

« Technologies émergentes et sagesse collective. Comprendre, faire comprendre, maîtriser » Un vaste programme de plus ?

DANIEL ANDLER

Daniel Andler, professeur émérite de philosophie de Sorbonne Université, est membre du Département d'études cognitives de l'École normale supérieure et de l'UMR "Sciences, normes, démocratie" de Sorbonne Université. À l'Académie des sciences morales et politiques dont il est membre, il est responsable du projet "Technologies émergentes et sagesse collective" (TESaCo).

Je remercie le Chancelier honoraire de l'Institut, qui s'exprimait aussi en tant que président de la Fondation Simone et Cino Del Duca, pour ses paroles de bienvenue et le soutien qu'il a apporté au projet « Technologies émergentes et sagesse collective » (TESaCo) dès l'origine. Ce projet s'inscrit dans le prolongement d'un groupe de travail inter-académique dirigé par Jean-François Mattei en 2017-2018, consacré aux biotechnologies et à leurs applications médicales, ainsi qu'à l'intelligence artificielle.

Avant de présenter TESaCo plus en détail, je veux remercier également les confrères qui lui ont fait bon accueil, et l'équipe administrative, en particulier la Secrétaire générale Véronique Duchaud-Fuselli et Marianne Tomi, chargée de mission, dont l'expérience est précieuse. Je remercie tout particulièrement les intervenants, qui ont accepté notre invitation, tard venue, et ont ainsi montré qu'ils prennent au sérieux notre proposition, qui s'écarte pourtant des chemins battus. Le public, composé pour une bonne part de spécialistes des questions qui vont nous occuper, doit être remercié pour les mêmes raisons.

Enfin, je voudrais rapidement présenter l'équipe qui m'entoure, et sans la participation décisive de laquelle le présent colloque n'aurait pas pu avoir lieu. Elle est actuellement constituée de quatre conseillers scientifiques : Fabienne Cazalis, Florian Forestier, Jean-Lorenceanu, Alex Peluffo, et d'une assistante chargée de l'organisation et de la communication, Margaux Berrettoni¹. Vous trouverez sur le site

¹ L'équipe s'est renforcée depuis ; l'ont rejointe Serena Ciranna en tant qu'attachée de recherche, Célia Zolynski, Sonia Desmoulin-Canselier, Bastien Guerry et Mehdi Khamassi, en tant que conseillers scientifiques. Anne Le Goff, Soraya de Chadarevian et Jacopo Domenicucci ont également rejoint l'équipe en tant que membres associés .

de TESaCo une brève biographie de chacun d'eux. L'équipe sera prochainement renforcée par trois ou quatre conseillers scientifiques et par un assistant ou une assistante de recherche. Au sein de l'Académie, Marianne Tomi assure l'interface.

1. Une idée qui tombe sous le sens

Il n'est nul besoin de souligner l'importance qu'ont prise certaines technologies nouvelles — qu'on peut dire « émergentes » pour indiquer qu'elles sont présentes, comme l'électricité et l'automobile, mais que, contrairement à ces dernières et à la quasi-totalité des technologies sur lesquelles repose encore aujourd'hui la société humaine, elles sont en plein essor, croissant et se transformant à vue d'oeil. Ces « néo-technologies » — autre terme parfois employé pour les désigner — forment un ensemble aux contours mouvants, mais comportant au moins les domaines à haut potentiel suivants : les technologies de l'information et de la communication (TIC), en y incluant l'intelligence artificielle et la robotique ; les biotechnologies ; les technologies dérivées des sciences cognitives et des neurosciences ; les nanotechnologies. S'en rapproche tout ce qui touche à la transition énergétique et au changement climatique. Une cartographie un peu plus précise sera proposée dans un court moment.

De tout cela, journalistes, essayistes et aujourd'hui « blogueurs » et autres « youtubers » se sont emparés, proposant en général une vision positive des progrès qu'apportent ou qu'apporteront bientôt les nouvelles technologies. Mais après cette phase d'enchantement, nous vivons à l'heure du « techlash »² — en réalité les deux phases forment un mélange complexe, où voisinent l'émerveillement et la répulsion, l'espoir et la consternation. La question est en tout cas au coeur du débat public.

Il peut cependant être utile de s'attarder sur trois aspects de la situation présente. Le premier est l'accélération du rythme des innovations, de leur propagation planétaire et de leur impact sur les sphères d'activité et d'existence de l'humanité entière. Si pour l'essentiel elles proviennent de percées scientifiques datant du siècle précédent, résultant également de choix politiques, économiques et techniques antérieurs au XXI^e siècle, il n'en est pas moins vrai que c'est au cours des vingt premières années de ce siècle que sont apparues les technologies qui bouleversent notre quotidien et bousculent l'ordre économique, social et politique, ou à tout le moins qu'elles ont pris les proportions qu'on leur connaît aujourd'hui. Quelques dates le montrent³ :

2001 iPod

2001 Wikipedia

2002 Paypal

2002 Réseaux sociaux (Linkedin 2002,

² Mot-valise formé de « backlash » (réaction violente, retour de bâton...) et de « technology », « tech » étant devenu une abréviation courante ainsi qu'un suffixe passe-partout (du vénérable « hi-tech », aux plus récents « fintech », « lawtech », « healthtech », « edtech »,...).

