



# Jean-Pierre DUPUY

Jean-Pierre Dupuy, ancien élève de l'École Polytechnique, Paris, Ingénieur général des Mines, Directeur de Recherche au CNRS, Professeur de philosophie sociale et politique à l'École Polytechnique et à l'Université Stanford (Californie), a fondé et dirigé jusqu'en 1999 le Centre de Recherche en Epistémologie Appliquée (le CREA).

Il vient d'y créer le Groupe de Recherche et d'Intervention sur la Science et l'Éthique (GRISE).

Il a publié : *La Panique (Les empêcheurs de penser en rond, 1991 ; nouvelle édition 2003) ; Le Sacrifice et l'Envie – le libéralisme aux prises avec la justice sociale (Calmann-Lévy, 1992 ; Hachette, 1997) ; Introduction aux sciences sociales (Ellipses, 1992) ; Aux origines des sciences cognitives (La Découverte, 1994 ; 1999) ; Éthique et philosophie de l'action (Ellipses, 1999) ; Les Limites de la rationalité, tome 1 : Rationalité, éthique et cognition (avec Pierre Livet ; la Découverte, 1997) ; The Mechanization of the Mind (Princeton University Press, 2000) ; Pour un catastrophisme éclairé (Seuil, 2002).*

Il a fait paraître en septembre 2002 un livre sur les attentats terroristes du 11 septembre 2001, *Avions-nous oublié le mal ? Penser la politique après le 11 septembre* (Bayard).

Jean Pierre Dupuy siège au Conseil Général des Mines et il est membre du Comité d'Éthique et de Précaution de l'INRA.

# **Les défis éthiques des nanotechnologies**



## INTRODUCTION

Il y a quelques années maintenant que je travaille comme philosophe sur l'éthique des nanotechnologies – la « nanoéthique » comme on la nomme maintenant – ou, plus précisément, sur l'éthique de la dénommée « convergence NBIC », qui est la convergence entre Nano, Bio, Info-technologies et Sciences Cognitives. J'aimerais tout d'abord défendre le point de vue que la convergence NBIC et les nanotechnologies en particulier soulèvent des questions éthiques originales. Cela a été fortement nié. Par exemple, Philip Ball, écrivain scientifique de *Nature*, écrit dans un essai intitulé « 2003 : nanotechnologie sur la ligne de feu »<sup>1</sup> :

En mars [2003], l'Institution Royale de Londres a accueilli un séminaire d'un jour sur les nanotechnologies intitulé « Atome par atome », que j'ai personnellement trouvé utile car il a permis d'entendre un large spectre d'opinions sur ce qui est maintenant connu sous le nom de nanoéthique. [...] Tout d'abord, une crainte est apparue fondée sur l'idée que ce qui est qualitativement nouveau dans les nanotechnologies est qu'elles permettent, pour la première fois, la manipulation de la matière à l'échelle atomique. Je suppose que cette vue est communément partagée et, si c'est le cas, cela doit nous pousser à nous demander : comment est-il possible que nous vivions dans une société où l'on ne voie pas que c'est précisément ce que la

chimie fait de façon rationnelle et informée depuis deux siècles et même plus? Comment avons-nous pu laisser s'installer une telle ignorance? Il devient de plus en plus clair que le débat à propos de l'étendue des possibilités ultimes offertes par les nanotechnologies se ramène à des questions sur les fondements mêmes de la chimie [...] La vacuité des connaissances dans laquelle la majeure partie du débat public sur les nanotechnologies prend place n'existe que parce que le public ne connaît presque rien à la chimie : ce qu'elle est, ce qu'elle fait et ce qu'elle peut faire.

Se penchant sur la nanoéthique, Ball poursuit :

Les questions relatives à la sécurité, l'équité, les implications militaires et la transparence des nanotechnologies sont identiques à celles que soulèvent d'autres domaines de la science et de la technologie. Ce serait une erreur grave, et peut-être dangereuse, que les nanotechnologies en viennent à être considérées comme une discipline qui pose des problèmes éthiques inédits. De ce point de vue, je pense qu'elles diffèrent fondamentalement de certains pans de la recherche en biotechnologie qui touchent à des questions morales entièrement nouvelles. Et cependant, c'est peut-être la première fois qu'une science, une science appliquée ou une technologie, comme on voudra, se développe dans un

<sup>1</sup> Nanotechweb.org, 23 décembre 2003.  
<http://www.nanotechweb.org/articles/society/2/12/1/1>

climat social sensibilisé par avance aux besoins d'un débat éthique.

[...] Plus encore, la vérité pragmatique est que, si les nanotechnologies ne reconnaissent pas qu'elles comportent une dimension éthique, elles y seront forcées quoi qu'il arrive. Ceux qui travaillent dans le champ savent que les nanotechnologies ne constituent en aucune manière une discipline, que leurs visées ne peuvent être réunies en un tout cohérent et qu'elles ne représentent pour aucun secteur industriel un objectif défini. Mais même les agences de financement en parlent comme si tel n'était pas le cas. Dans l'esprit du public, le simple fait que des opérations telles que l'Initiative nationale américaine pour les nanotechnologies existent suggère à coup sûr que les nanotechnologies ont une certaine unité, et c'est pourquoi il apparaît que les gens voudront être assurés que leurs aspects éthiques sont bien pris en considération.

Je pense que Philip Ball se trompe doublement. Je crois qu'il y a vraiment une unité derrière la convergence des NBIC ; mais cette unité se trouve au niveau du *programme métaphysique de recherche* sur lequel s'appuie cette convergence. Je crois aussi que les enjeux éthiques que ce programme soulève sont largement inédits et qu'ils trouvent leur source dans les *idées fortes* qui gouvernent le champ.

