



Les Cahiers de TESaCo n°1

GÉNÉTIQUE, BIOTECHNOLOGIES / NUMÉRISPHÈRE, RÉSEAUX,
DATA / SCIENCES COGNITIVES, NEUROSCIENCES /
IA, ROBOTIQUE / NANOTECHNOLOGIES

Technologies émergentes et sagesse collective
Comprendre, faire comprendre, maîtriser



ACADÉMIE DES SCIENCES
MORALES ET POLITIQUES
INSTITUT DE FRANCE



FONDATION
SIMONE ET CINO
DEL DUCA
INSTITUT DE FRANCE

TESACO

En l'espace de deux décennies, les technologies dites émergentes — biotechnologies, technologies de l'information et de la communication, technologies issues des neurosciences cognitives, nanotechnologies... — ont profondément modifié les conditions d'existence à l'échelle planétaire et affecté tous les secteurs d'activité humaine. Porteuses de solutions mais aussi de menaces pour nos équilibres fondamentaux, ces nouvelles technologies sont devenues si puissantes qu'on ne sait comment en reprendre le contrôle, alors même qu'elles continuent de se développer, ouvrant la voie à des conséquences et à des risques imprévisibles.

Cet état de fait appelle un effort pour mieux comprendre les technologies et leurs effets, informer le public et les responsables politiques, et proposer des dispositifs pouvant contribuer à maîtriser l'évolution en cours.

L'Académie des sciences morales et politiques a souhaité participer à cet effort, et avec l'appui de la Fondation Simone et Cino Del Duca elle a lancé en 2019 une enquête sur « Technologies émergentes et sagesse collective » (TESaCo). Un colloque tenu à l'Institut en janvier 2020 a marqué la fin des travaux préparatoires. Le présent numéro des Cahiers de TESaCo, qui est aussi le premier, est consacré à un compte rendu presque exhaustif de cette manifestation.

LES CAHIERS DE TESACO

Les Cahiers de TESaCo est une publication périodique qui présente les travaux de l'équipe du projet, organisée en six groupes de travail thématiques : biotechnologies, intelligence artificielle et robotique, sciences cognitives appliquées, éthique-droit-libertés, numérisphère, anthropologie numérique.

REMERCIEMENTS

Nous remercions Véronique Duchaud-Fuselli et Marianne Tomi de l'Académie des sciences morales et politiques ainsi qu'Ariane Bigenwald, Julien Benesteau et Mathias Dambuyant pour leur aide et leur soutien.

COMITÉ ÉDITORIAL

Daniel Andler, responsable du projet TESaCo

Margaux Berrettoni, chargée de coordination et de communication

Serena Ciranna, assistante de recherche

Juliette Cornet, design & graphisme

AVANT-PROPOS

GABRIEL DE BROGLIE

Gabriel de Broglie a été successivement directeur-général adjoint de l'ORTF, directeur général de Radio-France, président de l'INA (Institut national de l'audiovisuel). Historien et essayiste, il participe également, depuis 1981, aux différentes instances de la langue française, Haut Comité (1981-1982), Conseil supérieur (1984, 1986, et depuis 1999) et présidence de la Commission générale de terminologie et de néologie (1996-2006). Il est membre de l'Académie des sciences morales et politiques et de l'Académie française. Il a été Chancelier de l'Institut de France de 2006 à 2017. Il est Chancelier honoraire de l'Institut de France et Président de la Fondation Simone et Cino Del Duca.

En 1832, Louis-Philippe se laissait convaincre par Guizot de ressusciter la 3e classe de l'Institut dite des Sciences morales et politiques, créée par la Convention dans l'esprit des philosophes des Lumières, en ajoutant : " Cette académie sera très utile pourvu qu'elle ne soit pas trop pressée de se faire écouter et qu'ailleurs on ne fasse pas trop de bruit". L'une des modalités qu'adopta la nouvelle académie fut les enquêtes sociales fondées sur les statistiques, dont le modèle reste la grande enquête de Villermé publiée en 1840 sur les conséquences de l'industrialisation sur les conditions de travail et de vie des ouvriers. Menée dans un esprit scientifique, positiviste et sociologique, elle eut un effet d'entraînement pour d'autres enquêtes et pour l'évolution de la législation.

En 2007, je constatais que l'Académie des sciences morales et politiques manquait de moyens pour entreprendre des enquêtes et qu'elle ne bénéficiait d'aucun concours de la Fondation Del Duca. Un cycle de grandes enquêtes fut entrepris au sein de l'Académie, et a été illustré par Michel Albert et Raymond Barre : La France prépare mal l'avenir de sa jeunesse, par Raymond Boudon : Justice et égalité, par Jean Baechler : La Guerre, par Chantal Delsol : Intellectuels et penseurs en Europe centrale et orientale depuis 1945.

Le flambeau est maintenant repris par Daniel Andler pour une vaste enquête sur les technologies émergentes.

Il y a plusieurs décennies, le Haut Conseil de l'Audiovisuel multipliait les études sur les technologies de l'information et de la communication, les TIC. Mais composé de spécialistes et de professionnels de l'audiovisuel, et limité par sa mission, il n'a pas débordé sur d'autres champs de recherche.

Depuis quelques années, les sujets, outre les TIC, de la numérisation, du big data, des algorithmes, de l'intelligence artificielle et des nouvelles connaissances qu'ils entraînent sont en discussion dans tous les forums. L'Académie des sciences morales et politiques a créé en son sein un très vivant groupe de travail Humanisme et mondialisation qui a lui-même abordé ces sujets et bien d'autres, organisé nombre d'auditions, de conférences, de colloques, un cycle d'entretiens des académiciens à l'Auditorium de l'Institut, auxquels Daniel Andler a activement participé. Puis il a proposé de prendre la succession de Chantal Delsol pour une enquête de trois années sur les nouvelles technologies qu'il a rebaptisé technologies émergentes.

Le projet est ambitieux, original et novateur. Son thème est l'objet constant de descriptions alarmantes ou optimistes, et d'anticipations redoutables mais imprécises et incertaines. L'enquête dévoilée dans ces pages s'étend à la globalité du phénomène et s'efforce de dégager les responsabilités et les conditions de résultats favorables. D'où son titre : Technologies émergentes et sagesse collective — TESaCo.

Dans les mois qui ont suivi, le colloque d'ouverture dont on trouve ici les actes s'est tenu les 30 et 31 janvier 2020, associant seize personnalités remarquables, hautement compétentes dans les nombreux domaines abordés.

Il s'agit d'un état des lieux déjà enrichi. Les technologies émergentes sont inventoriées. Leurs caractères émergent, convergent, parfois contradictoire, accéléré et disruptif sont mis en valeur ; mais les incertitudes sur les accélérations, les enchaînements et les responsabilités sont soulignées. La vivacité des échanges montre bien que les enjeux sont considérables, que la tâche ne peut être que collective et que les conclusions attendues dépassent une énumération de propositions, qu'elles sont du ressort de l'information, de la compréhension, de la formation, de l'exhortation et de la conscience.

TESaCo n'est aussi pas seulement un colloque. C'est un programme établi par Daniel

Andler comportant un site internet, des publications, des séminaires, des cahiers trimestriels, la tenue de deux colloques de bilan et la publication d'un ouvrage collectif sous le timbre de l'Académie des sciences morales et politiques.

La Fondation Del Duca est heureuse de concourir à une mission historique de l'Académie des sciences morales et politiques, sur une transformation profonde qui n'est pas seulement technologique, mais qui est déjà et se trouvera davantage encore dans l'avenir au centre de l'évolution de notre société et de l'humanité dans son ensemble.

SOMMAIRE

<i>Avant-propos</i>	5
Gabriel de Broglie	
<i>Introduction</i>	9
Serena Ciranna	
<i>« Technologies émergentes et sagesse collective. Comprendre, faire comprendre, maîtriser » Un vaste programme de plus ?</i>	13
Daniel Andler	
Première partie. Numérisphère, data et société	27
1-1 <i>La sagesse collective comme moyen d'appropriation politique des technologies numériques</i>	29
Axelle Lemaire	
1-2 <i>Interactions Humain-Machine : quel cadre juridique penser pour demain ?</i>	35
Célia Zolynski	
1-3 <i>Serons-nous suffisamment sages pour continuer à protéger nos libertés ?</i>	41
Régis Chatellier	
1-4 <i>De la toxicité de certains termes associés à «technologie» : impact, éthique, problème, solution, usage...</i>	47
Daniel Kaplan	
Deuxième partie. Génétique et biotechnologies	53
2-1 <i>GATTACA : Sommes-nous prêts ?</i>	55
Alex Peluffo	
2-2 <i>Le forçage génétique ou « gene drive » : état des lieux et enjeux associés à cette nouvelle biotechnologie.</i>	63
Virginie Courtier-Orgogozo	

Troisième partie. Intelligence artificielle, robotique et nanotechnologies	69
3-1 <i>Qu'est-ce qu'une intelligence artificielle « réellement éthique » ?</i> Antonio Casilli	71
3-2 <i>Robotique : l'intelligence de la gravité.</i> Jean-Paul Laumond	79
3-3 <i>Traiter les maladies graves à l'aide de nanomédicaments : pour un Homme réparé mais pas augmenté.</i> Patrick Couvreur	87
 <i>Table ronde et débat</i>	95
Yannick Blanc, Alain Bravo, Nicolas Chaillet, Élisabeth Grosdhomme. Modérateur : Florian Forestier	
 <i>Conclusions</i>	125
Pierre Delvolvé	

INTRODUCTION

À PROPOS DU CAHIER N°1

SERENA CIRANNA

Serena Ciranna est philosophe et termine actuellement une thèse de doctorat sur les effets des médias numériques sur l'identité narrative. Elle est assistante de recherche pour TESaCo.