³ Beaucoup d'entre elles appelleraient des précisions voire des critiques. Dans de nombreux cas, l'innovation indiquée résulte d'une avancée relativement mineure sur le plan théorique, mais aux effets énormes; le deep learning (ou apprentissage profond) en est un exemple caractéristique, puisque pour l'essentiel il remonte aux années 1980. Une réserve analogue s'applique à l'impression 3D. L'internet des objets prend forme en 2004, même si l'idée et le nom

remontent à 1999. Les armes létales autonomes (ALA) ne sont parfaitement circonscrites ni dans leur définition (que doit-on entendre par « autonomie »?) ni dans leur apparition (elles n'ont pas soudain surgi des laboratoires militaires en 2013). Mais par-delà l'arbitraire de certains choix d'exemples et de datation, cette liste confirme largement une thèse qui n'a d'ailleurs rien de controversé. Une absence significative est celle des nanosciences, y compris la nanomédecine et la nanoparmacologie (dont Patrick Couvreur parlera le 31 janvier) : elles résultent en effet d'une progression continue remontant aux deux dernières décennies du XX^e siècle. Il n'en demeure pas moins que les nanotechnologies sont reconnues comme l'une des technologies clés du XXI^e siècle.

Facebook 2004, Twitter 2006, Instagram 2010...)

2002 Vote électronique

2004 Écran tactile équipé du balayage (swipe) et du pincement (pinch)

2004 IoT – internet des objets

2005 Googlemap

2005 Youtube

2006 Impression 3D

2006 Deep learning (apprentissage profond)

2006 Streaming

2007 Kindle

2007 Smartphones (iPhone 2007; premier smartphone Android 2008)

2008 Waze

2010 iPad

2012 CRISPR-Cas9

2013 Présentation des premières armes létales autonomes

2016 Blockchain et crypto monnaies

2017 Siri et autres assistants intelligents

La vague, née dans l'après-2^e Guerre mondiale, s'est abattue sur notre époque comme un tsunami.

Le deuxième aspect de la situation qui mérite notre attention est qu'il est impossible de localiser avec précision l'origine du caractère « disruptif » de chaque innovation, d'une part, son impact d'autre part. En premier lieu, les différentes technologies se nourrissent les unes des autres, et leur effet conjoint dépasse celui qu'elles peuvent avoir séparément. C'est pourquoi on parle à leur propos de technologies convergentes. En second lieu, elles pénètrent de nombreux domaines, les transformant les uns après les autres, et parfois en même temps, formant des combinaisons inédites et largement imprévisibles avec les écologies locales des secteurs d'activité et d'existence qu'elles envahissent. En ce sens, on pourrait dire d'elles qu'elles sont divergentes, étendant leurs branches dans toutes les directions.

Le troisième aspect, non moins évident que les précédents, est que nous avons largement perdu la maîtrise de ces technologies — ce dont témoigne le techlash dans son ensemble,

et de manière à la fois plus précise et plus profonde les inquiétudes que suscitent, dans le domaine privé, la mainmise sur des biens essentiels, tels que la privauté, l'attention, la maîtrise du temps, l'intimité, l'identité; dans le domaine public, la confiscation des prérogatives des États et des pouvoirs publics par des organismes privés transnationaux devenus des géants économiques et politiques, le bouleversement des relations internationales et des pratiques militaires, financières, commerciales, etc., tout autant que la transformation du totalitarisme en ultra-totalitarisme, et celle, moins apparemment dramatique mais insidieuse, d'institutions démocratiques en instruments de surveillance voire d'asservissement par des administrations publiques ou des organismes privés.

Tels des parents qui ont laissé leurs garnements d'enfants faire leurs expériences de chimie au grenier qui ont mis le feu à la maison, nous nous avisons que nous eussions dû faire preuve de davantage de vigilance. Partout dans le monde se constituent des comités de sages chargés d'établir des chartes éthiques censées mettre de l'ordre et éviter les conséquences manifestement indésirables d'un développement incontrôlé des technologies dont les dégâts sont manifestes et le potentiel de nuisance terrifiant.

Mais les parents partageaient d'un bon sentiment : ils voulaient laisser leurs enfants acquérir par eux-mêmes quelques rudiments de science. L'excuse vaut pour nous, et plus encore : les technologies dont certains effets nous accablent aujourd'hui résultent de progrès scientifiques et techniques auxquels il eût été inconcevable de renoncer d'avance, et qui de fait se traduisent aujourd'hui en avancées en maints domaines et en promesses qui ne sont pas nécessairement vaines. Tel est donc le défi : maîtriser sans étouffer, reprendre la main sans entraver. Mais qui doit la reprendre, et comment ?

La question centrale à laquelle il s'agit d'apporter une réponse est donc celle-ci : quels moyens nos sociétés peuvent-elles et doivent-

elles se donner pour acquérir une sagesse collective, c'est-à-dire la capacité de prendre les meilleures décisions quant au développement et à l'usage des nouvelles technologies ? Une condition préalable est de les comprendre, d'en saisir sinon le détail du moins ce qu'Auguste Comte appelait, à propos des sciences de son temps, les « généralités », de manière à en apprécier non seulement les effets actuels, mais les potentialités à court et moyen terme, et de mesurer leurs impacts anthropologiques, sociaux, éthiques, politiques, économiques. Cet objectif — comprendre en ce sens étendu — est d'autant plus difficile à atteindre que le phénomène est complexe, mouvant et mû par de nombreux facteurs.

Si difficile que ce soit, comprendre ne suffit pas. Quel que soit l'individu ou le collectif qui possède cette compréhension, il ne peut à lui seul agir en conséquence : c'est la société entière, ou en tout cas de nombreux acteurs en son sein, individuels et collectifs, qui en ont les moyens. Encore faut-il que ces acteurs aient eux-mêmes la compréhension requise, ce qui met ceux qui la possèdent en devoir de la transmettre. Du reste, chacun de nous est en quête de compréhension : les professeurs sont aussi des élèves. De plus, c'est lorsque nous nous interrogeons sur les moyens de la maîtrise que nous commençons à comprendre ce qu'il s'agit de comprendre. Comprendre, faire comprendre, maîtriser constituent ainsi les trois moments d'un apprentissage de la sagesse collective, trois moments analytiquement distincts mais étroitement imbriqués.