Je viens d'utiliser l'expression « programme métaphysique de recherche » : il s'agit d'une référence à la philosophie des sciences de Karl Popper. La philosophie positiviste qui sous-tend la majeure partie de la science moderne (et beaucoup de la philosophie contemporaine) considère que la « métaphysique » est une quête dénuée de sens pour trouver des réponses à des questions qui n'ont pas de réponses, mais Popper, à la suite d'Emile Meyerson<sup>2</sup>, a montré qu'il n'y a pas de programme de recherche scientifique (ou, pour notre sujet, technologique) qui ne repose sur un ensemble de présuppositions générales au sujet de la structure du monde. Certes, ces vues métaphysiques ne peuvent être testées empiriquement et elles ne peuvent faire l'objet d'une « falsification ». Cependant, cela n'implique pas qu'elles ne présentent pas d'intérêt substantiel et qu'elles ne jouent pas un rôle fondamental dans l'avancement de la science. Ceux qui dénie la métaphysique la rendent simplement invisible, et il est très vraisemblable que leur métaphysique cachée soit mauvaise ou inconsistante. Au grand étonnement de ceux qui, à tort, le considéraient comme un positiviste, Karl Popper répondait que la tâche du philosophe ou de l'historien des sciences était double : premièrement, exhumer et rendre visibles les idées métaphysiques qui reposent sous les programmes scientifiques afin de les amener à la critique ; deuxièmement, procéder à un examen critique de ces théories métaphysiques, différent certes de celui qui

<sup>2</sup> « L'homme fait de la métaphysique comme il respire, sans le vouloir et surtout sans s'en douter la plupart du temps. » E.Meyerson, *De l'explication dans les sciences* (Paris, 1927).

fonde la critique des théories scientifiques, puisque ici aucun test empirique n'est possible, mais néanmoins rationnel.

La question que je me suis posée ensuite est donc la suivante : quel est le programme métaphysique de recherche qui sous-tend ladite convergence NBIC ? L'une de ses caractéristiques majeures est que les technologies convergentes prétendent prendre le relais de la nature et de la vie et devenir les ingénieurs de l'évolution. L'évolution jusqu'à maintenant a fondamentalement consisté en un simple « bricolage ». Elle peut se bloquer elle-même dans des chemins indésirables ou des impasses. C'est pourquoi l'homme peut être tenté de prendre la relève et de devenir le *designer* des processus biologiques et naturels. *L'homme peut participer à la fabrication de la vie.*

Pouvons-nous évaluer de façon normative une telle entreprise ? Je ne crois pas que les réponses éthiques préexistent aux questions qui les sollicitent. La tâche du philosophe n'est certainement pas *d'appliquer* des doctrines éthiques prêtes à l'emploi à des problèmes nouveaux. Les normes et les règles émergent des problèmes mêmes qu'elles sont censées réguler. Cette vue *en boucle (bootstrapping)* de ce qu'est un jugement normatif est inévitable quand on aborde l'évaluation de sauts technologiques qu'on ne fait qu'anticiper.

Je voudrais commencer par quelques remarques méthodologiques.

## L'ÉTHIQUE N'EST PAS UN CALCUL COÛTS-AVANTAGES

Une première erreur à dénoncer est celle qui consiste à confondre éthique et prudence, et à comprendre « prudence » sur le mode de la gestion rationnelle du risque. 90% des rapports, articles ou livres que j'ai pu consulter sur le sujet commettent cette faute. Car c'en est une, aussi grave que celle que commettrait un physicien qui ne ferait pas la différence entre masse et poids. C'est une erreur sérieuse de traiter des questions éthiques en termes de bilan entre des coûts et des avantages, c'est-à-dire de réduire l'éthique à une sorte de calcul économique élargi. Dans un plateau de la balance, on met les bénéfices que l'on attend du progrès technologique et économique et dans l'autre, les coûts. L'incertitude affecte plus le second plateau que le premier, et c'est évidemment en termes de risques qu'on l'appréhende. Or le concept de risque s'analyse en trois éléments: a) il existe une éventualité de dommage, affectée normativement d'un signe moins; b) il est possible d'assigner un degré de vraisemblance à l'occurrence de ce dommage, sous la forme d'une probabilité par exemple; c) il est loisible de prendre pour étalon de l'appréciation du dommage un système d'évaluations individuelles et collectives, par exemple les « préférences », les fonctions d'utilité ou de satisfaction de la population d'individus potentiellement concernés par le dommage.

Il devrait être clair que les questions éthiques soulevées par les technologies avancées ne satisfont aucune des trois conditions que je viens d'énoncer. Lorsque le rapport de la National Science Foundation de juin 2002 intitulé *Converging Technologies for Improving Human Performance* énonce que la convergence NBIC va entraîner un «changement de civilisation<sup>3</sup>», bien malin serait celui qui s'aventurerait à mettre un signe, plus ou moins, devant cette éventualité, qui se prononcerait sur son degré de vraisemblance ou qui en évaluerait les conséquences en additionnant les différentiels d'«utilités» sur toute la population.

Il faut bien comprendre que l'obstacle méthodologique que je souligne n'est pas réductible à une forme d'incertitude sur les données. Il n'est pas épistémique, il est ontologique. Je prends l'exemple du très intéressant «Report on the Social and Ethical Issues of Genetic Engineering with Human Beings», de novembre 1982, plus connu sous le nom de «Splicing Life»<sup>4</sup>. Dans une section du chapitre "Social and Ethical Issues" intitulée "Concerns with Consequences", on lit ceci :

La tendance courante est de penser une personne comme un individu ayant un certain caractère et une personnalité qui, tout au long des étapes normales de son développement physique, social et psychologique, reste à peu près fixe dans un certain domaine. Mais ce con-

cept – et l'impression de prédictibilité et de stabilité qu'il confère aux relations interpersonnelles – pourrait rapidement devenir démodé si les gens se mettent à recourir à des techniques de greffe génétique pour opérer des changements fondamentaux sur eux-mêmes tout au long de leur vie. On peut d'ores et déjà subir des transformations profondes au moyen de la psychochirurgie, des techniques de modification du comportement ou de l'usage thérapeutique de drogues psychoactives. Mais il n'est pas à exclure que le génie génétique apporte des moyens plus rapides, plus sélectifs et plus faciles d'usage. Ici encore, l'incertitude sur la façon dont les concepts de base d'une personne peuvent évoluer au cours du temps s'accompagne d'une incertitude normative et éthique, parce que les concepts en question sont intimement liés à des valeurs et à des postulats éthiques. Il est peu probable que les gens puissent avoir une vision claire de ce que serait leur attitude éthique si leur notion de ce qui constitue l'identité personnelle ou bien de ce qui constitue les étapes normales du développement au cours d'une vie venait à varier profondément.