Ce premier Cahier rassemble l'essentiel des communications du colloque de lancement d'une enquête « Technologies émergentes et sagesse collective » (TESaCo)¹, qui a eu lieu à Paris, à l'Institut de France les 30 et 31 janvier 2020 sous le titre « Technologies émergentes : comprendre, faire comprendre, maîtriser. Les conditions d'une sagesse collective ».

L'objet de l'enquête est constitué par l'ensemble des technologies dites « émergentes », en raison de leur développement extrêmement rapide : elles se caractérisent, comme le souligne Daniel Andler dans son exposé liminaire, par « 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 ». Il s'agit des technologies de l'information et de la communication (TIC), de l'intelligence artificielle et de la robotique ; des biotechnologies ; des technologies dérivées des sciences cognitives et des neurosciences ; des nanosciences. Ces domaines en plein essor appellent une adaptation en continu de la part tant des usagers que des experts et des décideurs, qu'il s'agisse de l'adoption de nouveaux outils, de l'évaluation de leur fiabilité, de la prévision de leurs effets sur le long terme, des conditions économiques, industrielles et politiques de leur développement.

L'objectif du colloque était de présenter certaines des principales avancées dans les différents secteurs. Les intervenants ont su mettre en évidence les liens précis qui s'établissent

¹ C'est une enquête née en 2019 sous l'égide de l'Académie des sciences morales et politiques, avec le soutien de la Fondation Simone et Cino Del Duca.

entre les bases scientifiques et techniques des programmes en cours et les questions économiques, politiques et éthiques qu'ils soulèvent. Ils ont également illustré les effets de convergence, qui amplifient considérablement la puissance des nouveaux dispositifs. Comme le précise encore Daniel Andler : « Les différentes technologies se nourrissent les unes des autres, et leur effet conjoint dépasse celui qu'elles peuvent avoir séparément ».

C'est le cas, par exemple, de la robotique dont Jean-Paul Laumond (« Robotique : l'intelligence de la gravité ») retrace les origines dans la convergence entre la cybernétique et l'intelligence artificielle et dans le couplage des machines à calculer avec les machines tout court. Un effet de convergence est également à la base de la montée en puissance des nanotechnologies et de la génomique, avec la présence toujours plus marquée de l'informatique dans les laboratoires de biologie, mais également avec la possibilité inédite de communication à distance entre les chercheurs et de partage des données et des résultats. L'informatisation de la biologie ouvre le champ à une nouvelle gamme de possibilités et de problèmes, de la personnalisation du traitement à la tutelle des données des patients.

Les questions que ces technologies suscitent touchent à un grand nombre de domaines et de dimensions, qui font appel à autant de champs d'expertise spécialisés. C'est ainsi que la possibilité de manipulation du vivant pose de nouveaux défis éthiques qui interpellent les biologistes, les citoyens, les pouvoirs publics, les entreprises, les organisations internationales. C'est le cas, par exemple, des nombreuses applications potentielles du forçage génétique aussi bien en santé publique qu'en agriculture et en matière de conservation. Les enjeux et risques qui y sont associés sont mis en évidence par Virginie Courtier (« Le forçage

génétique ou "gene drive" : état des lieux et enjeux associés à cette nouvelle biotechnologie »). De même, comme le montre bien Alex Peluffo (« GATTACA : Sommes-nous prêts ? »), la génétique promet de révolutionner, notamment, rien de moins que la santé, la justice et le sport. Comment concilier les perspectives ouvertes par ces avancées considérables avec le respect des valeurs fondamentales que nous voulons préserver dans la société de demain, dès lors qu'elle serait largement modifiée sur le plan social, politique et économique par les techniques de génomique ? Mais du reste, de quelles valeurs parlons-nous ? Une question centrale formulée par André Choulika dans son intervention (« L'édition du génome ou l'aube d'une révolution en marche »²) est la suivante. Face à la séduction d'un programme visant à permettre à l'homme de « prendre le contrôle sur la vie », porté par une science de plus en plus globalisée, comment faire droit à la variété des cultures, des traditions et des législations ?

Non moins « disruptives » que les techniques d'édition du génome, les nanotechnologies offrent aujourd'hui de formidables possibilités thérapeutiques. Mais elles posent également la question des limites entre réparation et augmentation de l'humain, comme le souligne Patrick Couvreur (« Traiter les maladies graves à l'aide de nanomédicaments: pour un Homme réparé mais pas augmenté »). L'utilisation des nanomédicaments en médecine humaine vise exclusivement à guérir le patient mais elle pourrait aussi déboucher sur des projets visant à augmenter certaines capacités humaines, physiques ou cognitives, en dehors des champs habituels de la médecine et de la pharmacie. Telle l'introduction d'une partie du génome modifiée dans une population, l'introduction d'humains « augmentés » dans notre société pourrait en modifier les équilibres et avoir également un effet sur le rapport de l'in-

² Dont il n'a malheureusement pas été possible de faire figurer la version écrite dans le présent numéro.

dividu à soi-même. N'est-ce pas l'identité humaine elle-même qui est mise en péril, comme le suggère Patrick Couvreur, lorsqu'on fait reculer les limites d'un être humain en l'affranchissant de ses faiblesses ?

Les technologies émergentes mettent en question également, et ce dès aujourd'hui, les relations établies entre les différents acteurs au sein de la société : employeurs et travailleurs, producteurs, consommateurs, usagers, responsables politiques et citoyens. Des notions et principes juridiques tels la responsabilité, le sujet, la propriété doivent être alors constamment repensés. Certes, comme le rappelle Régis Chatellier (« Serons-nous suffisamment sages pour continuer à protéger nos libertés ? ») au cours des quarante dernières années la France puis l'Europe se sont dotées d'un cadre juridique pour s'assurer que l'informatique ne porte atteinte « ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques » (article 1, loi Informatique & Libertés). Il n'en demeure pas moins que la numérisation de nos vies quotidiennes fait constamment surgir de nouveaux enjeux pour les libertés et notre capacité à vivre en société. Parmi les phénomènes émergents qui requièrent une vigilance renouvelée : les tests ADN dits récréatifs, les assistants vocaux, les villes numériques. Les législateurs sont appelés à repenser le cadre juridique de ces technologies en évolution et des nouvelles formes d'interaction humain-machine. Particulièrement préoccupants sont les outils qui permettent l'exploitation de l'attention de l'utilisateur et la manipulation de ses émotions. Célia Zolynski (« Interactions Humain-Machine : quel cadre juridique penser pour demain ? ») met en garde contre les menaces que ces dispositifs font peser sur l'autonomie, et de manière générale elle souligne la nécessité de repenser la place de l'humain pour lui reconnaître la

qualité d'agent (et non de simple consommateur) de l'intelligence artificielle, condition de sa liberté.

L'intelligence artificielle suscite de nombreuses interrogations. Gilles Dowek (« L'évolution de la notion de responsabilité au temps des algorithmes »³) soulève celle de la responsabilité. Il rappelle que selon le code de la route, le conducteur d'un véhicule est responsable des infractions commises par lui dans la conduite dudit véhicule. Mais que se passe-t-il quand ce conducteur est un algorithme? Au-delà de l'opportunité de donner une personnalité juridique aux algorithmes, cette question nous amène à nous interroger sur la nature et le rôle de la sanction, quand une décision est prise par un algorithme. De son côté, Antonio Casilli (« Qu'est-ce qu'une intelligence artificielle « réellement éthique » ? ») s'interroge sur l'étonnante prolifération, ces dernières années, de chartes qui tentent de dégager les principes d'une IA éthique. Il observe que celles-ci se limitent généralement à l'usage des technologies, laissant dans l'ombre leurs moyens de production. Il est alors nécessaire de dévoiler l'impensé de l'IA et de ramener l'attention sur le rôle du travail humain nécessaire pour la conception et la fabrication des technologies intelligentes.

Une autre question centrale est celle du rôle des décideurs — responsables politiques et institutionnels, acteurs de la société civile, industriels... Comment doivent-ils se rapporter aux technologies émergentes ? Comment penser la dimension politique de ces technologies ? Ses fonctions au gouvernement ont amené Axelle Lemaire à se pencher sur cette question dans le cas du numérique. Dans son intervention (« La sagesse collective comme moyen d'appropriation politique des technologies numériques ? ») elle partage les acquis de son engagement dans une expérience inédite de

³ Dont il n'a malheureusement pas été possible de faire figurer la version écrite dans le présent numéro.

participation citoyenne à l'écriture du texte de la Loi pour une République numérique. C'est là un exemple concret dont le succès montre qu'un déploiement démocratique des technologies numériques est possible.

Faire débattre les experts, entre eux et avec les acteurs du domaine et les citoyens, relève de l'objectif principal de TESaCo : promouvoir une sagesse collective dans notre rapport aux technologies émergentes. La « sagesse » en question est d'ordre pratique, mais va au-delà d'une simple (quoique difficile) mise en adéquation des moyens aux fins. Il ne s'agit pas de trouver les bonnes solutions, ou les moins mauvaises, mais de poser la question des fins, des valeurs en jeu dans l'usage et le développement des technologies émergentes, prenant en compte les points de vue de toutes les parties prenantes. Dans cette optique, il faut aussi remettre en question, d'un point de vue philosophique et anthropologique, la manière dont nous pensons et construisons sémantiquement notre rapport aux technologies et à la technologie. Ce qui implique, comme le montre Daniel Kaplan (« De la toxicité de certains termes associés à «technologie» : impact, éthique, problème, solution, usage ») de soumettre notre vocabulaire à la critique et de questionner certains termes que nous employons couramment lorsque nous parlons de technologie(s).

