement être qualifiées d'explorations dans cette direction. Avec ces systèmes, la production sémantique reste

¹³³ Une attitude similaire se trouve finalement derrière le test de Turing ; Breton [1995, p. 126] la décrit ainsi : « Peu importe si la machine pense ou non, l'important est qu'elle se comporte comme si elle pensait. » Turing a pourtant sous-estimé la difficulté d'apprendre à la machine à faire semblant.

largement du côté des êtres humains tandis que les algorithmes apportent une plus-value par le croisement d'énormes quantités de marqueurs, de traces et d'appréciations.

Cette insistance sur une différence de catégorie et de capacité entre l'être humain et la machine, n'est-elle pas en contradiction avec notre critique formulée plus haut de cette même dichotomie ? Nous pensons que non, parce que nous avons dit que la symétrie entre l'humain et le non-humain était avant tout un geste méthodologique permettant de rester à l'écart des dichotomies reçues et des préconceptions faciles. Non encore, car malgré les différences de nature, ce serait une erreur de penser la production de sens comme le privilège du seul comportement polymorphe. La différence entre les deux archétypes d'actions n'est pas la même que celle qui sépare « action » et « fonctionnement » et le but du travail de Collins et Kusch [1998, p. 196] consiste explicitement à créer un « terrain de rencontre » (*meeting ground*) pour les actions et les comportements que la philosophie à l'habitude d'opposer. La frontière que ces auteurs établissent concerne avant tout la question du degré d'acculturation nécessaire pour effectuer ou reconnaître une action. Le fait qu'ils conservent une frontière entre l'humain et le non-humain ne veut pas dire qu'il s'agit de la même frontière essentialiste que celle de Searle et Dreyfus. En nous rapprochant des questions posées en SIC, nous ne sommes donc pas d'accord avec le constat de Robert Escarpit à propos des systèmes experts :

« Mais il faut bien se rendre compte qu'en fait ce choix décisionnel de la machine n'est pas vraiment une production d'information : c'est simplement l'application, selon un programme préétabli, de choix de base effectués par des esprits humains à des situations prévues par ces mêmes esprits. »
[Escarpit 1981, p. 197]

Escarpit limite la capacité de produire de l'information, notion qui nous semble chez lui quasiment identique à celle de sens, à l'être humain. La machine ne peut pas produire d'information parce qu'elle n'est que programmée ; elle fonctionne. Mais n'y a-t-il pas une conception mythique ou fabuleuse de la signification dans cette perspective ? Le comportement du programme est certainement prévu, et non devenu comme nous le sommes, mais qui filtre les emails ? Les constructeurs de *SpamBayes* ou le logiciel ? Avec Collins et Kusch nous ne pensons pas que le mimeomorphe soit dépourvu de sens, et une conception de la production de l'information et du sens qui exclut les apports pratiques des algorithmes nous semble victime d'un « facile humanisme » [Simondon 1958, p. 9].

Dans ce chapitre, nous avons essayé de montrer qu'il n'y a pas de réponse définitive à la question de savoir ce qu'on peut déléguer. Les deux membres de l'équation sont dynamiques et leur négociation produit à tout moment des nouvelles possibilités de croisement. La perspective que nous avons tenté de développer à travers la lecture du livre de Collins et Kusch se positionne entre les dogmes habituels. C'est un lieu inconfortable parce que nos concepts de base et surtout notre vocabulaire sont tellement marqués par des dichotomies vieilles qu'il est difficile de prendre son indépendance et de se dispenser des deux réflexes habituels, soit de les accepter en tant que vérité évidente, soit de les rejeter au nom d'un modernisme technique. Nous allons voir que la deuxième question, celle de l'autonomie n'est pas beaucoup plus facile.

1.4.2 La question de l'autonomie

En 1803 le jeune Hegel constate que les machines et l'argent « sont animés d'une vie qui se meut en soi-même - autonome – d'une réalité morte ». Le philosophe allemand reconnaît donc que la frontière entre la vie et la mort n'est pas nécessairement celle qui sépare les êtres biologiques de la matière inanimée. Certains éléments – dont apparemment les machines et l'argent – gagnent par pur *fonctionnement* une identité indépendante de la réalité objective qui les supporte. Cette phrase de Hegel est d'autant plus intéressante que le terme « autonomie » (*Autonomie*) en allemand est souvent défini comme « liberté de volonté » (*Willensfreiheit*) ; l'autonomie ainsi attribuée aux machines décrit donc non seulement une indépendance en termes d'énergie et de direction, mais quelque chose qui va au-delà de la simple absence de la dépendance de l'impulsion et de l'orientation apportées par un être humain. Mais comment penser ce « quelque chose de plus » ? Le problème est en même temps ambivalent et ambigu. Ambigu parce que le sens du terme « autonomie » diffère selon la discipline et ambivalent parce que les jugements portés en son nom divergent profondément.