Ce texte est remarquable par son aveuglement, en ce que, tout à la fois, il désigne une difficulté, qui est en vérité une impossibilité, comme Popper l'a montré : celle d'anticiper ce que seront nos principes éthiques et nos valeurs à l'avenir ; et qu'il

<sup>3</sup> NSF Report *Converging Technologies for Human Performance*, Mihail C. Roco & William S. Bainbridge (eds.), juin 2002, p. 21.

<sup>4</sup> President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research.

continue néanmoins à raisonner en termes conséquentialistes, le tour de passe-passe consistant à parler d'incertitude alors qu'il s'agit d'une indétermination radicale. On imagine très bien un champion du Bayésianisme se débrouiller pour placer des probabilités subjectives dans l'espace des systèmes de valeurs ! Ce n'est certes pas la première fois que cette confusion est commise, puisque la fameuse *Unbestimmtheitsrelation* de Werner Heisenberg est traduite par „Principe d'incertitude“ alors qu'il s'agit d'un principe d'indétermination.

Le rabattement de l'ontologique sur l'épistémique est une faute d'autant plus grave que les transformations que nos principes et valeurs éthiques connaîtront seront en partie causées par les choix technologiques mêmes que nous allons faire. Le rapport de la NSF *Converging Technologies* en est bien conscient qui écrit, à propos du «changement de civilisation» entraîné par la convergence des technologies: «Il est possible que des principes éthiques entièrement inédits vont prévaloir dans des secteurs où le progrès technologique sera radical : ainsi de l'acceptation des implants cervicaux, le rôle des robots dans la société et l'ambivalence de la mort devant les avancées en matière de clonage..<sup>5</sup>”

Si l'éthique pouvait être ramenée à un calcul moral d'avantages et de coûts, sa tâche pour les problèmes qui nous occupent

serait proprement désespérée, puisque nous ne saurions même pas dire dans quel plateau de la balance il faudrait placer tel ou tel aspect d'une évolution envisagée. Si l'éthique était un calcul, même si ce calcul se révélait impossible à effectuer en pratique, tout problème éthique serait réputé avoir une solution, même si nous étions incapables de la déterminer. Mais la situation morale de l'homme est d'une autre nature. Il peut arriver que la conquête du savoir et l'activité créatrice des hommes se révèle une arme à double tranchant, «mettant en danger la poursuite même des processus à laquelle elle est pourtant indispensable.<sup>6</sup>» Nous en verrons des exemples. Ce n'est donc pas que nous ignorions si l'usage de cette arme est une bonne ou une mauvaise chose: c'est que cet usage est bon et mauvais à la fois.

## L'ÉTHIQUE DES TECHNOLOGIES, PAS DES TECHNIQUES

Une autre erreur consiste à faire porter l'évaluation éthique sur la technique elle-même. Sur quoi donc devrait-on la faire porter ? Sur la technologie ! Le français moderne ne fait semble-t-il plus la distinction entre ces deux termes, marchant sur les traces de l'anglais, et c'est une perte. La technologie, c'est la technique enchâssée dans le discours [logos] qui l'accompagne, tout à la fois la porte et est porté par elle. Je prends une illustration dans l'œuvre majeure

<sup>5</sup> NSF Report, op. cit., p. 22; je souligne. Le rapport ajoute, cependant: "l'identité et la dignité humaine doivent être préservées."

<sup>6</sup> Henri Atlan, *Les Etincelles de hasard*. Tome 1: *Connaissance spermatique*, Paris, Seuil, 1999, p. 45.

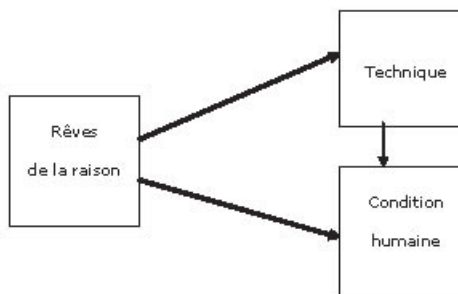


re de Hannah Arendt, *Human Condition*. Ce livre, publié pour la première fois en 1958<sup>7</sup>, commence par une méditation sur «cet événement, que rien, pas même la fission de l'atome, ne saurait éclipser» qu'a constitué la mise sur orbite d'un Spoutnik, le 4 octobre de l'année précédente. Arendt écrit :

La réaction immédiate, telle qu'elle s'exprima sur le champ, ce fut le soulagement de voir accompli le premier 'pas vers l'évasion des hommes hors de la prison terrestre.' [...] Ces opinions sont devenues des lieux communs. Elles prouvent que les gens ne sont nullement en retard sur les découvertes de la science et sur les progrès techniques et qu'au contraire *ils les ont devancés de plusieurs dizaines d'années. En ce cas comme dans d'autres, la science a réalisé et confirmé ce que les hommes avaient anticipé dans des songes qui n'étaient ni creux ni absurdes.* La seule nouveauté, c'est que l'un des plus respectables journaux américains aient enfin proclamé en première page ce qui jusqu'alors était enfoui dans la littérature fort peu respectable de la science-fiction (à laquelle malheureusement personne n'a encore accordé l'attention qu'elle mérite comme véhicule des sentiments et aspirations de masse). La banalité de la phrase ne doit pas nous faire oublier qu'elle était, en fait, extraordinaire; car si les chrétiens ont parlé de la Terre comme d'une vallée de larmes et si les philosophes n'ont vu dans le corps qu'une vile prison de l'es-

prit ou de l'âme, personne dans l'histoire du genre humain n'a jamais considéré la Terre comme la prison du corps, ni montré tant d'empressement à s'en aller, littéralement, dans la Lune. [...] La Terre est la quintessence même de la condition humaine [...]»<sup>8</sup>