La table ronde qui conclut le présent recueil reflète l'esprit prospectif du colloque et de TESaCO même. Dans sa composition, d'abord, puisqu'elle, réunissait Yannick Blanc, Alain Bravo, Nicolas Chaillet, Elisabeth Grosdhomme, quatre personnalités venant d'horizons divers — haute fonction publique, société civile, industrie, université —, souvent tout cela à la fois. Dans son contenu, puisque, comme l'indiquait Florian Forestier, modérateur de la

discussion, «il n'est pas question d'apporter des réponses toutes faites ou faire des exposés ex cathedra, mais bien de faire émerger des questions, des problématiques, de commencer à élaborer un langage commun, bref de construire les conditions d'une intelligence collective pour examiner les phénomènes multifactoriels dont TESaCo traitera tout au long de son enquête. »

Ce Cahier restitue les résultats d'une première étape de cette enquête. Nous espérons qu'il intéressera un public varié, présentant un panorama riche, à défaut d'être complet, du nouveau monde technologique et des questions qui l'accompagnent. Peut-être apportera-t-il même aux spécialistes une perspective complémentaire, par-delà leur champ d'expertise. Nous accueillerons en tout cas avec reconnaissance toute remarque ou suggestion susceptible de nous faire progresser dans la difficile recherche de la sagesse collective.⁴

⁴ Vous pouvez nous contacter à l'adresse courriel suivante : tesaco@asmp.fr ou via notre site internet <https://www.tesaco.fr/>

« 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

paraît, 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.

Première partie

NUMÉRISPHERE, DATA ET SOCIÉTÉ

La sagesse collective comme moyen d'appropriation politique des technologies numériques

AXELLE LEMAIRE

Juriste internationale de formation, Axelle Lemaire est associée au sein du cabinet européen de conseil en stratégie Roland Berger, où elle y dirige l'écosystème d'innovation Terra Numerata. Secrétaire d'Etat en charge du Numérique et de l'Innovation de 2014 à 2017, elle a accompagné l'envol de la French Tech, permis l'ouverture des données publiques, amplifié le Plan France Très Haut Débit, lancé la première stratégie nationale d'intelligence artificielle. Elle a défendu au Parlement la loi pour une République numérique, co-construite pour la première fois avec les citoyens et adoptée à l'unanimité. Elle a obtenu le Prix européen de l'Innovation politique en 2017 et le Women in Business Award en 2018.

1. La part du politique dans le numérique

L'actualité n'a de cesse de nous rappeler le rôle croissant joué par les technologies numériques dans la vie quotidienne des individus, des entreprises et des organisations publiques : intelligence artificielle, antennes mobiles dont la 5G, reconnaissance faciale, application mobile de suivi des cas contact pendant la crise sanitaire, caméras de surveillance, usage des téléphone portables pour filmer la police, télétravail... Pourtant, il n'existe aucune doctrine homogène des pouvoirs publics en France quant à la manière d'introduire ces technologies dans l'espace public et privé : l'appropriation se fait uniquement à l'usage, par le rejet ou le succès populaire, et par des encadrements réglementaires spécifiques construits en dehors de tout débat démocratique et de processus de débat public. Cette manière de gouverner les technologies par les seules approches d'efficacité, de sécurité, de consommation et/ou de débouché industriel, sans s'interroger par exemple sur le respect des libertés publiques, la prévention des discriminations, l'impact environnemental ou encore la nécessité des usages, dilue la mise en perspective des objectifs poursuivis et renforce la dépolitisation des enjeux sous-jacents à l'utilisation des technologies. Doit-on se satisfaire de cet état de fait, qui ne contribue pas à la paix sociale et peut alimenter des formes d'angoisse individuelle et collective, ou est-il envisageable de développer des moyens d'appropriation politique autres ? La sagesse collective peut-elle jouer un rôle dans une telle appropriation ?

Cette intervention se propose d'offrir un té-

moignage de la réalité de l'action politique en ce domaine, pour la confronter aux questionnements plus théoriques des chercheurs.

L'appropriation des technologies par la sagesse collective passe forcément par le questionnement sur la possibilité pour le numérique d'être considéré comme un objet politique démocratique. Je pars du postulat que l'une des conditions de la bonne intégration des technologies dans la société - avec comme finalité poursuivie le progrès humain pour tous - est que celle-ci se fasse par la société c'est-à-dire selon des processus d'élaboration démocratique. Si le numérique constitue de facto un objet politique, la conscience de ce caractère est assez récente. On a longtemps considéré que les technologies appartenaient au seul domaine de la science et qu'il valait mieux séparer ce champ de celui des décisions politiques, suivant l'idée que celles-ci relèvent plutôt de l'arbitraire et du discrétionnaire. Or le numérique n'est en rien un objet neutre et apolitique - il est plus que jamais un objet politique, dans sa capacité d'impact, dans ses conséquences économiques et la vision sociétale qu'il sous-tend. Si les institutions démocratiques ont toujours du mal à voir le numérique comme un objet politique à part entière et à le traiter en tant que tel, les régimes totalitaires ou autocratiques s'en servent pour manipuler les contenus ou les contenants et préserver le pouvoir en place. La question est donc de savoir comment le numérique peut se mettre au service d'une ambition démocratique.

En tant que membre du gouvernement en charge du sujet numérique entre 2014 et 2017, j'ai vécu cet écart de perception - objet politique et objet considéré comme « neutre » - à de multiples reprises. Un exemple concret nous est fourni par les discussions auxquelles il m'a été donné de participer au sein de l'ICANN (Internet Corporation for Assigned Names and Numbers), l'instance internationale qui attribue les noms des domaines de premier niveau sur le Web. L'ICANN est, par exemple, à l'origine

de l'autorisation attribuée à l'État français pour utiliser le domaine “.fr”. Cet organisme était à l'origine constitué en société commerciale anonyme, domiciliée en Californie, l'ICANN produisant des revenus à travers la vente des noms de domaines par des mises aux enchères. Les décisions sont prises par un collège dit « multipartite », composé des représentants des États, des entreprises privées, des consommateurs et des experts techniques. Or, même si les décisions à l'ordre du jour dans le cadre des discussions de l'ICANN ont souvent un caractère politique, celui-ci n'est pas assumé explicitement : la catégorie du politique n'y a pas sa place. Or, comment ignorer par exemple la portée politique de la décision de réserver unilatéralement et dans le plus grand secret le domaine « .army » à l'armée américaine, ou que le domaine « .amazon » puisse être acheté par le géant américain du commerce en ligne malgré les contestations du Brésil ? La question de la transparence dans la gouvernance de l'ICANN, sous-jacente à celle du caractère démocratique de l'institution et de sa prise en compte de l'intérêt général, s'est posée avec tout autant d'acuité : qui devrait donc se charger de décisions qui ne relèvent pas seulement de la technologie mais aussi du politique ? Le système en place à l'ICANN se base sur des décisions par consensus, avec une capacité de blocage minoritaire. Certaines résolutions aux enjeux plus politiques ont souvent été bloquées par le même groupe d'États, à savoir les États-Unis, le Royaume Uni, le Canada, l'Australie et la Nouvelle Zélande (ceux qui composent le groupe des “Five Eyes” dans le monde du renseignement). Cette remarque pourrait nous amener à tirer la conclusion que la capacité à débrancher certains noms de domaines est très étroitement liée à des enjeux de souveraineté, de sécurité et de maîtrise de l'internet. Tout en étant face à des sujets éminemment politiques, des décisions ont été prises selon des processus opaques sous prétexte de leur caractère technique et exclusivement commercial. Depuis quelques années, sous la pression de certains gouvernements, la gouvernance de l'ICANN

a néanmoins commencé à évoluer. C'est ainsi qu'en France, des viticulteurs indépendants se sont mobilisés pour s'opposer à la vente aux enchères des noms de domaines “.vin” et “.wine”, qui aurait permis d'ignorer les appellations d'origine. La gouvernance de l'ICANN a été alors revue et des processus plus démocratiques mis en place.

L'ICANN est devenue une organisation internationale avec son siège à Genève, et en 2016 le président Barack Obama a décidé de couper définitivement le lien organique entre l'organisme et le secrétariat d'État au Commerce américain. Il ne s'agit là que d'une illustration parmi beaucoup d'autres du caractère éminemment politique des choix relatifs à l'utilisation d'internet et des technologies numériques, qui sont pourtant réservées aux seuls experts. Ce décalage entre la réalité politique des technologies et la compréhension de cette réalité par les décideurs politiques eux-mêmes peut s'expliquer par des formes diverses d'ignorance ou de négligence, menant à une délégation de la responsabilité décisionnelle aux « sachants ». Pendant ce temps, on observe une défiance et une contestation croissantes de la légitimité politique des responsables publics dans les Etats démocratiques, qui peinent à établir un dialogue de qualité entre les citoyens et leurs représentants.

2. Une expérience démocratique numérique

Le numérique ne pourrait-il pas contribuer à construire un dialogue démocratique de plus grande qualité ? C'est la promesse des « civic tech », les technologies mises au service du renouveau démocratique. Et l'objectif posé lors de l'expérience lancée, en 2016, pour la préparation de la loi pour une République numérique. Nous avons proposé l'écriture d'un texte législatif émanant de l'exécutif avec les citoyens, de manière transparente, en lançant un processus de concertation en ligne, dont le

but n'était pas uniquement de consulter, mais de permettre les conditions d'une participation citoyenne active (par des commentaires, propositions d'amendements, suggestions de nouveaux articles) pour rédiger le texte avant que celui-ci n'arrive en examen au Parlement. La consultation a duré trois semaines et le défi a notamment consisté à faire rentrer cette manière de faire dans le processus démocratique “normal” ou classique, très structuré, rythmé par les arbitrages interministériels, le contrôle par le Conseil d'Etat et la navette parlementaire. L'avant-projet de loi ainsi partagé en amont a attiré 23 000 contributeurs et plus de 140 000 votes sur Facebook. A l'issue de cette procédure, cinq nouveaux articles ont été insérés dans le texte, devenu par la suite l'un des rares exemples sous la cinquième République de loi adoptée à la quasi-unanimité à l'Assemblée Nationale et au Sénat. En observant ces résultats positifs, l'on pourrait voir dans le numérique une promesse démocratique forte. En effet, la transparence du processus a ultimement amélioré la qualité des débats parlementaires, car les nœuds, les polémiques et contentieux possibles, ont été identifiés très en amont, travaillés et discutés en ligne et dans les circonscriptions des parlementaires, avant d'arriver au Parlement. La légitimité même du texte de loi s'est vue renforcée par le soutien de communautés qui avaient participé à son élaboration, aussi diverses que des associations représentatives de personne en situation de handicap ou des jeunes « gamers » prenant part à des compétitions en ligne de jeux vidéo. Or, si l'on compare cet exercice à d'autres, en apparence similaire, comme celui du Grand débat national, on se rend compte qu'il existe certaines conditions d'un exercice véritablement démocratique et transparent des outils technologiques, et qu'il ne suffit pas d'adopter les mêmes plateformes de consultation pour garantir le même résultat. Des initiatives de démocratie numérique qui ne respectent pas ces conditions peuvent se révéler contre-productives en générant beaucoup d'attentes et des frustrations chez les citoyens.