Cet apprentissage est indispensable, et nous devons y contribuer : chacun de nous, mais aussi, entre autres organismes, l'Académie des sciences morales et politiques, qui a pour mission d'informer et de conseiller les pouvoirs publics sur toute question engageant des enjeux éthiques et politiques. Telle est l'idée qui tombe sous le sens.

2. Une tâche au-dessus de nos forces ?

L'objection n'est pas moins évidente. Le programme en question est en effet inscrit depuis des années à l'ordre du jour de l'humanité entière, et en France comme dans le monde d'innombrables et puissantes organisations spécialisées, publiques et privées, y travaillent, enrôlant les compétences d'une multitude de chercheurs et d'experts. De plus, si puissantes qu'elles soient, elles ne s'attachent généralement qu'à une partie de la tâche : ce ne sont pas les mêmes qui s'occupent d'intelligence artificielle et d'ingénierie génétique, de l'économie d'internet et des effets psychologiques des réseaux sociaux sur les adolescents, des nouveaux nano-médicaments anti-cancéreux et d'armes létales autonomes, des effets de l'automatisation sur le marché du travail et du statut juridique des données personnelles. Tout près de notre Compagnie, l'Académie des sciences, l'Académie des technologies, l'Académie de médecine, comme le Comité consultatif national d'éthique, France-Stratégie et d'autres instances encore, abordent certaines de ces questions, chacune de son côté et dans des comités distincts. Pour espérer progresser, on n'a d'autre choix que de diviser le domaine en secteurs, d'autant que chacun d'entre eux exige des connaissances approfondies qu'on n'acquiert pas en un jour, même au niveau des « généralités ». Qui trop embrasse mal étreint. Seuls ont des chances de porter des fruits les projets dont la cible est convenablement circonscrite.

La réplique est en deux temps : oui, et non. Oui, bien entendu, chaque secteur, chaque dimension doivent être pris en charge séparément par des équipes compétentes, capables de développer une pensée cohérente, critique et progressive, au plus près des connaissances pertinentes et appuyée sur une expertise reconnue en matière d'évaluation et de prospective⁴. Loin de prétendre nous substituer à ces équipes, nous comptons nous reposer sur leurs

travaux et faire appel à leur jugement. Nous sommes d'ailleurs en état de contribuer à leur réflexion, s'agissant de certains secteurs et certaines dimensions. Et c'est précisément parce qu'elles existent que nous pouvons et devons nous assigner une tâche différente.

Mais si se cache derrière cette reconnaissance de la nécessité d'une approche sectorielle l'idée qu'elle est non seulement nécessaire mais suffisante, non seulement possible mais la seule possible, alors il faut selon moi rejeter cette idée. À cela trois raisons qui se recoupent en partie.

La première est qu'il semble exister entre les problématiques associées aux différentes technologies des similitudes qu'il est probablement profitable de dégager. Il vaut sans doute la peine d'examiner dans quelle mesure ce qui se fait, par exemple, dans le domaine des armes létales autonomes, ou celui de la protection des données privées sur internet, ou celui de la stimulation cérébrale profonde, peut nous éclairer quant à la régulation de l'ingénierie génétique, ou à l'intervention de l'État dans les choix des individus par les « nudges », ou aux procédures de la justice prédictive. Il ne s'agit pas seulement de repérer ce que ces technologies ont en commun, mais aussi, au second degré, la façon dont elles sont développées, la manière dont les problèmes qu'elles soulèvent sont abordés, par qui, à quel moment de leur développement, avec quels effets, etc.

La deuxième raison, à laquelle il a été fait allusion plus haut, est qu'il existe une puissante synergie entre les néotechnologies. Cette convergence a fait l'objet d'une théorisation au

tout début du siècle, sous l'étiquette « NBIC », à partir des quatre secteurs que sont les *Nanotechnologies*, les *Biotechnologies*, les technologies de l'*Information* et les sciences *Cognitives*. M. Roco et W. Bainbridge, les géniteurs de NBIC⁵, faisaient valoir que, séparément, chacun de ces domaines avait une influence importante mais que s'amplifiant mutuellement ils exerceraient conjointement sur la société une force de transformation inédite. En dépit de certaines exagérations et extrapolations hasardeuses, l'idée a fait son chemin, à juste titre. De nombreuses études ont été à l'époque engagées sur le thème de la convergence. La Commission européenne mandatait très rapidement un groupe d'experts « de haut niveau » (dont le signataire du présent texte), afin de dégager une perspective proprement européenne et humaniste sur NBIC — c'est-à-dire précisément une voie vers la maîtrise de ces technologies : TESaCo s'inscrit dans la continuité du rapport que nous avons remis à l'époque⁶. L'accélération tout comme la convergence prédites par Roco et Bainbridge ont été pleinement confirmées, alors même qu'ils ne pouvaient prévoir que très incomplètement les percées dont certaines ont été énumérées ci-dessus. Pour ne prendre qu'un exemple parmi les plus criants : on ne parlerait pas aujourd'hui d'intelligence artificielle comme d'une révolution sans la percée du deep learning, percée qui ne serait rien sans les données massives, elles-mêmes rendues possibles par internet et les ordinateurs portables et les smartphones, ces derniers largement tributaires des nanotechnologies. De manière plus large, ce que la convergence a en-

⁴ Les ouvrages de chercheurs et d'essayistes, qui ne sont pas tous intégrés dans des équipes professionnelles, sont nombreux et constituent également une ressource appréciable.

⁵ Mihail C. Roco, Williams S. Bainbridge, *Converging Technologies for Improving Human Performance*, National Science Foundation & Department of Commerce, 2002; disponible en ligne : <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/bioecon-%28%23%20023SUPP%29%20NSF-NBIC.pdf>; également republié par Springer.