Je tiens ce texte pour une contribution essentielle aux problèmes qui nous occupent. Il nous dit que les hommes rêvent la science avant de la faire et que ces rêves, qui peuvent prendre la forme de la science fiction, ont un effet causal sur le monde: ils s'incarnent dans la technique et celle-ci transforme la condition humaine. L'objet de l'évaluation éthique doit donc être, non pas la technique seule, mais cette structure de cause commune:



Celui qui croit que c'est la technique seule qui a un effet sur la condition humaine se doit de faire le départ entre ce qui est techniquement réalisable et ce qui ne l'est pas. On observe effectivement que

<sup>7</sup> Chicago, The University of Chicago Press, 1958. J'utilise la traduction française de Georges Fradier, *Condition de l'homme moderne*, Calmann-Lévy, 1961.

<sup>8</sup> Ibid., p. 7-8.

les travaux déjà existants en nanoéthique prennent un soin extrême à distinguer ce qu'ils tiennent pour la science sérieuse de ce que tous appellent la «science fiction». Mais le domaine de celle-ci varie beaucoup d'un rapport à l'autre. Si tous s'accordent à rejeter les élucubrations d'un Michael Crichton dans le royaume de la littérature de bazar, on se bat encore sur le point de savoir si les nanorobots autorépliquants imaginés par Eric Drexler, l'«inventeur» des nanotechnologies, sont ou non condamnés à faire à jamais partie de cette science fiction dédaignée. Le principal intéressé lui-même semble avoir changé complètement d'opinion et range désormais dans la catégorie du «non-sérieux» ce qui lui a assuré la célébrité. Le récent rapport de la Royal Society britannique, quant à lui, se permet de traiter par le mépris le rapport de la NSF américaine, avec un perfide: «On nous pardonnera de dédaigner la plupart des contributions, car elles portent moins sur la science et la technologie sérieuses que sur la science fiction.<sup>9</sup>»

Le changement de perspective que je propose a pour première implication que « anything goes », comme aurait dit Feyerabend, en ce qui concerne ce qu'il convient de mettre dans la boîte «Rêves de la raison» : le non-sérieux n'est pas moins important que le sérieux lorsqu'il s'agit de nourrir l'imaginaire de la science. La méta-physique qui sous-tend la convergence NBIC

se trouve donc dans la boîte, à côté de l'idéologie de battage nourrie par les livres d'un Ray Kurzweil<sup>10</sup>, d'un Eric Drexler<sup>11</sup> ou d'un Damien Broderick<sup>12</sup>. On y trouvera aussi la pratique de la langue, les arts, la littérature populaire et bien d'autres choses.

Comme Hannah Arendt qui réagissait à l'expression d'un journaliste comparant la terre à une prison, le chercheur en nanoéthique doit rester attentif aux torsions que l'on fait subir à la langue et qui, tels des lapsus ou des symptômes, révèlent les couches les plus profondes de l'imaginaire scientifique ou technologique. Je ne résiste pas à la tentation de donner un exemple. En novembre 2003, des scientifiques israéliens ont construit des transistors à partir de nanotubes de carbone en utilisant l'ADN comme gabarit. Un scientifique du Technion-Israël a dit : « ce que nous avons fait a consisté à amener la biologie à *auto-assembler un mécanisme électronique* dans un tube à essai [...]. L'ADN est utilisé comme un échafaudage ou un gabarit qui détermine où les nanotubes de carbone vont se trouver. C'est la beauté de l'utilisation de la biologie.»<sup>13</sup> Cette utilisation transitive d'un verbe réflexif trahit à elle seule, mieux qu'un long discours, l'ambition des nanobiotechnologies *d'arraisonner* [*le Gestell* heidggérien] les propriétés auto-organisatrices du vivant pour les mettre au service des fins humaines.

<sup>9</sup> The Royal Society, *Nanosciences and nanotechnologies: opportunities and uncertainties*, RS Policy document 19/04, juillet 2004, p. 55.

<sup>10</sup> Ray Kurzweil, *The Age of Spiritual Machines*, Texere Publishing, 2001.

<sup>11</sup> Eric Drexler, *Engines of Creation*, Anchor Books, 1986.

<sup>12</sup> Damien Broderick, *The Spike*, Forge, New York, 2001.

<sup>13</sup> Kenneth Chang, "Smaller Computer Chips Built Using DNA as Template", *New York Times*, November 21, 2003: <http://www.nytimes.com/2003/11/21/science/21DNA.html?ex=1075525200&en=67948bd27029a142&ei=5070>.

Pourquoi la démarche que je propose s'applique-t-elle particulièrement bien aux nanotechnologies, et, plus généralement, à la convergence NBIC ? Parce que, pour l'essentiel, ces technologies n'existent pas encore. Mais le «rêve» qui les porte – avec ses dimensions métaphysiques, idéologiques, populaires, etc. - , lui, est bien présent, et sans doute depuis bien longtemps. Je vais d'ailleurs le montrer, en citant des œuvres philosophiques datant de plusieurs décennies, voire de plusieurs siècles, et dont on peut dire, regardant en arrière, qu'elles avaient prévu beaucoup de choses et, peut-être, contribué à les mettre en place.

J'utilise l'expression «rêves de la raison» à dessein. Je fais référence à la gravure terrifiante de Goya qui porte le titre «El sueño de la razón produce monstruos»<sup>14</sup>. Titre ambigu au possible, puisque le mot «sueño», en espagnol, signifie indifféremment «sommeil» ou «rêve». En français et en anglais, on traduit souvent par «Le sommeil de la raison produit des monstres», et l'on comprend: «Lorsque la raison est en sommeil, c'est-à-dire mise entre parenthèses, alors l'imagination produit des monstres.» Mais un autre sens n'est pas moins possible: «Les rêves de la raison engendrent des monstres.» C'est la raison elle-même, et non son absence, qui a cette capacité de faire advenir par ses rêves des choses monstrueuses. J'aime cette ambivalence fichée au cœur même des rapports entre la science et l'imaginaire.