3. Les conditions d'un exercice démocratique du numérique

Deux conditions se posent alors pour que ces processus de concertation citoyenne puissent améliorer le dialogue démocratique. D'abord des garanties de transparence, qui se traduisent concrètement par un paramétrage de la plateforme numérique utilisée qui permette de limiter au maximum la possibilité de contrôler les contenus et écarte les suspicions de manipulation, autorise les échanges directs et itératifs entre contributeurs, et accueille tous les types d'avis, négatifs comme positifs. Pour créer une participation constructive et génératrice d'idées nouvelles, il est en effet important de permettre aux participants de commenter les contributions de leurs pairs et de créer des « dialogues parallèles » qui nourrissent véritablement l'intelligence collective, dans un exercice qui ne se veut pas directif mais encadré. Aussi, lorsqu'un propos à teneur illicite ou déplacée est détecté, le processus de suppression doit-il rester transparent et participatif. La censure de pair à pair par une « corbeille ouverte » crée de la confiance et augmente le niveau des attentes vis-à-vis de la qualité des contributions. Cela suppose un « lâcher prise » de la part des gouvernants et des administrations, et l'expression d'une confiance à l'égard des administrés, qui n'est pas souvent intégrée dans les processus décisionnels internes et peut constituer un obstacle culturel important. La deuxième condition à la réussite d'une expérience démocratique numérique est que la « promesse » initiale soit clairement énoncée, quitte à rester modeste, et que celle-ci soit ensuite tenue, là aussi en toute transparence. Dans le cas de la concertation sur la loi pour une République numérique, la finalité « propositionnelle » de l'exercice a été explicitée : il ne s'agissait pas d'introduire un exercice de démocratie directe ou de référendum, à rebours des attributions du parlement, mais d'ouvrir le processus d'écriture de la loi à l'ensemble des parties intéressées, particuliers

inclus. Les participants étaient prévenus que le vote massif de certaines propositions ne garantissait pas leur adoption directe par le gouvernement, mais uniquement un examen attentif. Il fallait donc établir cette règle dès le départ et la rendre très claire pour les participants. En revanche, le gouvernement s'engageait à un devoir de réponse, en expliquant les raisons de la position finalement adoptée par rapport à une proposition citoyenne. En l'occurrence, les administrations ont été fortement mobilisées pour préparer des réponses écrites à plus de deux cents propositions, celles qui avaient reçu le plus de votes sur la plateforme.

4. Conclusion

Le numérique doit être pensé comme objet politique pesant dans les choix démocratiques, et comme outil politique de renouveau du dialogue démocratique et des fonctionnements institutionnels. Néanmoins, une autre réalité s'impose aussi aux responsables politiques : celle des dérives néfastes observées sur les réseaux sociaux, liées à la prolifération de la désinformation, des activités illicites et des contenus haineux. Vis-à-vis de ces phénomènes, la politique se considère comme démunie et se tourne vers les entreprises technologiques pour demander une gouvernance et une régulation que les États ne sont pas en mesure d'offrir. Ces dérives offrent aussi l'occasion de renforcer des modes de gouvernance plus autoritaires et moins démocratiques. La promesse initiale du numérique au service de la démocratie n'est donc pas encore tenue aujourd'hui. Au contraire, la perception dominante chez les responsables politiques semble être celle d'un numérique qui les éloigne plus encore des citoyens qu'ils doivent représenter. La sagesse collective pourrait trouver à s'exprimer par les technologies numériques, mais à la condition d'entourer son expression par une doctrine, des principes, des modes de gouvernance qui aujourd'hui ne sont pas encore assez bien définis. Un grand travail reste à faire pour établir les

conditions d'émergence d'une sagesse numérique collective.

Peut-on donc imaginer que les conditions d'une telle gouvernance démocratique s'établissent dans l'avenir ? D'un côté, on assiste à l'incapacité actuelle des institutions de créer les conditions de cette émergence de la sagesse collective via le numérique. Les leçons du traumatisme laissé par la naissance du fichier "TES", créé en 2016 par le Ministère de l'Intérieur (un fichier central des données des citoyens français visant à réunir différents fichiers dans une base unique) ne semblent pas avoir été tirées. A l'époque, l'administration avait imposé une décision éminemment politique concernant l'ensemble de la population française sans aucun débat démocratique, mais par le détour d'un obscur décret n'ayant même pas fait l'objet d'un arbitrage interministériel. Plus récemment - avec le projet Alicem - le gouvernement a paru plus conscient de la sensibilité du sujet, celui de la reconnaissance faciale sur une application mobile gouvernementale permettant de se connecter à des services publics et de transmettre des informations personnelles. Pour autant, un décalage s'est de nouveau manifesté entre la machine administrative qui avait créé cette solution technologique et l'appropriation par les utilisateurs, car les conditions de cette appropriation, à commencer par celles du débat, de la maturité des usages et de l'envie d'utilisation, n'étaient pas toutes réunies. Le sujet a fait l'objet d'un examen juridique savant par le Conseil d'Etat, bien éloigné du pouls battant du débat démocratique. Mais l'idée qu'il faille impliquer des collectifs citoyens, des philosophes, des sociologues, des designers et des scientifiques dans des projets de construction technologique au service de l'agenda politique n'est pas encore ancrée dans l'appareil institutionnel. Cette situation peut bien sûr évoluer - des exemples positifs nous viennent du Canada et des Pays scandinaves, entre autres. Il existe pourtant une situation structurelle actuelle qui nous amène à être plus pessimistes : de plus en plus la gou-

vernance des technologies ne se fait plus par les États ni par les citoyens, mais par les entreprises et les consommateurs. Or si le consommateur recherche essentiellement l'utilité et la facilité d'usage, le citoyen a une exigence démocratique et exige aussi la « redevabilité » au sens anglo-saxon d'« accountability ». Aujourd'hui, les grandes entreprises du numérique productrices de solutions techniques mènent ce sujet de la gouvernance des technologies, en testant des nouveaux produits sur le marché et en faisant dialoguer les juristes et les développeurs. Dans le domaine de l'éthique de l'intelligence artificielle, par exemple, si plusieurs organisations internationales, dont la Commission Européenne, définissent les grands principes, ce sont les ingénieurs, développeurs et chercheurs des entreprises privées qui en définissent les usages concrets. Pourtant, on voit bien que derrière la définition des standards techniques subsistent des questions à caractère politique. Face à ce défi les responsables politiques et plus largement les institutions républicaines ne maîtrisent pas encore suffisamment les codes, au sens figuré (codes sociaux du monde technologique) comme au sens propre (programmation) du terme.

Interactions Humain-Machine : quel cadre juridique pour demain ?

CÉLIA ZOLYNSKI

Célia Zolynski est professeur de droit privé à l'École de droit de la Sorbonne de l'Université Paris 1 Panthéon-Sorbonne où elle codirige le Département de recherche en droit de l'immatériel (IRJS-DreDis). Membre du Comité pilote national d'éthique du numérique, elle est en outre personnalité qualifiée au sein de la Commission consultative nationale des droits de l'Homme (CNCDH) et du Conseil supérieur de la propriété littéraire et artistique (CSPLA) après avoir été membre du Conseil national du numérique (CNNum). Ses activités de recherche et d'enseignement portent sur le droit du numérique, le droit de la propriété intellectuelle, le droit du marché et les libertés fondamentales. Elle anime plusieurs groupes de travail interdisciplinaires et projets de recherches collectives sur la protection et la valorisation des données et la régulation des systèmes algorithmiques.

Analyse de comportements sur les réseaux sociaux, développements d'interfaces utilisateurs, déploiement d'objets intelligents : l'Intelligence artificielle nous invite au cœur de nouvelles interactions. Songeons par exemple aux milliers d'agents virtuels, conversationnels qui permettent de rendre des services aussi variés qu'améliorer l'expérience client, optimiser la stratégie de comportement ou donner des conseils médicaux. Ces nouvelles formes d'interaction humain machine suscitent l'interrogation, comme en attestent les travaux menés par Laurence Devillers sur les machines émotionnelles et les mécanismes de co-adaptation humain-machine¹. Cela rejoint certains des questionnements placés au cœur de la stratégie portée par l'Union européenne pour développer une approche européenne de l'Intelligence artificielle centrée sur l'humain qui entend déterminer comment l'IA peut être mise au service de l'humanité, et faire de l'Union européenne le foyer d'une IA éthique et source de progrès social². Le groupe d'experts de haut niveau nommé par la Commission préconise en ce sens de penser une IA « digne de confiance » reposant sur le respect des droits fondamentaux³. Protéger la personne humaine interagissant avec l'IA constitue dans ce contexte la première posture

¹ L. Devillers, Les robots émotionnels, L'observatoire, 2020 – v. également CNIL, A votre écoute, Exploration des enjeux éthiques, techniques et juridiques des assistants vocaux, Livre Blanc n°1, 2020.