⁶ Alfred Nordmann, ed., *Converging Technologies - Shaping the future of European Sciences*, Bruxelles: Commission européenne, 2004; en ligne https://www.researchgate.net/publication/311456580_Converging_Technologies_-_Shaping_the_Future_of_European_Societies ou sur mon site <http://andler.dec.ens.fr/> (et ailleurs).

gendré est un univers parfois désigné comme « le numérique », terme auquel je préfère un néologisme que je croyais avoir inventé mais n'ai fait que ré-inventer, la « numérisphère », sorte d'englobant au sein duquel une part considérable des activités humaines se déploie désormais⁷. Les néo-technologies sont invasives, à la manière dont l'électricité a fini par envahir toutes les activités humaines, mais de manière beaucoup plus large, ce que je soulignais plus haut en parlant de technologies non seulement convergentes, mais divergentes.

Enfin, troisième raison pour ne pas se contenter d'une approche sectorielle, ces innovations sont rendues possibles par un contexte social, politique, économique qui est commun aux différentes technologies. Au total, il existe une technosphère, dont une part importante est la numérisphère, qu'il faut essayer d'appréhender dans sa totalité, dans son organicité. Ce qui ne nous condamne pas à nous enfoncer dans un holisme confus, mais nous incite à mettre en oeuvre concrètement une stratégie d'aller-retour entre analyse et synthèse, entre approches monographiques, approches comparatives et approches globales.

Encore faut-il partir d'une première cartographie du domaine qu'il s'agit, progressivement, d'embrasser. Celle que je propose, à titre de simple point de départ, est composée de cinq « couches » ou « dimensions » (ni l'un ni l'autre de ces termes ne convenant parfaitement).

La première est celle des ressources de base, sans lesquelles les technologies n'existeraient pas même à titre d'hypothèse sérieuse. Elles comportent :

- Les sciences fondamentales : mathématiques et informatique théorique ; physique-chimie; biologie; sciences cognitives (certaines branches de ces disciplines sont plus directement concernées que d'autres, mais il

n'existe pas de frontière fixe et nette les délimitant)

- L'informatique à nouveau, mais cette fois en tant que cadre technoscientifique, incluant des pratiques, un savoir-faire, une vaste trousse à outils évolutifs

- Les télécommunications (constituant, comme l'informatique, un cadre technoscientifique)

- La numérisphère, recouvrant tout ce qui relève du « numérique », résultant de la symbiose entre l'informatique et les télécommunications, et comprise comme une source fondamentale, ubiquitaire, et comme un cadre de pensée, un « englobant ».

La deuxième couche est celle des technologies elles-mêmes, non pas conçues comme l'ensemble sans cesse croissant d'outils, gadgets, dispositifs, que je place dans la couche suivante, mais comme des corps de métier, chacun pourvu d'une certaine unité, de cohésion et se comprenant comme communauté partageant un ethos, un savoir-faire. Trouvent une place dans cette deuxième couche :

- L'intelligence artificielle (au sens large communément utilisé aujourd'hui)

- La science des données (data science)

- La robotique, le génie mécanique, la biomécanique, la théorie du contrôle

- Les nanotechnologies

- Les biotechnologies

- Les neurotechnologies et technologies cognitives

Cette sous-liste comporte une bonne part d'arbitraire, d'autant que chacune de ses lignes cache des subdivisions importantes, et qu'entre certaines rubriques et sous-rubriques existent des relations si étroites qu'elles justifieraient d'autres regroupements.

⁷ Le terme reçoit chez mes prédécesseurs une acception plus restreinte. Ma numérisphère est parfois appelée «

cybersphère », ce vocable ayant lui-même une diversité d'acceptions.

La troisième couche rassemble les outils. Produits par la couche précédente, ils sont déployés par la couche suivante : ce sont eux dont celle-ci fait l'acquisition et s'efforce de maîtriser pour l'exploiter au mieux. L'impact des technologies est médié par les outils. Voici quelques-uns de ces outils, à titre illustratif, étant bien entendu qu'ils se comptent probablement par centaines⁸ :

- Blockchain, cryptomonnaies
- Bases de données
- Algorithmes, progiciels, logiciels, « apps »
- Prothèses, exosquelette
- Robots autonomes, dont : armes létales ; véhicules autonomes...
- Détection à distance par satellites⁹

La quatrième couche est celle des impacts, des secteurs d'activité et des modes d'existence que les technologies transforment par le biais des outils qu'elles produisent. Ces impacts se rangent en deux grandes catégories. La première est celle des effets transversaux : effets communs, partagés par plusieurs secteurs ; effets globaux, résultant de la combinaison des différentes technologies, déployées dans les différents domaines.

La première sous-catégorie inclut des phénomènes tels que :

- Le double effet : un même outil, dans un domaine donné, peut donner le pire et le meilleur, ou encore être utilisé à des fins thérapeu-

tiques (pour « réparer » l'individu lésé) ou d'augmentation (pour « doper » l'individu) ;

- L'universalité : un même outil peut être déployé dans différents domaines (ce qui était rare jadis, mais devient presque la règle) ;
- L'extension des risques : un même outil peut constituer, dans différents domaines, le même risque ou danger, par exemple menacer la vie privée, accroître la surveillance sociale, priver de leurs pleins droits certaines personnes ou populations, etc.

La seconde catégorie rassemble les effets résultant de la combinaison de différentes technologies, déployées dans différents domaines et à différents niveaux et affectant des pans entiers de l'existence :

- La sursaturation ou obésité informationnelle
- La multiplication indéfinie des choix, des décisions
- Le formatage du quotidien, rythmé par un enchaînement toujours plus serré de routines
- La formation et la dissolution des communautés
- Le transhumanisme, comme horizon et comme repoussoir...