Je voudrais maintenant, non pas dresser un programme de recherches en nanoéthique qui éviterait les erreurs que je viens de dénoncer, mais simplement esquisser sur quelques dimensions la direction que pourrait prendre un tel programme.

### DÉCLENCHER LA COMPLEXITÉ

On affirme souvent que le point de départ des nanotechnologies est la conférence désormais classique donnée par Feynman en 1959<sup>15</sup>. J'exprime mon désaccord. Il me semble plus juste de dire que l'origine de ce nouveau champ doit être recherchée dans une autre conférence classique, celle que John von Neumann donna en 1948 au cours d'un colloque organisé par la Fondation Hixon au California Institute of Technology (CalTech), et qui apparaît rétrospectivement comme l'un des moments fondateurs des sciences cognitives. Von Neumann y présenta sa théorie des automates auto-reproducteurs et, à cette occasion, émit une conjecture portant sur la notion de complexité<sup>16</sup>.

On peut concevoir, affirma von Neumann dans ce contexte dominé par les cybernéticiens, que des automates complexes soient non seulement capables de s'auto-reproduire, mais aussi de produire des automates plus complexes qu'eux. Von Neumann s'opposait ainsi au projet de la cybernétique,

<sup>14</sup> <http://www.bne.es/Goya/c75.html>

<sup>15</sup> « There is Plenty of Room At the Bottom ».

<sup>16</sup> Voir Jean-Pierre Dupuy, *The Mechanization of the Mind*, Princeton University Press, 2000.

lequel consistait à réaliser (au moins sur le papier), pour chaque fonction classiquement attribuée à l'esprit humain, une machine (de Turing) capable de la reproduire ou de la simuler. Mais quid, objectait von Neumann, si la manière la plus simple de caractériser la fonction est de présenter la machine elle-même ? Il serait dans ces conditions dépourvu de sens de découvrir que tel comportement peut être incarné dans une certaine structure puisqu'il ne serait pas possible de définir le comportement autrement qu'en décrivant la structure. Von Neumann mettait ainsi en lumière les limites de la démarche descendante (*top-down*) qui est classiquement celle de l'ingénieur. Dans le cas des systèmes complexes, seule a du sens la démarche ascendante (*bottom-up*), qui consiste à explorer ce dont est capable un automate donné. Bientôt, prophétisait von Neumann, le constructeur d'automate serait aussi désarmé devant sa création que le scientifique l'est devant les phénomènes naturels complexes.

L'ingénierie consiste classiquement à concevoir et à fabriquer des structures dont le comportement reproduira les fonctionnalités que l'on juge désirables. Cependant, avec les nanotechnologies, et surtout la convergence nano-biotechnologique, une nouvelle conception se met en place, pour devenir peut-être un jour la conception dominante. Il s'agit cette fois de «se donner» des structures complexes (éventuellement en les puisant dans le réservoir que nous offrent la nature et la vie, par exemple le

cerveau humain, ou bien en les reproduisant artificiellement, par exemple sous la forme d'un réseau de neurones formels) et d'explorer les fonctionnalités dont elles sont capables, en essayant de dégager le rapport structure/fonction: démarche ascendante, donc, et non plus descendante. De manière plus imagée, on peut dire que l'ingénieur, ici, loin de souhaiter la maîtrise, estimera que son entreprise est d'autant plus couronnée de succès que la machine qu'il aura mise au point le surprendra. On observe déjà une telle attitude dans la recherche sur les algorithmes génétiques. Il s'agit de simuler les capacités évolutives d'une «soupe» primitive constituée de programmes d'ordinateur, les plus performants se reproduisant davantage que les autres. On obtient ainsi des algorithmes très performants en effet, puisqu'ils ont été «sélectionnés» selon ce critère, mais on est dans l'incapacité de comprendre pourquoi ils ont ces propriétés. Celui qui veut fabriquer - en fait, *créer* - de la vie ne peut pas ne pas ambitionner de reproduire sa capacité essentielle, qui est de créer à son tour du radicalement nouveau.

Ce serait une erreur de penser que bien qu'originale, notre situation actuelle devant les conséquences de nos choix technologiques n'est pas le produit d'un long processus historique. Discontinuités et ruptures doivent toujours être analysées sur fond de dynamiques continues. Dans son étude magistrale des faiblesses de l'action humaine, *Human Condition*, Hannah Arendt a mis en évidence le paradoxe fondamental de notre

époque : tandis que les pouvoirs humains s'accroissent sous la poussée du progrès technologique, nous sommes de moins en moins à même de maîtriser les conséquences de nos actions. Une longue citation prend ici toute sa valeur car sa pertinence par rapport à notre sujet ne peut être exagérée – et nous devons garder en mémoire que ceci fut écrit en 1958 :

[...] en essayant de supprimer l'action à cause de son incertitude et de préserver de la fragilité les affaires humaines en les traitant comme si elles étaient ou pouvaient devenir les produits d'une technique, on a abouti d'abord à concentrer la faculté d'agir, d'entreprendre des processus nouveaux et spontanés qui n'existeraient pas sans l'homme, sur une attitude envers la nature, qui jusqu'au dernier stade de l'époque moderne, a consisté à explorer les lois naturelles et à fabriquer des objets avec des matériaux naturels. A quel point nous avons commencé à agir sur la nature, au sens littéral du mot, on peut l'entrevoir d'après une remarque faite en passant par un savant qui déclarait fort sérieusement : *'La recherche fondamentale, c'est quand je fais ce que je ne sais pas que je fais.'* [Wernher von Braun, décembre 1957].