Les autonomies

En philosophie, la question de l'autonomie est directement liée à celle de la liberté. La faculté de se « gouverner par ses propres lois » (la traduction la plus exacte du terme grecque *autonomia*) se réalise seulement quand ces lois ont été *librement* choisies. L'extension et les limites de la liberté des êtres humains occupent l'esprit des philosophes depuis longtemps et au moment de la rédaction de ces pages il n'y avait toujours pas de consensus en vue. Les deux grands courants de la philosophie française de la deuxième moitié du vingtième siècle se sont opposés parmi maints autres sujets aussi sur ce terrain : l'existentialisme déclarant l'être humain condamné à la liberté et le structuralisme l'enfermant dans un jeu de forces qui le dépassent.

Dans les disciplines techniques qui tentent de construire pratiquement des machines dotées d'autonomie, le mot n'a pas le même sens qu'en philosophie. Il n'est pas lié au principe (métaphysique ou ontologique) de la « liberté de volonté », mais à quelque chose de bien plus pragmatique, que Dietrich Dörner [2003, p. 112] appelle la « liberté de choisir ». En ce sens, un système autonome est défini comme une machine qui choisit parmi différents comportements possibles en prenant en compte les conséquences positives ou négatives de ces comportements. Pour procéder ainsi, il lui faut la capacité de se représenter son environnement ainsi que la compétence de « raisonner » sur les effets des changements qu'il y provoque. Sur le plan technique il y a aujourd'hui une série d'approches pour modéliser un tel système, et quasiment tous les agents « intelligents » contiennent un tel moteur.

Les sciences humaines contestent souvent que les modèles employés par les techniciens produisent une « véritable » autonomie. Selon elles, il ne s'agit de rien d'autre que des relations causales interférées par un moteur logique. Il ne pourrait pas y avoir d'*action* mécanique, mais seulement de la

réaction. Comme dans le cadre du débat autour de l'intelligence, il existe, à ce sujet, une opposition plutôt dogmatique et essentialiste entre les deux cultures. Et encore une fois, nous souhaitons avancer une position intermédiaire qui prend en compte la vie des tous les jours des usagers dans des contextes pratiques.

Le philosophe Daniel Dennett [1987] aborde ce problème à partir de notre attitude quotidienne envers d'autres êtres humains que nous traitons couramment en tant qu'agents intentionnels et non comme des automates. Nous empruntons aussi cette « posture intentionnelle » (*intentional stance*) envers des objets dont *nous ne connaissons ou ne comprenons guère* la composition interne. Afin de pouvoir prédire leurs comportements, malgré notre ignorance, nous leur attribuons des croyances (*beliefs*), des désirs (*desires*) et des buts (*goals*).¹³⁴ Avec une telle *grille de lecture* nous pouvons habiller l'entité en question d'un manteau de sens qui nous permet d'interagir avec elle sans connaître les principes techniques réels qui la font fonctionner. Dans cette perspective, il est secondaire que l'autonomie affichée soit « réelle » ou non parce que l'utilisateur ne juge qu'à partir de ce qu'il voit. C'est une position assez similaire à celle que propose Alan Turing dans son fameux *jeu d'imitation* : il n'est guère important que l'intelligence affichée par la machine soit « vraie ». Il suffit qu'elle réussisse à convaincre un observateur de son humanité.

Ce qui nous intéresse dans le contexte de ce travail est de concevoir les NTIC à partir de la relation usager-objet-créateur et non seulement en partant de l'objet. L'autonomie nous concerne donc en tant qu'autonomie *de* l'utilisateur et autonomie *perçue par* l'utilisateur. Le premier aspect porte sur la question pratique de savoir quel niveau d'intervention de l'utilisateur sera nécessaire pour accomplir la tâche en question. Le deuxième aspect concerne le regard que porte l'utilisateur sur l'objet technique. Mais il y a quand même une troisième dimension importante : l'autonomie de l'objet par rapport à son créateur.