Ici encore on peut contester la pertinence de la taxinomie proposée, mais on ne peut guère douter de l'importance des phénomènes ainsi regroupés : tous font l'objet d'une multitude d'études et de conversations, et tous sont pour une bonne partie l'effet des outils néotechno-

⁸ Les différents champs technoscientifiques affectent les activités via les outils, lesquels outils à leur tour affectent différents domaines d'activité, différents outils affectant un domaine de différentes façons, et beaucoup d'outils résultant d'une combinaison de résultats et de méthodes venant de différents champs de la couche des technologies. On peut parfois se demander, pour un item donné, s'il relève de la deuxième ou de la troisième couche. Ainsi, les désormais célèbres « algorithmes » du deep learning sont le produit de l'intelligence artificielle, mais ce sont aussi des outils que les différents secteurs économiques prennent « sur les rayons ». Il en va de même des outils de modification du génome.

⁹ Cet outil (en réalité une vaste famille d'outils) est intéressant. En premier lieu, du fait qu'il résulte de la convergence de trois champs : l'IA pour la reconnaissance des formes; les télécommunications; la technologie spatiale pour le satellite. En second lieu parce qu'il est « divergent » au sens indiqué précédemment : il intervient dans différents domaines, avec des valences différentes — dans le guidage des véhicules, dans la surveillance militaire et civile (mouvements de troupe, incendies volontaires en Amazonie, bases nucléaires...), en archéologie, en agronomie... mais aussi en surveillance des individus. Notons que les progrès en précision changent la donne (un phénomène qu'on observe dans d'autres applications de l'IA et de la robotique).

logiques, même s'ils s'inscrivent aussi dans des logiques politiques, économiques, sociologiques, anthropologiques.

Plus familière est la seconde catégorie d'effets, ceux qui concernent des secteurs particuliers d'activité. La longue liste qui suit n'est nullement exhaustive :

- Santé, handicap
- Droit
- Éducation
- Media d'information
- Recherche scientifique
- Politique, décisions stratégiques, conduite de projets
- Processus démocratiques, participation, intelligence collective
- Travail : industrie
- Travail : services et administrations
- Guerre
- Sécurité, police, surveillance
- Économie, finance, commerce
- Environnement, écologie
- Énergie
- Transports
- Urbanisme, aménagement du territoire
- Loisirs
- Organisation des rapports sociaux (ex. réseaux sociaux)

La cinquième et dernière couche rassemble les réponses à la question : De qui est-ce l'affaire ? Il s'agit des disciplines et instances qui réfléchissent aux impacts énumérés à l'instant, s'en inquiètent, et sont pour certaines chargées de réguler les technologies, de pallier ou de prévenir les effets néfastes des outils qu'elles produisent, et qui tentent plus généralement d'orienter le processus de transformation technologique, aux différents niveaux et dans les différents domaines et dimensions (spécifiques ou générales).

Bref, la cinquième couche est celle à qui échoit la responsabilité de comprendre, de faire comprendre et de maîtriser, pour reprendre le slogan du présent projet. On y distingue trois grandes composantes :

- Les disciplines académiques, dont

- o La philosophie des sciences et l'épistémologie sociale

- o Le domaine STS (science-technologie-société), la philosophie de la technologie, la prospective

- o La philosophie politique, les sciences politiques

- o Le droit

- o L'éthique

- o L'économie

- o La sociologie et l'anthropologie

- o Les théories de la complexité

- Les instances politiques et citoyennes

- Les technologies elles-mêmes (les corps de métier de la troisième couche).

Ces différentes composantes interagissent mais tendent à se grouper en trois pôles : un pôle scientifique (auquel participent plus particulièrement la philosophie des sciences, l'épistémologie sociale, la philosophie de la technologie, les théories de la complexité...) ; un pôle normatif (philosophie et sciences politiques, droit, éthique...) ; un pôle anthropologique (économie, sociologie, prospective). Ces trois étiquettes sont très discutables et importent peu. En revanche, on peut noter que les corps de métier technologiques (ce qui est paradoxal), tout comme les instances politiques et citoyennes (à quoi on peut s'attendre particulièrement dans notre pays) interviennent surtout dans les deux derniers. L'une des tâches auxquelles TESaCo peut contribuer est d'établir des ponts entre ces trois pôles.

Dans l'ensemble, cependant, de cette cartographie nous pourrions tirer une conclusion doublement pessimiste. D'une part, ce domaine est tellement vaste, multidimensionnel, compliqué, hétérogène qu'il ne peut faire l'objet d'une approche globale — qui se mettrait en tête de considérer, que sais-je, l'existence humaine comme un domaine d'étude (à moins de le restreindre justement à une dimension ou un aspect particulier) ? En prétendant en dresser une carte, je n'aurais fait que nous engager dans une impasse. D'autre part, à supposer qu'il s'agisse malgré tout d'un domaine

d'étude, il est bien trop considérable pour qu'avec nos modestes forces nous puissions prétendre nous y attaquer — nous serions ainsi revenus au point de départ de la présente section, clairement indiqué par son intitulé.

3. Une intelligence collective pour un hyper-problème

En réalité, c'est tout le contraire : c'est précisément en raison de sa nature particulière que notre domaine appelle une approche particulière, qui ne se réduit pas à une stratégie purement analytique. Et inversement, qu'une telle stratégie ne soit pas adéquate ne montre pas que le domaine échappe à toute stratégie. En un sens, ce qui se joue ici n'est rien d'autre que le vieux débat autour du holisme, en particulier en sciences de la vie et en sciences sociales — débat qu'il n'est pas utile de restituer ici. Mais s'agissant de notre domaine, deux de ses caractéristiques, en se combinant, en font un objet d'un genre nouveau. La première est qu'il s'agit d'un système complexe, au sens semi-technique du terme : il est constitué de parties entre lesquelles existent des interactions multidirectionnelles qui empêchent de rendre compte de son évolution selon les schémas mécanistes classiques, car la contribution de chaque constituant dépend de la dynamique

globale du système ; il est ainsi le siège de phénomènes dits émergents qui ne peuvent être ni expliqués ni prédits par une analyse hiérarchique de ses constituants¹⁰.