Cela commença de façon assez inoffensive par l'expérimentation dans laquelle les hommes ne se contentèrent plus d'observer, d'enregistrer et de contempler ce que la nature telle qu'on la voit

était prête à livrer, mais se mirent à prescrire des conditions et à provoquer des processus naturels. L'évolution qui perfectionna sans cesse l'art de *déclencher des processus élémentaires* qui, sans l'intervention de l'homme, seraient demeurés virtuels et n'auraient peut-être jamais eu lieu, aboutit finalement à un véritable art de *'faire la nature'*, autrement dit de créer des processus 'naturels' qui n'existeraient pas sans l'homme et que la nature terrestre semble incapable d'accomplir elle-même [...]

Le simple fait que les sciences naturelles soient devenues exclusivement des sciences de processus et, au dernier stade, des *sciences de 'processus sans retour' virtuellement irréversibles, irrémédiables*, indique très clairement que, quelle que soit la puissance cérébrale nécessaire pour les déclencher, la faculté humaine sous-jacente qui seule peut être à l'origine de cet état de choses n'est pas une faculté 'théorique', contemplation ou raison : c'est la *faculté d'agir, de déclencher des processus sans précédent, dont l'issue demeure incertaine et imprévisible* dans le domaine, humain ou naturel, où ils vont se dérouler.

Dans cet aspect de l'action [...] on déclenche des processus dont l'issue est imprévisible, de sorte que *l'incertitude plus que la fragilité devient la caractéristique essentielle des affaires humaines*<sup>17</sup>.

<sup>17</sup> Hannah Arendt, *Condition de l'homme moderne*, op. cit., p. 259-261.

Aucun doute qu'avec une incroyable prescience, cette analyse s'applique parfaitement à la convergence NBIC, en particulier sur deux points. Premièrement, l'ambition de (re-) faire la nature est une dimension importante de ce que j'ai appelé la métaphysique sous-jacente au champ. J'y reviendrai dans un moment. Deuxièmement, comme cela a été expliqué plus haut, ce sera une tentation inévitable, pour ne pas dire une tâche ou une obligation, pour les nanotechnologistes de l'avenir, de déclencher des processus sur lesquels ils n'ont aucun contrôle. Le mythe de l'apprenti sorcier doit être actualisé : ce n'est ni par erreur ni par terreur que l'homme sera dépossédé de ses propres créations mais à dessein.

### « JOUER À ÊTRE DIEU »

Le fait de concevoir la nature comme un artefact a des implications éthiques et épistémiques considérables. Il est intéressant d'analyser ce que les promoteurs de la convergence NBIC imaginent être l'état d'esprit de ceux qu'ils prennent pour leurs «ennemis», ou en tout cas leurs critiques. Une expression revient toujours pour désigner cet état d'esprit présumé: les êtres humains n'auraient pas le droit d'usurper des pouvoirs qui n'appartiennent qu'à Dieu: «jouer à être Dieu» [Playing God] serait un jeu interdit<sup>18</sup>. On ajoute souvent que ce tabou serait spécialement le fait du «judéo-christianisme».

Je passe sur le fait que cette imputation méconnaît complètement tant la leçon talmudique que la théologie chrétienne. Elle les confond tout bonnement avec la conception que les Grecs avaient du sacré: les Dieux, jaloux des hommes coupables d'*hybris* (la démesure), leur dépêchent la déesse de la vengeance, *Némésis*. Mais la Bible représente au contraire l'homme comme co-créateur du monde. Comme l'écrit le biologiste Henri Atlan, grand spécialiste du Talmud, analysant la littérature sur le Golem: «On n'[y] trouve, au moins au départ, contrairement à la légende de Faust, aucun jugement négatif sur le savoir et l'activité créatrice des hommes, à l'image de Dieu'. Bien au contraire, c'est dans l'activité créatrice que l'homme atteint la plénitude de son humanité, dans une perspective d'*imitatio Dei* qui lui permet d'être associé à Dieu, en un processus de création continue et perfectible.<sup>19</sup>» J'y reviendrai en conclusion. Quant au christianisme, toute une série d'auteurs importants, de Max Weber à Louis Dumont, de Marcel Gauchet à René Girard, l'analysent comme «la religion de la fin de la religion»: ils le rendent responsable de la désacralisation du monde (le fameux «désenchantement»), et donc de l'élimination progressive de tout tabou, interdit ou limite. C'est d'ailleurs pourquoi les mêmes auteurs font du christianisme la principale cause du développement scientifique et technique de l'Occident, car la science et la technique reposent précisément sur l'affranchissement par rapport à toute limite.

<sup>18</sup> Voir la section "Concerns about 'Playing God'" du rapport *Splicing Life*, op. cit., p. 53 sq.

<sup>19</sup> Henri Atlan, *Les Étincelles de hasard*, op. cit., p. 45.

C'est la science elle-même qui a pris le relais de cette désacralisation du monde opérée par les religions de la Bible, en dépouillant la nature de toute valeur prescriptive ou normative. Il est donc parfaitement vain de vouloir opposer la science et la tradition judéo-chrétienne sur ce point. Le kantisme a donné ses lettres de noblesse philosophiques à cette dévaluation de la nature en en faisant un monde sans intentions ni raisons, habité uniquement par des causes, et en le séparant radicalement du monde de la liberté, où les raisons d'agir tombent sous la juridiction de la loi morale.

Où donc se situe le problème éthique, s'il y en a un ici? Il n'est à l'évidence pas dans la transgression de je ne sais quel tabou ou limite garantie par le sacré, puisque l'évolution conjointe du religieux et de la science a privé de tout fondement le concept même de limite morale, et donc de transgression. Mais c'est précisément cela le problème. Car il n'y a pas de société humaine libre et autonome qui ne repose sur un principe d'autolimitation, même quand elle croit que ce principe lui vient d'une quelconque transcendance. Rousseau puis Kant ont défini la liberté ou l'autonomie comme l'obéissance à la loi que l'on se donne à soi-même. Rousseau voulait que les lois de la Cité aient la même extériorité par rapport aux hommes que les lois de la nature, alors même que ce sont les hommes qui font les premières et qu'ils le savent. Mais dans une société qui rêve de façonner et de fabriquer la nature selon ses désirs et ses besoins, c'est l'idée

même d'une extériorité ou d'une altérité qui perd tout sens. La substitution du faire au donné participe évidemment de ce même processus. Traditionnellement, la nature était définie comme ce qui restait extérieur au monde humain, avec ses désirs, ses conflits, ses turpitudes diverses. Mais si la nature devient dans nos rêves intégralement ce que nous faisons d'elle, il est clair qu'il n'y a plus d'extérieur et que tout dans le monde reflètera tôt ou tard ce que des hommes ont fait ou n'ont pas fait, ont voulu ou bien négligé.