² Commission européenne, L'intelligence artificielle pour l'Europe, Communication du 25 avr. 2018 COM(2018) 237 final - v. également la résolution du Parlement européen du 12 fév. 2019 sur une politique industrielle européenne globale sur l'intelligence artificielle et la robotique (2018/2088(INI)).

³ Groupe d'experts de haut niveau, Lignes directrices en matière d'éthique pour une IA digne de confiance, 2018.

à adopter dans le contexte du déploiement de ces systèmes (I). Cette approche « défensive » ne saurait toutefois suffire à porter une vision de l'IA centrée sur l'Humain. Certains proposent par conséquent d'aller au-delà afin de mettre en pouvoir d'agir l'utilisateur. Il s'agirait de lui reconnaître un nouveau statut propre lui permettant de devenir agent du système afin de conserver son autonomie (II). La protection et mise en pouvoir d'agir peuvent ainsi constituer les deux composantes de l'équation qui doit permettre d'assurer l'autonomie humaine dans le contexte du développement de l'IA.

I- Protéger l'Humain lors de ses interactions avec la Machine

Au-delà des formidables perspectives qu'elles représentent, les interactions de l'humain avec les systèmes d'IA peuvent engendrer des risques de manipulation déloyale, de tromperie, de conditionnement voire d'asservissement. Parmi les solutions envisagées pour répondre à de tels risques, deux principes peuvent être mobilisés : un principe de transparence (A) ainsi qu'un principe de non-discrimination (B).

A. Protection et transparence

Un principe de transparence devrait se déployer à plusieurs niveaux. Il pourrait tout d'abord se traduire sous la forme d'une obli-

gation mise à la charge des opérateurs d'informer l'utilisateur de la nature de leurs interactions. Une première confusion peut en effet résulter de la méconnaissance par l'utilisateur de la nature réelle de son interlocuteur, ce qui explique la récente adoption par l'Etat de Californie d'une réforme relative à l'usage des chatbots⁴ qui prévoit de sanctionner toute personne utilisant un chatbot afin de communiquer ou d'interagir avec autrui pour inciter à l'achat ou vendre des produits ou services dans le cadre d'une transaction commerciale ou pour influencer le vote dans un cadre électoral. Il est également imposé à tout opérateur de révéler à l'utilisateur que l'interaction est réalisée via un chatbot. Ni le droit français ni le droit européen ne reconnaissent pour l'heure d'obligation équivalente même si différentes propositions sont formulées en ce sens, notamment afin que l'utilisateur puisse identifier qui s'adresse à lui au travers de la machine⁵. On pourrait toutefois soutenir la proposition qu'à terme, une obligation légale soit expressément consacrée afin de l'informer spécifiquement de la nature de son interlocuteur (humain ou machine ?).

Une autre source de confusion peut résulter des mécanismes au travers desquels les systèmes d'IA pourraient conduire à influencer, voire déterminer les décisions de leurs utilisateurs. Cette réflexion se déploie tout particulièrement avec l'exploitation croissante des biais cognitifs tels les nudges, techniques pour inciter des personnes ou une population ciblée à changer leurs comportements ou à faire certains choix sans être sous contrainte ni obligations et qui n'implique aucune sanction. Ces techniques font l'objet de nombreuses études,

⁴ Business and Professions Code - BPC, Division 7, Part 3, Chapter 6 : Bots.

⁵ K. Sein, "Concluding Consumer Contracts via Smart Assistants: Mission Impossible Under European Consumer Law?", EuCML 2018, 179. La réforme « Nouvelle donne pour le consommateur » paraît aller en ce sens dès lors qu'elle admet l'application des dispositions de la directive

2011/83 visant les places de marchés au fonctionnement de leurs interfaces en ligne. V. également le projet de recommandation du PE, 21/01/2020, Assurer la protection du consommateur et la libre circulation des produits et services – v. plus généralement, Lignes directrices en matière d'éthique pour une IA digne de confiance, préc., point 131.

dans le prolongement des travaux du prix Nobel d'économie Richard Taler. S'agissant du design des interfaces numériques, une étude de la CNIL démontre alors comment ces mécanismes de manipulation peuvent être sanctionnés sur le fondement du RGPD⁶. Au-delà des enjeux, certes majeurs, du traitement des données personnelles, la réflexion pourrait également s'ouvrir pour déterminer comment appréhender la captation voire de la manipulation de deux composantes essentielles de l'autonomie humaine que constituent l'attention et les émotions dès lors que le risque est celui d'une « perte de liberté, de réflexivité et de capacité de négociation » de l'humain s'il n'est plus en mesure de « se mettre à distance »⁷. Nombreux considèrent en effet que le déploiement des nouvelles interfaces humain-machine pourraient conduire à transformer nos modalités d'interactions. Par nature engageantes, ces interactions supposent de l'implication ; elles supposent de « faire corps »⁸.

Ces nouvelles formes d'interaction pourraient d'une part amplifier les questions qui surgissent aujourd'hui du fait de l'essor sans précédent de l'économie de l'attention. L'angle mort de la protection de l'attention pourrait ainsi justifier de reconnaître un nouveau droit subjectif au bénéfice des utilisateurs sur le modèle du droit à la protection des données personnelles : *un droit à la protection de l'attention*. Il s'agirait ainsi de répondre à l'asymétrie

informationnelle subie par les individus lors de leurs usages des services numériques en leur permettant de visualiser la captation de leur attention. A ce titre, tout opérateur de traitement de l'attention aurait une obligation d'informer ses utilisateurs dans des termes clairs et intelligibles des dispositifs de captation de l'attention qu'il déploie sur son service ; il devrait mettre en oeuvre une signalétique spécifique voire offrir un accès aux métriques de captation de l'attention. Les utilisateurs pourraient ainsi choisir de quitter un service aux mauvaises pratiques pour privilégier d'autres services, plus respectueux de leur libre arbitre. Cette protection devrait en outre être assurée par la reconnaissance d'obligations à la charge des opérateurs de traitement de l'attention, sur le modèle des obligations imposées au responsable de traitement de données personnelles⁹.

D'autre part, le déploiement de la robotique émotionnelle annonce l'exploitation des émotions de l'utilisateur de ces systèmes à l'occasion de ses interactions notamment s'agissant de l'analyse de la voix humaine qui est une source d'informations incroyablement riche¹⁰. Une telle exploitation pourrait alors déterminer la prise de décision de l'utilisateur. Si une manipulation est avérée, divers fondements juridiques pourraient être mobilisés telle la sanction des vices du consentement du droit des contrats¹¹ ou de l'abus de faiblesse en droit pénal¹². On pourrait toutefois réfléchir à faire

⁶ CNIL, La forme des choix. Données personnelles, design et frictions positives, Cahier Prospective n°6, 2019 -v. également CNIL, SAN – 2019-001, 21 janv 2019 : Comm. com. électr. 2019, comm. 32 et 43, obs. N. Metallinos ; Dalloz IP IT 2019, p. 165, note E. Netter ; JCP E 2019, 1059, note J. Déroutel.

⁷ H. Guillaud et A. Masure « Retro-design de l'attention : limites, angles morts et autres propositions », <http://www.internetaactu.net/2019/01/25/retro-design-de-lattention-limites-angles-morts-et-autres-propositions/>

⁸ H. Guillaud et A. Masure, op. cit.

⁹ Sur ces propositions, v. tout particulièrement M. Leroy, F. Levin, C. Zolynski, « L'économie de l'attention saisie par le droit : plaider pour un droit à la protection de

l'attention », Dalloz IP IT 2019, n°11, pp. 614-622

¹⁰ Livre blanc CNIL préc. – adde, L. Devillers, préc.

¹¹ C. civil, art. 1137.

¹² C. pénal, art. 223-15-2 : « Est puni de trois ans d'emprisonnement et de 375 000 euros d'amende l'abus frauduleux de l'état d'ignorance ou de la situation de faiblesse soit d'un mineur, soit d'une personne dont la particulière vulnérabilité, due à son âge, à une maladie, à une infirmité, à une déficience physique ou psychique ou à un état de grossesse, est apparente ou connue de son auteur, soit d'une personne en état de sujétion psychologique ou physique résultant de l'exercice de pressions graves ou répétées ou de techniques propres à altérer son jugement, pour conduire ce mineur ou cette personne à un acte ou à une abstention qui lui sont gravement préjudiciables » (nous soulignons).

évoluer le cadre légal pour saisir ces nouvelles pratiques qui tendent à « mettre à nu » l'utilisateur d'une façon inégalée par le passé. Il s'agirait a minima ici encore de consacrer une obligation d'information à l'égard de l'opérateur ayant recours à un système d'IA exploitant les émotions de l'utilisateur. On pourrait en outre préconiser de proscrire de telles pratiques dès lors qu'elles visent à tromper son consentement ou exploiter sa vulnérabilité à des fins qui lui seraient préjudiciables.