Autonomie de l'utilisateur

L'autonomie de l'objet par rapport à l'utilisateur qu'il est censé servir est la plus essentielle dans la question de la délégation. Un degré d'indépendance élevé se traduit par un faible taux de contrôle. Les métatechnologies que nous avons présentées sont toutes dotées d'une certaine autonomie – ce qui fait même partie de leur définition. Bien que nous ayons rapproché certains de nos exemples, surtout le moteur de recherche de Google, de la logique de l'outil, il faut reconnaître que même dans ce cas la distance entre l'utilisateur et l'objet reste assez grande. La participation de l'internaute est certainement requise et le fait qu'il formule la requête lui garantit certainement de conserver une influence et un

¹³⁴ Il est intéressant de noter que le modèle le plus utilisé pour construire du comportement autonome, le *belief-desire-intention model*, s'inspire en partie du travail de Dennett et d'autres philosophes de tradition analytique comme Michael Bratman.

contrôle sur le résultat – à la différence de *Google News* où aucun apport humain n'est nécessaire. Mais même dans le cas du moteur de recherche, la plupart des paramètres, l'indexation, l'algorithme de *ranking*, etc. sont hors de portée des internautes. L'autonomie s'accompagne donc d'un double mouvement : elle nous fait gagner en puissance mais nous fait perdre en contrôle. En ce sens, le transfert d'une tâche implique un transfert secondaire qui est peut-être involontaire mais néanmoins inséparable de son apport.

La complexification des techniques de communication et d'information, c'est-à-dire leur étroite imbrication avec des algorithmes puissants, accentue la question de l'autonomie au sens où les capacités d'agir des artefacts grandissent, que ses facultés de « se débrouiller tout seul » progressent. Cependant, ce n'est pas seulement le gain de capacités de cognition et de raisonnement qui rendent certaines interactions entre l'être humain et la machine obsolète, mais aussi le fait que le fonctionnement interne est devenu trop complexe pour qu'un usager puisse encore le paramétrer facilement.

L'exercice d'un quelconque contrôle sur l'objet technique ne se fait donc plus par *manipulation directe* de sa structure interne, mais seulement de façon médiatisée, par le biais de la communication. *SpamBayes* en est le meilleur exemple : ici, il n'y a quasiment plus d'interaction de type « paramétrage » entre le programme et son usager. Ce dernier se contente de marquer un courriel non-souhaité comme spam et le logiciel se charge du reste.

Suite à ces constats, nous pouvons définir le degré d'autonomie d'un logiciel par rapport à l'utilisateur comme équivalent à sa capacité d'agir selon ses propres règles, indépendamment de l'intervention humaine. Bien évidemment, cette question ne peut pas être validée de manière globale mais uniquement en se référant à des exemples précis : l'autonomie qui peut être affichée n'est pas seulement une question d'avancées techniques, mais dépend en large partie de la nature de la tâche ou responsabilité que le logiciel est censé accomplir. De ce fait, la question de l'autonomie pratique recouvre celle, plus générale, des capacités et performances réelles et potentielles des machines.

Autonomie perçue par l'utilisateur

La perception par l'utilisateur de l'autonomie d'un programme qu'il emploie est bien sûr liée au point précédent, c'est-à-dire à la capacité d'une application d'accomplir des tâches sans ou avec peu d'intervention de la part de l'utilisateur. Un logiciel qui ne demande pas d'inputs ou paramètres pour fonctionner est probablement perçu comme plus autonome qu'un programme qui propose beaucoup de possibilités de configuration. Le problème ne se réduit pourtant pas à la question de l'influence véritable de l'utilisateur sur le fonctionnement de l'objet. Il s'agit surtout d'un problème de *lecture*. Chaque logiciel possède une certaine apparence. Habituellement, il est doté d'une interface, son fonctionnement laisse des traces dans le système d'exploitation et produit des effets sur son domaine d'application. Même un programme qui souhaite rester invisible (chevaux de Troie, *spyware*, etc.) ne

l'est pas pour quelqu'un qui sait chercher. Et toute la question est là : la *visibilité* ne garantit pas la *lisibilité*. Cette dernière ne peut pas être définie seulement en regardant l'objet en question, parce que l'acte de lecture est toujours la rencontre entre un texte (un objet) et un lecteur [cf. de Certeau 1990]. Il existe des marqueurs « objectifs » appartenant à l'objet, certes, mais l'interprétation de ces marqueurs dépend de l'alphabétisation, du degré de cyberculture [Baltz 2005] de la personne qui les perçoit. L'autonomie perçue (mais aussi l'intelligence perçue) est donc avant tout un *effet* de lecture.

Une façon très simple de produire l'effet d'autonomie et d'intelligence consiste à choisir une forme anthropomorphe ou zoomorphe pour représenter un programme. Même des applications dont les capacités s'avèrent très faibles peuvent ainsi apparaître comme des « agents intelligents ». Nous reviendrons de nouveau à ce sujet – fort contesté – de la représentation de la délégation. Mais il est important de noter que la perception de l'autonomie par l'utilisateur nous amène dans un espace très complexe où le pur fonctionnement technique n'est plus qu'un paramètre parmi d'autres. L'attribution d'un comportement autonome ou intelligent, la posture intentionnelle de Dennett, n'est pas déterminée par l'objet. Nous « anthropomorphisons » couramment des artefacts qui ne le méritent par aucun élément « objectif ». Voilà la *tache aveugle* de beaucoup de techniciens : la production du sens, la fixation de significations ne sont l'affaire unique des constructeurs, mais se forment dans la rencontre entre un objet et un être humain.