Ce problème éthique est autrement considérable que les questions spécifiques qui portent, par exemple, sur l'« amélioration » [*enhancement*] de telle ou telle capacité cognitive au moyen de diverses techniques. Mais ce qui le rend encore plus insoluble est qu'alors que la responsabilité des hommes sur le monde s'accroît sans limite, les ressources éthiques dont nous disposons s'amenuisent au même rythme. [...]

## LA NATURE ARTIFICIELLE

Au cœur du programme métaphysique de recherche qui sous-tend la convergence NBIC, on trouve un énorme paradoxe. La métaphysique en question se veut clairement *moniste*: on ne dirait plus aujourd'hui que tout, dans l'univers, procède de la même *substance*, mais on dira que tout est soumis aux mêmes *principes d'organisation*: la nature, la vie et l'esprit. Le mot d'ordre

des sciences cognitives est: «*naturaliser l'esprit*». Il s'agit de redonner à l'esprit (et à la vie) leur place pleine et entière au sein du monde naturel. Or il se trouve que les principes d'organisation supposés être communs à tout ce qui existe dans l'univers sont des principes mécanistes. La machine à traiter de l'information selon des règles fixes, c'est-à-dire l'algorithme, constitue le modèle unique de tout ce qui existe. Chronologiquement, et contrairement peut-être à certaines idées reçues, c'est d'abord l'esprit qui a été assimilé à un algorithme (ou machine de Turing : modèle de McCulloch et Pitts, 1943) ; puis ce fut le tour de la vie, avec la naissance de la biologie moléculaire (Max Delbrück et le groupe du phage, 1949) ; et, seulement plus tard, la thèse selon laquelle les lois de la physique sont récursives (ou Turing computables). La naturalisation de l'esprit se confond donc avec la mécanisation de l'esprit. C'est encore une fois la littérature de battage qui dit le mieux cela dans la mesure où dans sa grande naïveté philosophique, elle ne s'embarrasse pas de prudences rhétoriques. Le futurologue américain Damien Broderick fait un condensé saisissant de l'histoire de l'évolution biologique dans les termes suivants. Une fois de plus, chacun des mots employés est révélateur, à commencer par celui par lequel Broderick désigne les êtres vivants: ce sont des «réplicateurs vivants » :

Des *algorithmes génétiques* en nombre astronomique titubaient à la surface de la Terre et sous la mer, jusqu'à des

niveaux très profonds, comme nous le savons maintenant, pendant des milliards d'années, se répliquant, mutant, tout en étant sélectionnés en fonction du succès de leurs expressions, c'est-à-dire des êtres biologiques qu'ils *fabriquaient*, ceux-ci se livrant à une compétition pour la survie dans le monde macroscopique. Finalement, toute l'écologie du vivant sur la planète a *accumulé, et représente une quantité colossale d'information, comprimée, schématisée*<sup>20</sup>.

Les cellules eucaryotes et procaryotes par lesquelles la vie a commencé se trouvent assimilées à des productions de l'esprit humain, les algorithmes génétiques, qui n'apparaîtront que dans les dernières décennies du vingtième siècle. Ces êtres sont un condensé *d'information*, le *blueprint* pour la *fabrication* des êtres vivants eux-mêmes. Le monisme matérialiste de la science moderne est devenu soudainement un monisme spiritualiste. Si l'esprit ne fait qu'un avec la nature, c'est que la nature est interprétée comme si c'était une production de l'esprit. Le tour fait penser au célèbre clown suisse Grock. Magnifique concertiste, il s'approchait de son Steinway et découvrait que son tabouret était trop éloigné du piano. Il se mettait alors à tirer péniblement celui-ci pour le rapprocher du tabouret. Le piano, c'est la nature, et le tabouret, l'esprit. C'est la réécriture de la nature dans des termes qui pourraient faire croire que l'esprit en est le créateur, qui permet de dire que l'on a rapproché l'esprit de la nature. Une ex-

<sup>20</sup> Damien Broderick, *The Spike*, op. cit., p. 116. Je souligne.



pression en forme d'oxymore résume fort bien tout cela: la nature est devenue *nature artificielle*.

L'étape suivante consiste évidemment à se demander si l'esprit ne pourrait pas prendre le relais de la nature pour accomplir plus efficacement et intelligemment son œuvre créatrice. Broderick interroge de façon rhétorique: «Ne peut-on penser que des nano-systèmes, conçus par l'esprit humain, court-circuiteront toute cette errance darwinienne pour se précipiter tout droit vers le succès du design?<sup>21</sup>»

Dans une perspective d'études culturelles comparées, il est fascinant de voir la science américaine, qui a dû se battre de haute lutte pour chasser de l'enseignement public toute trace de créationnisme, y compris dans ses avatars les plus récents, comme l'*intelligent design*, retrouver par le biais du programme nanotechnologique la problématique du design, avec simplement désormais l'homme dans le rôle du démiurge.

contraire comme le couronnement d'un long chemin d'ascension vers la sainteté et la connaissance, les deux allant de pair dans la perspective d'une *imitatio Dei* : «En effet, comment savoir que l'initié a réussi à déchiffrer et à bien comprendre les lois de la création du monde, si ce n'est en vérifiant que son savoir est efficace en ce qu'il lui permet, à lui aussi, de créer un monde ? Comment savoir que sa connaissance de la nature humaine est correcte, si ce n'est en vérifiant qu'elle lui permet de créer un homme?<sup>22</sup>»