B. Protection et non-discrimination

La protection de l'utilisateur doit également conduire à lui reconnaître un droit de ne pas faire l'objet de décision prise sur des biais injustifiés. Les techniques d'IA peuvent en effet induire des biais, sources de discrimination hautement préjudiciable pour un individu ou un groupe d'individus telles des discriminations tenant à l'origine raciale, au genre ou encore au lieu de résidence¹³. Certaines informations sont d'ores et déjà ponctuellement imposées par le droit de la consommation, à l'image de celles relatives au prix qui doivent permettre d'encadrer les pratiques de tarification dynamique¹⁴ auxquelles s'ajoutent les dispositions du droit des données personnelles dès lors que les variations de prix résultent du traitement de ces données. Mais au-delà, il demeure encore difficile d'évaluer et de rapporter la preuve de possibles discriminations préjudiciables à l'individu¹⁵. Différentes propositions entendent

désormais lever ces obstacles. Concernant l'évaluation de la discrimination, le rapport Villani préconise ainsi la réalisation par les concepteurs d'IA d'étude d'impact de discrimination, sur le modèle des études d'impact concernant les risques générés par le traitement des données à caractère personnel¹⁶. Son traitement paraît également appeler des adaptations du cadre juridique¹⁷, par exemple pour faciliter la preuve d'une discrimination illicite. Il pourrait être ici envisagé de construire un droit de la non-discrimination algorithmique organisant un régime probatoire spécifique ; celui-ci pourrait organiser un renversement de la charge de la preuve d'une possible discrimination préjudiciable à l'individu afin de le faire peser le risque probatoire sur le professionnel déployant un système d'IA.

Divers instruments sont ainsi consacrés ou envisagés afin d'assurer la protection de l'humain lors de ses interactions avec l'IA. Pour autant, penser une IA centrée sur l'humain nécessite d'aller au-delà pour lui offrir la possibilité de prendre une décision éclairée et garantir pleinement son autonomie. Il s'agirait de mettre l'utilisateur en pouvoir d'agir afin de contrôler les décisions prises par les systèmes d'IA avec lesquels il interagit.

¹³ CNIL, Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'Intelligence artificielle, rapport 2017, p. 34 &s. ; CNIL et Défenseur des droits, Algorithmes : prévenir l'automatisation des discriminations, 2020.

¹⁴ En ce sens v. l'art. 6 de la directive 2011/83 révisé par la directive 2019/2161 2020, pp. - Adde, F. Marty, « Plateformes numériques, algorithmes et discrimination », Revue de l'OFCE, 164 (2019).

¹⁵ Sur ce point, v. CNIL, Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'Intelligence artificielle, 2017, p. 31 &s. – v. également, J. Charpenet et C. Lequesne Roth, « Discrimination et biais générés. Les lacunes juridiques de l'audit algorithmique », D. 2019, p. 1852.

¹⁶ Donner un sens à l'Intelligence artificielle, rapport de la mission présidée par C. Villani, 2018, p. 147.

¹⁷ CNIL-DDD, op. cit.

II- L'Humain, un nouvel agent des systèmes d'IA

Différentes propositions entendent désormais reconnaître un nouveau rôle à l'utilisateur des systèmes d'IA. Il est ainsi préconisé de le concevoir comme un véritable agent du système en déployant une nouvelle notion : l'agentivité.

La notion d'agentivité est issue des sciences sociales, notamment des travaux du prix Nobel d'économie Amartya Sen, consacrés à la « personal agency »¹⁸ pour nommer la capacité offerte aux individus, leur liberté à utiliser leurs biens pour choisir leur propre mode de vie.

En effet, dans ces premiers travaux, la notion de Personal agency fonde la mise en pouvoir d'agir (empowerment) de l'individu sur deux composantes principales : d'une part la capacité pour l'individu d'endosser le rôle d'agent c'est-à-dire de décider d'agir dans le sens de ses valeurs »¹⁹ ; d'autre part, l'opportunité d'exercer de façon effective son agentivité c'est-à-dire d'être en mesure de contribuer à rendre ses propres décisions effectives²⁰. Cela caractérise différents vecteurs de mise en pouvoir liés au développement humain, à la réduction de la pauvreté ou à l'amélioration de la condition des femmes.

La notion de « personal agency » paraît pouvoir être traduite au travers un nouveau concept, celui d'« agentivité » de l'utilisateur d'un système d'IA. Il s'agirait ainsi de mettre en exergue sa capacité d'action sur le système, sa capacité à le transformer, à l'influencer. Transposer cette notion d'agentivité permet-

trait ainsi de proposer une nouvelle lecture du rôle confié à l'utilisateur interagissant avec ces systèmes. Sortant de son rôle passif pour acquérir le statut d'agent, il serait alors mis en mesure de prendre des décisions autonomes éclairées à l'égard des systèmes d'IA. Une telle proposition pourrait s'inscrire dans le prolongement des déclarations de Tim Berners-Lee, fondateur d'Internet et lauréat du prix Turing, lorsqu'il dénonce la réalité actuelle du Web et les monopoles de situation qu'elle engendre sur la gestion des données personnelles²¹. Cette notion d'agentivité de l'individu pourrait reposer sur deux piliers. Elle devrait tout d'abord porter sur le traitement de données nourrissant les systèmes d'IA. Elle se prolongerait ensuite au travers le fonctionnement de ces systèmes *per se*.

S'agissant des données personnelles, les fondations d'une telle construction paraissent progressivement se dessiner. La modernisation du droit des données personnelles réalisée par le RGPD entend en effet considérer la personne concernée comme un acteur à part entière qu'il convient d'armer pour le mettre en pouvoir d'agir dans le contexte de la circulation des données ; en atteste la reconnaissance d'une nouvelle prérogative à son profit au travers le droit à la portabilité de ses données personnelles pensée à des fins concurrentialistes. L'utilisateur du service se voit ainsi conférer un rôle actif en tant qu'opérateur de marché et sa souveraineté se matérialise au travers ce libre choix soit de changer d'opérateur, soit de migrer vers un système de gestion individuelle de ses données. C'est dans ce contexte que se sont

¹⁸ A. Sen définit l'agentivité comme dimension de sa théorie des « capacités » (Well-being, agency and freedom: the Dewey Lectures 1984, The Journal of Philosophy, n°82, p. 206).

¹⁹ R. Alsop et al. (2006), « Measuring Empowerment in Practice: Structuring Analysis and Framing Indicators », p. 6, World Bank Policy Research Working Paper 3510, February 2005

²⁰ R. Alsop et al. (2006), préc., p. 3.

²¹ V. également C. Lazaro et D. Le Métayer, Control over personal data : true remedy or Fairy tale ?, SCRIPTed, 12 :1, 2015.

développées les solutions de Cloud Personnel à l'image de celle portée par Tim Berners-Lee avec le système SOLID qu'il développe avec le MIT. Déclinaison technique de l'exercice du droit à la portabilité, le cloud personnel de l'individu est doté d'un ensemble de connecteurs lui permettant de récupérer l'ensemble de ses données personnelles. Avec ces offres, il peut regrouper toutes ses données dans un système unique, avec la possibilité d'en moduler les accès au profit de services innovants. A partir de ces évolutions juridiques et techniques, il pourrait être proposé de franchir une nouvelle étape : faire de l'utilisateur un véritable agent du système en lui permettant d'agir directement sur son architecture afin qu'il soit pleinement en mesure d'exercer sa souveraineté²².

L'agentivité offrirait alors à l'utilisateur la possibilité de contrôler l'exécution des décisions grâce à des solutions techniques. Cela supposerait que l'individu puisse orchestrer sous son contrôle les traitements de données correspondants à ses décisions, avec certaines garanties techniques ce qui lui assurerait que les traitements qu'il diligente sont conformes aux décisions qu'il prend, intègres et confidentiels²³.

Au-delà du traitement des données, l'agentivité pourrait plus largement se déployer pour s'étendre aux systèmes d'IA *per se*. Il s'agirait alors de reconnaître à l'utilisateur un droit à agir sur le fonctionnement de ces systèmes afin de contrôler les décisions le concernant. En France, des premières pistes de réflexion commencent à se dessiner en ce sens. Dans son rapport sur les enjeux éthiques des algorithmes et l'IA, la CNIL propose ainsi de consacrer le concept de « jouabilité » en reprenant l'expérience « MesSystèmes » lancée par la Fondation pour un internet nouvelle génération

(FING)²⁴.

Il est ainsi recommandé de permettre aux utilisateurs de « jouer » avec ces systèmes en leur permettant d'en faire varier les paramètres, les métriques. Devenu agent de la décision prise par le système d'IA, l'utilisateur deviendrait l'un des rouages de la boîte noire ce qui permettrait, sans avoir à l'ouvrir, d'en moduler les résultats, ce qui supposerait de dessiner des systèmes algorithmiques pour mettre l'individu en capacité d'exercer une telle agentivité. Au-delà, certains préconisent de dépasser la simple expérience privée vers laquelle paraissent nous conduire les interactions agents conversationnels-utilisateurs, machine-humain pour construire des expériences partagées, des espaces collaboratifs. L'anthropologue Stefana Broadbent propose ainsi de penser comment, avec ces outils, organiser nos interactions avec les autres afin de mobiliser collectivement des expériences partagées et des modalités démocratiques de contrôle de la technique par les usagers. Il s'agirait d'éviter le déploiement d'une économie de l'infantilisation, et ses risques de manipulation comportementale en prônant un nouvel agencement des interactions²⁵.

Autant de nouvelles perspectives qu'il convient certainement d'explorer afin de garantir la pleine autonomie de l'humain lorsqu'il interagit avec un système d'Intelligence artificielle.

²² Sur ces propositions, v. N. Anciaux et C. Zolynski, « Empowerment et Big data sur données personnelles : de la portabilité à l'agentivité », in *Droit et Big data*, dir. F. G'Sell, Dalloz, coll. Thèmes et commentaires, 2020, pp. 219-237.

²³ N. Anciaux et C. Zolynski, préc.

²⁴ CNIL, Comment permettre à l'homme de garder la main ?, préc., p. 57.

²⁵ Cité par H. Guillaud et A. Masure, art. préc.

Serons-nous suffisamment sages pour continuer à protéger nos libertés ?

RÉGIS CHATELLIER

Régis Chatellier pilote des projets d'études et d'explorations prospectives de sujets émergents liés aux données personnelles et à la vie privée au sein du LINC¹ (Laboratoire d'innovation numérique de la CNIL). Ses travaux se situent à la rencontre entre innovation, technologies, usages, société, régulation et éthique. Diplômé de l'Institut d'études politiques de Grenoble, il enseigne la « Gouvernance de la donnée urbaine » à l'École urbaine de Sciences Po Paris.