Autonomie du créateur

Un troisième niveau d'autonomie doit être distingué, que nous pourrions appeler « autonomie par complexité ». Cette problématique, au bout du compte, revient à la question du contrôle par ceux que Robert Escarpit appelle des « tireur de ficelles » :

« Il [le programme] n'a pas plus d'autonomie qu'une marionnette et, comme pour la marionnette, le tireur de ficelles reste caché. C'est ce manipulateur qui, à travers l'analyse et la programmation, pose les problèmes, raisonne, tire les conclusions et décide. » [Escarpit 1980, p. 31]

Escarpit met le doigt sur une chose très importante qui concerne pleinement le regard que nous portons sur les NTIC : l'objet technique résulte toujours d'un processus de production qui n'est ni magique, ni neutre, mais pleinement rempli d'enjeux sociaux, de jugements de valeurs et de motivations très éloignés de la technique. Mais il ne faut pas non plus surestimer la prouesse des individus qui interviennent dans ce processus ; l'idée du « manipulateur » ne devrait pas réinstaller le créateur dans la position du génie qui maîtrise à tout moment tous les fils qu'il tient dans sa main. Il faut trouver un juste milieu et la notion d'une autonomie par complexité est l'un des éléments qui met en question l'omnipotence du « tireur de ficelles ».

Nous avons déjà rencontré cette question plus haut, lors de l'anecdote sur la prétendue préférence politique de *Google News*. L'incapacité des créateurs à anticiper un tel effet politique – l'apparente préférence donnée à un candidat sur l'autre pendant la campagne électorale – devrait nous alerter sur

quelque chose de plus général : les résultats des algorithmes avancés (non-triviaux), souvent basés sur des modèles probabilistes, génétiques ou connexionnistes, ne sont que difficilement prévisibles et il y a toujours des surprises. Cette imprévisibilité concerne d'abord les usagers, mais même les créateurs sont loin d'exercer un contrôle complet sur le comportement d'un tel programme, surtout quand il traite un terrain informationnel que sa taille met hors de la portée de la perception humaine. Au moment où s'y ajoute l'apprentissage automatique, le programme devient autonome au sens où il n'est plus évident de savoir où il va conduire, les données qu'il traite et les interactions avec les usagers définissant en grande partie ses états futurs.¹³⁵ Cette autonomie par complexité ne devrait pas nous faire croire qu'il s'agit de l'émergence d'une véritable « intelligence », qui se détacherait magiquement d'un fond mécanique, au seul motif que nous ne comprenons plus les détails du traitement. Mais il faut accepter qu'il y ait des limites à la quantité et à la complexité des données que le cerveau humain peut digérer. Le traitement d'information par la machine peut produire des résultats inattendus parce que l'anticipation nécessite l'exécution¹³⁶ et que nous ne pouvons pas le faire sans aide mécanique.

Les ficelles tirées par le programmeur peuvent s'avérer plus poreuses qu'on pourrait le penser : non seulement le comportement des usagers, mais aussi progressivement le comportement de la machine peuvent échapper aux créateurs d'objets techniques. Le programme qui est distribué chez les usagers n'est plus dans la main des designers et la prévision des comportements atteint vite des limites.

Discussion

Une bonne partie des pistes que nous avons essayé d'ouvrir jusqu'ici nous ont conduit sur le terrain de « l'hypercomplexité propre au phénomène anthropologique » dont parle Edgar Morin [2005, p. 52]. Au moment où nous intéressons au rapport que la technique entretient avec l'être humain en tant que réalité dense et non pas seulement en tant que modèle formel, les catégories deviennent glissantes et difficiles à saisir parce qu'un terrain commun, c'est-à-dire un milieu conceptuel qui ne soit pas le résultat d'un geste impérialiste d'un domaine sur l'autre, ne se construit que lentement et à travers maints détours. Cette condition d'insécurité ontologique est l'une des conséquences du processus d'hybridation.

¹³⁵ Sans vouloir entrer dans les détails, indiquons que tous les problèmes traités par la théorie de complexité et notamment ceux de NP-complétude interviennent ici.

¹³⁶ Voilà le fameux problème de l'arrêt (*halt*) de Turing.