Le domaine des technologies émergentes¹¹ est complexe en ce sens, de toute évidence. Mais il est aussi singulièrement étendu. Pour tenter de saisir ce trait, Timothy Morton parle d'hyperobjets, néologisme désignant des entités qui échappent aux catégories habituelles de l'espace et du temps¹² : tous nos efforts pour les localiser sont vains. Ses exemples sont, entre autres, le changement climatique, la pollution, Internet. Dans l'ignorance complète où je suis de la pensée de cet auteur, je ne lui emprunte que le terme, et je propose de considérer le domaine des technologies émergentes comme un hyperobjet, et donc comme un hyperobjet complexe.

En elle-même, cette décision terminologique ne semble pas faire avancer d'un iota notre problème. Pourtant, en le caractérisant comme un « hyper-problème », elle nous incite à le considérer sous un angle nouveau. Nous pouvons nous y préparer en reprenant l'un des exemples de T. Morton, le changement climatique. De cet hyperobjet, qui pose un hyper-problème, une myriade d'acteurs, individuels et collectifs, se saisit. Aucun d'entre eux ne peut prétendre en rassembler dans son enquête tous les aspects. Il existe cependant un organisme qui s'en approche, à savoir le

¹⁰ Il existe plusieurs définitions techniques d'un système complexe, correspondant à des concepts légèrement différents. De même et corrélativement, la notion d'émergence ne fait pas l'objet d'un consensus, mais elle apporte aux thèses holistes une clarification importante. Il importe en tout cas de distinguer complexité et complication : un système compliqué n'est pas nécessairement complexe, un système complexe n'est pas nécessairement compliqué, quoique beaucoup de systèmes naturels sont à la fois compliqués et complexes. Enfin, la complexité est affaire de degré : un système peut présenter plus ou moins de complexité, selon la structure des interactions, du nombre de niveaux d'émergence, de l'intensité des synergies en son sein. Dans l'abondante bibliothèque spécialisée, un ouvrage original et injustement méconnu permet de progresser dans la compréhension fine

de la question : Sunny Y. Auyang, (1999), *Foundations of Complex-system Theories: in Economics, Evolutionary Biology, and Statistical Physics*, Cambridge University Press. L'article de Wikipedia est également utile.

¹¹ C'est peut-être le moment d'admettre que l'adjectif « émergent » peut être compris ici dans les deux sens : initialement pris dans son acception temporelle (« déjà présent mais encore en phase de formation »), il peut aussi renvoyer au caractère complexe de l'ensemble.

¹² Timothy Morton (2013). *Hyperobjects: Philosophy and Ecology after the End of the World*, University of Minnesota Press. Trad. fr. *Hyperobjets. Philosophie et écologie après la fin du monde*, Cité du Design, 2018.

GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat). L'originalité du GIEC n'est pas suffisamment perçue, en partie à cause des controverses sur la question, mais aussi précisément parce que par son mode de constitution, son fonctionnement, sa méthodologie, il est sans précédent¹³. Imparfait comme toute oeuvre humaine, le GIEC m'apparaît comme la meilleure manière possible de relever le défi de l'hyper-problème¹⁴.

Ce que le GIEC met en oeuvre avec méthode, perfectionnant constamment ses pratiques, c'est l'intelligence collective. Laissant de côté cet exemple, demandons-nous ce que signifie cette expression et en quoi le concept peut nous être utile. Au premier niveau, l'intelligence collective désigne la capacité d'une communauté d'agents à acquérir et déployer collectivement un savoir, la compréhension d'un ensemble de phénomènes, la capacité à agir dans une situation. Cette capacité repose sur la division du travail : le problème entier dépassant les capacités d'un seul agent, une analyse préliminaire permet de le décomposer en sous-problèmes dont la résolution est à la portée d'un seul individu ou du moins d'une seule équipe. Ce schème, celui de la spécialisation, qui organise en particulier la recherche scientifique, est bien entendu nécessaire mais très insuffisant : la juxtaposition des solutions partielles obtenues par les individus ou les équipes spécialisées ne suffit généralement pas, sauf dans les cas les plus simples. Dans certains autres cas, on peut lier ensemble ces solutions au prix d'un travail complémentaire consistant à étudier, deux par deux, les articulations entre les différentes parties. L'interdisciplinarité dont on parle tant, à juste titre, se

ramène en général à ce genre de tâche. Mais lorsqu'on a affaire à un objet très complexe, et a fortiori à un hyperobjet complexe, une articulation exhaustive des descriptions ou modèles partiels est impossible. La tentation est alors de rechercher un point de vue totalisant, une *Aufhebung* qui par une opération miraculeuse livrerait une représentation complète de l'objet, ou du moins quelque chose comme son « sens profond », abstraction faite d'une masse de détails. À cette tentation cèdent nombre d'auteurs. Ce qu'ils produisent n'est pas sans intérêt, mais constitue un point de vue partiel de plus, à charge pour nous d'en trouver la place dans la mosaïque des représentations partielles.

Sommes-nous donc acculés à une limite indépassable de la connaissance humaine ? Non : c'est ici que la notion d'intelligence collective prend sa pleine portée. Si la cristallisation consciente dans un seul esprit, ou même dans les esprits interconnectés des membres d'une équipe ou d'une école de pensée, n'est pas possible dans certains cas, dont celui qui nous occupe, il est possible d'atteindre à une compréhension supérieure distribuée sur une population d'individus qui ne sont pas tous en relation entre eux. Les exemples sont légion. Aucune des grandes disciplines scientifiques ne donne lieu à autre chose qu'une intelligence collective de ce genre. La gestion d'une grande ville n'est à la portée d'aucun maire, d'aucun préfet, ni même d'aucune administration si l'on entend par là un organisme transparent à lui-même, au sein duquel le savoir circule entre toutes les parties. Pourtant, les mathématiques sont possibles et constituent un savoir bien réel, et les grandes villes sont bel et bien gérées et échappent, dans les meilleurs cas, au chaos.