La loi Informatique et Libertés naît un 6 janvier de l'année 1978, mais sa genèse remonte à quatre années plus tôt, avec un article paru dans *Le Monde* intitulé, « Safari, ou la chasse aux français ». Ce projet gouvernemental de « Système Automatisé pour les Fichiers Administratifs et le Répertoire des Individus » prévoyait d'instituer un identifiant unique pour interconnecter les fichiers administratifs, et « au moyen d'un ordinateur puissant », de « rassembler la masse énorme des renseignements grappillés sur tout le territoire ». De là, selon les mots du journaliste, se trouvait « posé un problème fondamental, celui des rapports des libertés publiques et de l'informatique ». Les quatre années suivantes permirent de rédiger et publier la première loi Informatique et Libertés, qui prévoyait la création de la Commission nationale de l'informatique et des libertés (CNIL), première autorité administrative indépendante.

Si l'on fait l'exercice consistant à relier prospective et sagesse collective, le thème de ce colloque, il est intéressant de constater que l'article 1 de cette loi n'a pas pris une ride en quarante ans : « L'informatique doit être au service de chaque citoyen. Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques ». Les rédacteurs posaient ainsi un socle, un cadre dans lequel on pouvait évoluer et penser le futur.

A la fin des années 70, seules la France et l'Allemagne (1977) s'étaient dotées d'un cadre juridique relatif à la protection des données. L'Europe a depuis harmonisé son cadre, d'abord avec une Directive en 1995, puis avec le Règlement général à la protection des données, le fameux RGPD, entré en application en 2018. La protection des données personnelles et des libertés est inscrite aux articles 7 et 8 de

¹ Le Laboratoire Innovation numérique de la CNIL (LINC) est constitué d'une équipe orientée humanités numériques qui s'ouvre sur différentes disciplines pour alimenter ses réflexions, de trois manières différentes :

- Explorer : Opérer une veille pour explorer le futur de la société numérique, pour mieux anticiper l'impact de l'usage des innovations technologiques sur la vie privée et les libertés.
- Expérimenter : LINC pilote des projets d'expérimentation, pour mieux cerner les usages numériques émergents.
- Échanger : LINC crée du lien entre les acteurs de la société numérique (entreprises, institutions, associations, société civile...), pour mieux les informer face aux nouveaux enjeux reliant éthique, libertés, et vie privée.

la Charte des droits fondamentaux de l'Union européenne de 2002.

L'Europe a ainsi su faire preuve d'une forme de « sagesse collective », même si les débats autour de la rédaction du RGPD furent intenses, pour réussir à s'entendre collectivement sur un continent de 500 millions d'habitants et établir un cadre commun sur un sujet aussi sensible que celui de la protection des données. Ce cadre reste général et applicable aujourd'hui à l'ensemble de nos activités tant les données personnelles sont partout. Nous nous promenons tous avec le plus puissant des capteurs dans la poche, le smartphone. Toutes nos activités sont enregistrées. Là où la protection des données personnelles consistait d'abord à ses prémisses en une protection face aux fichiers de l'État. Progressivement elle a concerné de plus en plus les activités du secteur privé, et notamment de ce que l'on appelle parfois le « capitalisme de surveillance » (Shoshana Zuboff).

Prévenir les enjeux actuels et futurs

Des textes comme le RGPD ou la Loi Informatique et Libertés, bien que de portée générale, font de plus en plus figure de textes pivots dans un contexte de numérisation de la société, d'hyper-personnalisation des services, et de collecte exponentielle des données. Voici une série non exhaustive d'exemples pour lesquels nous devons agir et rester vigilants, aujourd'hui et demain.

La santé : des tests ADN récréatifs mais pas inoffensifs

Les données génétiques sont non seulement des données personnelles sensibles, mais on les qualifie de données pluri-personnelles. Elles sont liées à l'ADN de leur porteur et permettent

de le distinguer parmi tous les autres êtres humains, mais elles sont également partagées pour partie par plusieurs personnes. L'ADN d'un individu est construit à partir de l'ADN de ses parents, ainsi, plus on est proche dans l'arbre généalogique d'une personne, plus nos ADN sont proches. Pourtant, depuis une dizaine d'années des sociétés comme 23andMe proposent des tests génétiques récréatifs à des prix très compétitifs et détiennent désormais des bases de données génétiques extrêmement importantes.

Des personnes font ces tests, sans connaître ou penser aux impacts possibles pour elles-mêmes et pour d'autres. Il existe donc un fort enjeu éthique lié au consentement pour soi, qui entraîne de fait l'ensemble de son arbre généalogique proche. En termes de santé, un test qui révélerait une anomalie génétique et le risque de maladie associée pour une personne, pourrait donner des indications sur les risques pesant sur les parents, enfants et proches de cette même personne. Vient alors un problème de conciliation entre le respect du secret médical, la nécessité de protéger la santé des membres de la famille, mais aussi le droit de ne pas savoir quelles maladies nous sommes susceptibles de développer dans les dix années à venir.

En matière de recherche, l'impossibilité d'anonymiser totalement les données génétiques pourrait notamment impliquer de laisser à l'individu la possibilité de choisir au cas par cas les recherches et acteurs avec lesquels il consent à partager ses données. Il s'agirait alors de mettre en place un système de consentement dynamique et non pas délivré une fois pour toutes lors du partage initial comme le proposent ces acteurs.

Le fait que les français aient recours aux tests récréatifs doit nous amener à nous pencher sur les avantages et les inconvénients qu'il y aurait à organiser un régime juridique permettant le développement sous contrôle d'une offre domestique de tests génétiques de filiation à des fins personnelles. Organiser le marché permettrait de le réguler, du point de vue de la protection des données, les matériaux bio-

logiques et les analyses génétiques qui sortent des radars lorsqu'ils sont envoyés hors UE, mais aussi parce qu'aucun système ne garantit la qualité des tests effectués à l'étranger.

Protection des données et environnement : l'exemple des assistants vocaux

A l'heure de la prise de conscience de l'urgence à agir sur les conséquences de nos modes de vies sur l'environnement, les nouvelles offres de services et produits numériques et leurs usages associés semblent parfois bien éloignés de ces considérations, presque d'un autre temps. Des outils comme les assistants vocaux ne sont absolument pas pensés en terme écologiques. Lorsque vous éteignez une lampe dont l'ampoule est connectée avec Google Home par exemple, une simple demande "OK Google, éteins la lumière" entraîne toute une chaîne d'actions. L'assistant doit d'abord convertir le son en texte, l'envoyer dans un data center, souvent de l'autre côté de l'Atlantique, où le texte sera analysé pour produire une réponse. Ici, éteindre la lumière. La commande devra à nouveau traverser l'Atlantique pour qu'enfin, votre ampoule ait « l'intelligence » de s'éteindre. S'il est à ce jour difficile de quantifier la dépense énergétique d'une telle action, il n'en reste pas moins facile d'en constater l'hérésie d'un point de vue écologique. Pour des actions qui peuvent être opérées en local et d'un geste de la main, on utilise des systèmes extrêmement consommateurs en énergie mais aussi en données personnelles.

Les assistants vocaux vont en effet écouter nos conversations, beaucoup plus qu'il ne le faut. Ces outils demandent à être entraînés le plus possible. Ils restent en veille, en attente du mot clé (Ok Google, Hey Alexa, etc.) pour

pouvoir se connecter, écouter notre demande et produire une réponse. Ces nouvelles interfaces ouvrent une nouvelle forme d'interaction qui n'est plus individuelle. Nous ne sommes plus seuls face à notre écran, mais potentiellement plusieurs dans la pièce où se trouve l'assistant, dans l'espace du foyer ou dans des lieux publics. Se pose alors de nouvelles questions quant à l'information et la transparence des personnes qui n'ont plus toujours conscience d'être écoutées, et même enregistrées. Tout ce que l'assistant écoute après avoir cru entendre le mot clé (parfois il fait erreur), est conservé et accessible depuis un tableau de bord où l'on peut rejouer les conversations - un système qui pose la question de la capacité des personnes à garder le contrôle de ces nouvelles interfaces qui permettent l'invisibilisation de l'écoute par la machine.

À cet enjeu il faut ajouter le risque de resserrement de la fenêtre ouverte sur le web, à l'hyper-personnalisation et à l'accès à l'information. Lorsque vous tapez une requête sur un moteur de recherche, vous obtenez une série de réponses, sur plusieurs pages. Lorsque vous posez une question à votre assistant, vous n'obtenez qu'une réponse. Le moteur de recherche devient un moteur de réponse (au singulier). La dimension des bulles de filtres, conceptualisées par Eli Pariser dès 2010, se rétrécit plus encore. La CNIL a publié en septembre 2020 un livre blanc sur les assistants vocaux², qui permet d'en cerner les enjeux en même temps qu'il donne des clés à ceux qui souhaiteraient promouvoir des usages de manière la plus respectueuse des données.