Le critère de la vérité du savoir du sage, comme aujourd'hui le critère de la vérité scientifique, est, selon l'expression fameuse de Gianbattista Vico, le *verum factum* : nous ne connaissons vraiment que ce que nous sommes capables de faire ou de refaire. Le cas de Warren McCulloch, l'artisan des conférences Macy et, en ce sens, le véritable fondateur de la cybernétique, bien plus que Norbert Wiener, est de ce point de vue éclairant<sup>23</sup>. Neuropsychiatre de son état, McCulloch fut au long des années de plus en plus déçu par les méthodes des neurosciences. Il se tourna vers la logique et ce qui ne s'appelait pas encore l'intelligence artificielle. Les étudiants ou disciples qui l'entouraient au MIT s'appelaient Seymour Papert ou Marvin Minsky - lequel Minsky allait d'ailleurs, bien plus tard, former un certain Eric Drexler. Le neurophysiologiste Jerome Lettvin a décrit en ces termes l'évolution intellectuelle de McCulloch, qu'il admirait énormément :

## JÉRÉMIE ET LE GOLEM

J'aimerais conclure cette réflexion par un récit talmudique du XIII<sup>e</sup> siècle que m'a rapporté le biophysicien français Henri Atlan. Ce récit met en scène le prophète Jérémie alors qu'il vient de réussir la création d'un golem. Le récit ne présente pas du tout cette création comme un acte de révolte contre Dieu, mais bien au

<sup>21</sup> Ibid., p. 118.

<sup>22</sup> Henri Atlan, *Les Etincelles de hasard*, op. cit., p. 49.

<sup>23</sup> Voir Jean-Pierre Dupuy, *The Mechanization of the Mind*, op. cit.

Il se vouait à la tâche de savoir comment le cerveau fonctionne dans les termes mêmes où le créateur d'une machine connaît ses rouages. La clé d'un tel savoir n'est pas dans l'observation, mais dans la fabrication de modèles, que l'on confronte ensuite aux données. Mais la *poiesis* doit venir d'abord. Et McCulloch préférerait risquer l'échec dans sa tentative de créer un cerveau que rencontrer la réussite dans l'amélioration de la description des cerveaux existants.<sup>24</sup>

Mais revenons à Jérémie et son homme artificiel. Contrairement à d'autres golems, celui-ci parle. Tout naturellement, il s'adresse en premier lieu à son créateur et lui dit ceci, en appelant à sa conscience : « Te rends-tu compte de la confusion que

tu viens d'introduire dans le monde ? A partir d'aujourd'hui, quand on rencontrera un homme ou une femme dans la rue, on ne saura plus si c'est une créature de Dieu ou la tienne ! ». Il se révèle que Jérémie n'avait pas pensé à cela. Très troublé, il demande conseil à son golem pour réparer ce qu'il a fait. Et l'homme artificiel lui répond : « Tu n'as plus qu'à me défaire comme tu m'as fait. » Jérémie s'exécute mais tire pourtant la leçon suivante : nous ne devons pas renoncer à atteindre la connaissance parfaite qui nous rend capables de créer un homme, mais lorsque nous y sommes parvenus, nous devons nous abstenir de le faire. Atlan conclut : « Grande leçon, qu'il nous revient de méditer.<sup>25</sup> »

C'est ce que je me permets de nous inviter à faire, avant qu'il ne soit trop tard.

**Jean-Pierre DUPUY**

*Professeur à l'École Polytechnique, Paris et  
à l'Université Stanford (Californie)  
Membre de l'Académie des Technologies*

<sup>24</sup> Jerome Lettvin, "Warren and Walter", inédit; archives personnelles de Heinz von Foerster. Cité dans Jean-Pierre Dupuy, *The Mechanization of the Mind*, op. cit., p. 137.

<sup>25</sup> Récit raconté dans *Les Étincelles de hasard*, op. cit., p. 49.





## C A L L

Octobre 2002

Scientists can be proud of having brought so much new knowledge thereby contributing to free man from multiple sufferings and painful tasks.

Never before has science given man so much power without leaving him the time to adapt to it. This misadjustment has entailed profound disfunctions, enormous disparities and even threats for the future of humanity.

Never have scientists had more duties. These duties are two-fold :

- on one hand, contribute to the ACQUISITION of new knowledge, without any limitation, in all domains,
- and on the other hand, bring their expertise to a rational APPLICATION of the technologies which ensue from it and for which they are also morally responsible.

In order to do this, scientists should come out of their ivory towers and get involved in society, not only as experts with the authorities, but also by addressing the general public by all modern means of broadcasting.

Scientists have a particular responsibility for presenting discoveries and especially the practical applications which ensue from them or would ensue from them with all the loyalty, the light

and the necessary objectivity, putting themselves within the reach of the listener, be he enlightened or ignorant, avoiding the sensational (too often used in media) and any useless alarmism.

By being informed, and informed at the right time, the general public opinion will use all its weight to influence political or industrial decision-makers, so that new technologies are not used in excessive or wrong ways. In a democracy, politicians listen to the voters and all around the world manufacturers listen to potential buyers and both voice public opinion. It is this huge force which we ask you to use *so that man does not have to undergo any more his lot but can henceforth direct clearly his fate.*

To be effective, a big movement of international dimension has been created, a non-profit association, without any political, religious or financial ties : the Universal Movement of the Scientific Responsibility, branches of which exist already in France, in Mexico and in Japan.

We urge you to join and to develop in your respective countries branches setting similar objectives but using natural approaches adapted to every culture.

Never have scientists had so evident and so pressing duties.

**Ismail SERAGELDIN**  
**President**  
**UMSR - International**

**Jean DAUSSET**  
**Président d'Honneur**  
**MURS - France**

MURS-France    Hôpital Saint-Louis  
1, avenue Claude Vellefaux

Quadrilatère Historique  
75475 PARIS cedex 10

☎ 01.47.03.38.21    📠 01.47.03.38.22  
courriel : [secretariat@murs-france.asso.fr](mailto:secretariat@murs-france.asso.fr)