¹³ Le GIEC est devenu à son tour un objet d'étude pour philosophes et sociologues des sciences. V. notamment l'article de D. Andler et al., "IPCC Assessment Reports as an epistemological puzzle" (soumis) et les références qu'il contient.

¹⁴ On sait que le GIEC est devenu un modèle que certains aimeraient appliquer à d'autres hyper-problèmes. Il semble par exemple que notre président de la République appelle de ses voeux un GIEC pour l'intelligence artificielle — qui n'est qu'une partie de l'hyperobjet que nous considérons ici.

Tout en illustrant les pouvoirs de l'intelligence collective, ces exemples sont encore relativement simples, parce qu'il est toujours en principe possible d'établir un pont entre deux parties du système ou deux représentations du phénomène, même si en pratique on y renonce souvent, soit parce que cela n'apporterait rien (il est heureusement souvent le cas qu'un aspect soit sans pertinence pour un autre), soit parce que le coût de l'enquête serait trop élevé, ou que personne ne soit disponible pour la mener, ou que sa conclusion arriverait trop tard. Ces cas semi-simples sont donc ceux dans lesquels une représentation unifiée est impossible, mais une représentation composite et en principe pleinement articulée est possible.

C'est à quoi il faut renoncer lorsqu'on a affaire à un domaine comme le nôtre. Il faut accepter que coexistent des représentations partielles que nous ne savons pas recoller, et qui peuvent même parfois s'exclure mutuellement. De plus, le terme de « représentation » est trop restrictif, dans la mesure où il semble renvoyer à un pur rapport de connaissance. Or les technologies émergentes impliquent d'autres rapports, relevant de l'action sous de multiples formes, et sont à la fois objet et sujet, puisqu'elles modifient les conditions de vie et les ressources des enquêteurs et jusqu'aux termes de l'enquête.

Ainsi, la bonne stratégie, la seule en réalité, consiste à mobiliser l'intelligence collective dans son extension la plus grande, c'est-à-dire mettre à contribution la plus grande diversité possible d'enquêteurs et d'agents. L'interdisciplinarité académique en est certes une composante, mais elle ne suffit pas. La diversité doit s'étendre aux formes d'activité, incluant la poursuite d'objectifs politiques et sociaux ; aux types d'organisations, incluant les comités d'éthique, les débats citoyens, les associations et les corps intermédiaires ; aux formes de pensée, incluant la prospective, le design, différents secteurs de la culture...

Tout cela exige que la réflexion, comme l'action, soient assises sur la meilleure compréhension possible des faits. Si évident que cela

paraisse, c'est là une condition qui détermine un objectif incontournable : faire comprendre. Dans le nécessaire partage des tâches, celle-là échoit plus particulièrement à des organismes comme le nôtre, à un projet tel que TESaCo. Encore faut-il se garder d'une illusion, selon laquelle les faits seraient strictement séparés des valeurs. Maintenir une distinction, tant que c'est possible, est une maxime louable, reprise rituellement dans les débats par les scientifiques soucieux de rappeler l'exigence d'objectivité et de neutralité axiologique : « Nous disons les faits, c'est à la société et ses représentants légitimes que revient de fixer les lignes d'action compte tenu des valeurs de cette société. » Cette maxime ne peut jamais être l'unique principe à l'oeuvre, et elle est particulièrement insuffisante lorsqu'on a affaire à l'objet hypercomplexe que constituent les technologies émergentes.

4. De l'intelligence collective à la sagesse collective

C'est que, contrairement au climat par exemple, ou encore aux pandémies et aux cancers, les technologies (quelles qu'elles soient du reste) sont entièrement entre nos mains, celles de l'humanité. Sans rouvrir l'éternel débat sur la responsabilité de l'inventeur quant aux usages possibles de son invention, il est surabondamment clair qu'au stade où nous en sommes, les technologies émergentes engagent de manière essentielle les valeurs. Que veux-je, pour moi-même, pour mes proches, pour la société ? Que veulent nos sociétés, pour leurs membres, leurs communautés, leur culture, leurs rapports mutuels ? Les technologies émergentes s'immiscent dans toutes ces questions, et inversement.

Or dès que les valeurs s'en mêlent, on quitte le domaine de la pure intelligence pour entrer dans celui de la sagesse. On l'avait déjà abordé, en réalité, en prônant à l'instant une contri-

bution à l'effort commun de groupes d'action. Mais on pouvait encore espérer échapper aux questions les plus difficiles, en se reposant sur une compréhension partagée du bien et du mal. Il est mal que les technologies émergentes aggravent les inégalités, il est bien qu'elles rendent possible des modes de participation démocratique plus efficaces. Il est bien qu'elles conduisent à des interventions thérapeutiques sur des désordres jusqu'ici incurables, il est mal qu'elles soient employées à conférer des avantages indus à des individus fortunés, fussent-ils malades, etc. Il est bien qu'elles permettent de retrouver des enfants perdus, il est mal qu'elles renforcent la surveillance policière de populations entières. Il est bien que le commerce soit stimulé, il est mal qu'en résulte un monopole, et ainsi de suite.

Mais les valeurs ont ceci de singulier qu'elles sont multiples, qu'elles entrent en conflit entre elles et avec la nécessité, qu'elles sont portées par des sujets distincts, qu'elles jouent sur différentes échelles temporelles, qu'elles ne sortent pas toujours toutes armées de nos codes, de nos coeurs ou de nos esprits, mais émergent ou se découvrent au cours des événements, et enfin qu'elles ne portent pas en elles-mêmes les instructions pour leur application aux cas particuliers qui se présentent dans la vie des sociétés et des individus.

C'est ici qu'intervient la sagesse pratique — en un sens proche de la *phronesis* aristotélicienne (souvent traduite par « prudence »), et qui est la capacité raisonnée d'agir en tenant compte de ce qui est bon et mauvais pour l'homme¹⁵. De manière plus précise, selon une définition largement acceptée dans les discussions contemporaines et qui ne prétend pas en subsumer toutes les acceptions, la sagesse pra-

tique consiste en la disposition à agir en sorte de satisfaire au mieux les valeurs, normes et objectifs des parties prenantes, c'est-à-dire en trouvant un équilibre entre ces exigences, compte tenu des faits présents et des conséquences à court, moyen et long terme, de son action. Maîtriser les nouvelles technologies, ce n'est rien d'autre que de rechercher les voies de la sagesse pratique, depuis leur conception jusqu'à leur mise au point, leur déploiement, leur régulation et leur usage, l'usage espéré ou envisagé faisant retour sur la conception, la mise au point, le déploiement, la régulation.

Comme y insistait déjà Aristote, la sagesse pratique ne s'obtient pas par le simple effort intellectuel : il ne suffit pas de réfléchir au bien et au mal en général. Il faut apprendre à agir de manière appropriée dans les circonstances vécues que l'on rencontre, ce qui nécessite de l'expérience, résultant en habitudes, en savoir-faire. Posséder la sagesse pratique est du même ordre, sur le plan moral, qu'être un artisan habile ou un bon violoniste.

Ce dont parlent Aristote et ses disciples contemporains est la sagesse comme disposition individuelle. Pour nous, la difficulté est que la maîtrise des technologies émergentes échappe à tout individu, fût-il capitaine d'industrie ou chef du gouvernement d'une grande puissance. Nous avons admis que comprendre les technologies émergentes ne pouvait être l'affaire que de l'intelligence collective. Mais c'est de sagesse pratique que nous avons besoin s'il s'agit de les maîtriser.

Une sagesse collective est-elle possible, et comment l'atteindre ou du moins s'en approcher ? Je ne sais pas s'il existe une réponse à cette question, prise de manière générale¹⁶. L'expérience nous donne quelques

¹⁵ La référence est le livre VI de l'Éthique à Nicomaque, particulièrement le §5. Ce qu'il faut entendre exactement par *phronesis* est l'objet de débats sans fin entre spécialistes d'Aristote. La définition moderne qui suit suffit à notre propos.

¹⁶ V. Landemore, H. & Elster, J., eds., *Collective Wisdom: Principles and Mechanisms*, Cambridge: Cambridge University Press, 2012; en particulier D. Andler, "What has collective wisdom to do with wisdom?", pp. 72-84.

raisons d'espérer : certains pays ont acquis une certaine sagesse en matière d'armes à feu, de circulation automobile, de sécurité alimentaire, de gouvernance démocratique, d'administration de la justice, de certification des faits. Le pessimiste n'aura aucun mal à contester ces exemples, et à trouver dans l'histoire et dans l'actualité des exemples de folie collective. Mais le désaccord porte en réalité sur les chances de succès : saurons-nous collectivement devenir sages ? Sur ce qu'il convient de faire, en revanche, le doute n'est pas permis. Il convient, au premier sens du terme, de mobiliser toutes nos ressources pour orienter les technologies émergentes dans la direction que nous jugeons bonne et les détourner des directions que nous jugeons néfastes. Le « nous » désigne, selon le contexte, l'une ou l'autre des communautés auxquelles chacun se rattache : n'espérons pas une prescription universelle. Au second sens, il convient de mettre en oeuvre les méthodes que nous connaissons et avons commencé à explorer, dans les nombreux groupes de réflexion, d'action, de régulation mentionnés au début de cet exposé. Si particulières qu'elles soient, on y a suffisamment insisté, les technologies émergentes donnent prise à des approches éprouvées au cours de l'histoire du développement technologique, industriel, économique et social.

*

En proposant le cycle d'étude « Technologies émergentes et sagesse collective », l'Académie des sciences morales et politiques n'a d'autre prétention que de se joindre à cet immense effort, en restant, avec sagesse, consciente des limites du projet.

Elle ne compte pas y travailler seule. Pour une bonne part, ses travaux seront menés en collaboration avec d'autres académies,

des équipes de recherche, des instituts de prospective, des associations, en France et à l'étranger. Des colloques seront organisés avec leur participation. Des groupes thématiques seront mis sur pied, composés d'académiciens, de chercheurs et d'experts extérieurs. Des études seront produites, et une documentation mise à disposition sur le site du projet. Les activités se partageront entre des études sectorielles et des approches globales, se nourrissant mutuellement. Une ambition du projet est de tisser ensemble un certain nombre de thèmes, en recherchant les similitudes entre différents secteurs, et en transférant des méthodes, concepts, aperçus utiles d'un secteur à l'autre. Nous espérons ainsi préciser la vue d'ensemble de notre hyperobjet. Enfin, chaque fois que cela sera possible, nous traduirons nos acquis en éléments de pédagogie, en nous appuyant, ici encore, sur le savoir-faire d'équipes compétentes.

Les interventions que nous allons entendre au cours des deux journées qui s'ouvrent donneront une idée plus précise de l'entreprise. Elles ne couvriront pas toute l'extension de notre hyperobjet, mais en dévoileront certaines régions particulièrement importantes, et ouvriront sans doute des pistes pour en explorer des aspects moins visibles. Ainsi commence ce qui sera, nous l'espérons, une aventure intellectuelle, morale et politique, menée de commun.