² https://www.cnil.fr/sites/default/files/atoms/files/cnil_livre-blanc-assistants-vocaux.pdf

Le design n'est pas l'angle mort de la régulation de la protection des données

Le design des interfaces n'a pas attendu le Règlement général sur la protection des données (RGPD) pour influencer nos vies. Bien avant le numérique, nos déplacements et nos actions ont été guidés par les architectures de choix conçues et mises en oeuvre par d'autres. C'était et c'est toujours le cas pour nos déplacements en ville et par l'aménagement urbain, avant que nous ne soyons téléguidés par nos applications. La grande distribution également a depuis longtemps modélisé ses hypermarchés par la construction de chemin préétablis mais non perceptibles directement par leurs clients afin de maximiser l'acte d'achat, depuis l'emplacement des packs d'eau à l'extrémité du magasin aux friandises disposées sur la caisse. Pourtant, nous explorions en 2018 dans un cahier *Innovation et prospective* comment ces questions prennent un tour inédit dès lors qu'elles s'appliquent à des services numériques qui usent de méthodes de design pour parvenir à capter notre attention, en lien avec les techniques associées à l'économie de l'attention, et à traiter toujours plus nos données. Dans un contexte de surcharge informationnelle, et où il n'est pas possible d'appréhender et de comprendre aisément la manière dont les plateformes numériques sont conçues et développées, les individus sont dans l'incapacité de comprendre comment les systèmes sont formés et quelles sont les décisions qu'ils prennent. Leur compréhension va dépendre du design, de la manière dont ils sont mis en scène et comment ils interagissent avec les services. Il en va ainsi pour l'usage même de ces services, mais aussi pour la perception et la compréhension des moments critiques où se joue la protection des données.

Nous appelions ainsi dans notre cahier *Innovation et prospective* à faire entrer plus directement le design dans un triangle de régulation, avec les analyses juridiques et les analyses techniques. Une telle approche trouve en particulier son sens dans la mise en application du principe de transparence, dans l'expression du consentement et dans la conception de l'exercice des droits des personnes concernées (accès, rectification, suppression, portabilité, ...). Pour répondre à ces enjeux et pour encourager les porteurs de projet et designers à co-construire des parcours utilisateurs respectueux du RGPD et de la vie privée, la CNIL a lancé la plateforme Données & Design³. L'objectif est de permettre aux concepteurs de services de se saisir des solutions offertes par le design pour accompagner positivement les utilisateurs dans la compréhension et la maîtrise du fonctionnement des services numériques et des traitements de données afférents.

Les libertés : " Nobody knows you're a dog " ?

Avec le numérique, nous avons gagné la possibilité d'évoluer sous pseudonyme. Cependant, il est de plus en plus difficile pour les utilisateurs de naviguer en ligne sans être tracé. Dans l'espace public, comme nous le décrivions dans notre cahier *IP - La Plateforme d'une Ville*, nous sommes de plus en plus connectés, collectés et captés, par l'entremise de notre smartphone, ou par les dispositifs de surveillance, privés ou publics.

La ville reste traditionnellement associée à un espace permettant à ses habitants de se sentir familiers et anonymes à la fois. Le paradoxe de la modernisation des villes par le numérique est qu'elle reconfigure cet anony-mat, qui se rapproche en cela de l'image de la

³ <https://design.cnil.fr>

place du village où les faits et gestes de chacun finissent par être connus de tous. L'anonymat dans la ville est ainsi en train de s'évanouir alors que c'est une notion centrale de sa modernité. Cette captation des comportements peut prendre différentes formes, et avoir pour conséquence la modification des comportements des personnes se sentant surveillées, au détriment de leurs libertés. Un exemple concret de cet effet de modification des comportements par la connaissance de la surveillance avait été fourni en 2016. Des chercheurs de l'université de Oxford avaient constaté que suite aux révélations d'Edward Snowden sur les outils de surveillance à la disposition des autorités étatsuniennes, la consultation de certaines pages Wikipédia informant sur des sujets sensibles (terrorisme, radicalisation, ...) avait chuté drastiquement (jusqu'à 20 % pour certaines pages).

Tous nos faits et gestes peuvent ainsi être enregistrés, calculés et tracés. Bien que le RGPD et la Loi Informatique et Libertés fournissent les garanties pour que les individus puissent conserver des droits quant à ces données collectées, il reste important que chacun soit conscient de l'usage qui pourrait être fait de leurs données, et que les autorités politiques comme le législateur soient vigilants sur ces questions, et notamment sur la reconnaissance faciale. En effet, aujourd'hui si l'usage des caméras de vidéosurveillance traditionnelles est très encadré, ce n'est pas le cas de l'usage de la reconnaissance faciale dans l'espace public, qui reste interdit tant qu'aucun cadre légal ne l'encadre.

La CNIL demande depuis septembre 2018, puis dans une note publiée en novembre 2019⁴, la tenue d'un réel débat public « à la hauteur des enjeux ». Dans cette note, elle rappelle que La reconnaissance faciale appelle des choix politiques : « sur le rôle dévolu à la technolo-

gie, sur ses effets sur les libertés fondamentales des individus, sur la place de l'humain à l'ère numérique. Ces choix dessineront certains contours du monde de demain. Le débat ne doit donc pas se résumer à un examen technique des potentialités d'usage et de l'efficacité de cette technologie. Il ne peut davantage avoir pour simple objectif de savoir comment rendre acceptable par les citoyens une technologie dont la nécessité s'imposerait de manière évidente. Car tel n'est pas le cas : le sujet est complexe et mérite un débat lucide et approfondi. »

Penser le partage des données dans la ville

En 2017, toujours dans le cahier *IP La Plateforme d'une ville*, le LINC s'est intéressé à l'influence des modèles économiques des plateformes sur la transformation de la ville. Ce fait est assez récent. Google a par exemple commencé à s'intéresser au marché urbain seulement vers 2015-2016. La particularité de certains outils comme Google, Waze, Uber et d'autres réside dans leur capacité à produire du service au public - et non pas du service public - suffisamment en masse pour transformer la ville. Par ailleurs, ces acteurs interviennent sans avoir besoin d'être en relation contractuelle ni même en contact avec l'acteur public, pourtant en charge de l'aménagement du territoire. On se retrouve dans une nouvelle configuration avec de grands acteurs privés qui peuvent modifier la ville, sur la base d'outils qui le plus souvent fonctionnent à partir de la collecte et du traitement de données personnelles. Par exemple, si Waze a la capacité de changer la manière dont on se déplace, c'est parce que nous acceptons en masse de partager notre localisation. Il en va de même pour un certain nombre de services. C'est pour cette

⁴ <https://www.cnil.fr/fr/reconnaissance-faciale-pour-un-debat-la-hauteur-des-enjeux>

raison que nous avons choisi de nous intéresser aux conséquences de ces services sur les politiques publiques, pour faire des propositions.

L'acteur public se retrouve parfois en porte-à-faux entre le choix de se passer de ces plateformes, pourtant très efficaces, et la tentation de les utiliser, avec le risque d'imposer à ses administrés l'usage d'une plateforme dont le modèle repose sur un échange données contre services, au détriment de leur droit à ne pas accepter. Dans notre *Cahier*, plutôt que d'en rester aux constats, nous proposons des scénarios de partage des données qui pourraient permettre d'engager un rééquilibrage privé/public par les données. Beaucoup de données sont produites hors du périmètre organique du service public (régie directe, concession). Or ces données seraient très précieuses pour remplir des missions d'intérêt général. Le cœur du pouvoir appartient à celui qui détient la donnée et qui a la capacité de l'utiliser de la meilleure manière. Fort de ce constat, au sein de la CNIL nous avons donc essayé de concevoir des scénarios de partage de la donnée afin d'opérer une forme de rééquilibrage et pour créer des missions de service public. Cela avait pris la forme d'une matrice à quatre entrées, avec quatre formes différentes de partage des données : l'obligation de l'open data pour certains acteurs privés, sous un format anonymisé et agrégé ; l'extension dans la loi de la notion de données d'intérêt général qui imposerait aux acteurs détenteurs de ces données fines – et donc personnelles – de les mettre à disposition de l'acteur public pour des missions d'intérêt public ; le développement de plateformes techniques d'accès et de partage des données qui ouvrirait la possibilité aux acteurs publics de requêter ces bases de données, sans avoir à les récupérer et les traiter eux-mêmes ; enfin la Portabilité citoyenne, sur la base de l'article 20 du RGPD et le droit à la Portabilité, des com-

munités d'utilisateurs pourraient exercer leur droit à la portabilité vis-à-vis d'un service pour mettre leurs données à disposition d'un acteur public, pour une finalité spécifique en lien avec une mission de service public. Ces quatre scénarios⁵ offrent un cadre d'analyse applicable au partage des données, dans le respect de la protection des données, qui pourrait s'appliquer bien au-delà du seul cas de la ville numérique.

Conclusion

Ces différents exemples et axes travail que je vous présente aujourd'hui ont aussi pour but de contredire l'idée que la CNIL et le RGPD sont binaires. Au contraire, nous voulions montrer qu'il y a plusieurs manières de faire et d'innover, par un partage de données respectueux de la protection des données ou par l'amélioration de la transparence et l'information dans le design des interfaces, par exemple. Aller vers une forme de sagesse collective implique d'abord de se rappeler des grands principes fondamentaux à respecter. Mais surtout il faut envisager le RGPD non pas comme un cadre qui empêche d'innover mais au contraire comme un cadre dans lequel on peut innover et faire évoluer les choses. Depuis 2018, la Californie, le Brésil et bientôt l'Inde se sont dotées d'un cadre législatif inspiré du RGPD, le Japon avait demandé son adéquation au RGPD dès 2017. Ce cadre, bien que contraignant, offre de nouvelles perspectives pour faire émerger de nouvelles formes d'innovation respectueuses des individus, et l'Europe de rayonner et faire valoir sa conception et sa philosophie de la protection des données et de la vie privée.

⁵ <https://www.cnil.fr/fr/reconnaissance-faciale-pour-un-debat-la-hauteur-des-enjeux>

De la toxicité de certains termes associés à «technologie» : impact, éthique, problème, solution, usage...

DANIEL KAPLAN

Pionnier du numérique et de l'internet, entrepreneur et prospectiviste, Daniel Kaplan a cofondé la Fondation internet nouvelle génération (Fing) et l'a dirigée jusqu'en 2016. Il développe depuis le Réseau Université de la Pluralité, réseau international de celles et ceux qui mobilisent les imaginaires pour élargir le champ des futurs pensables : artistes, utopistes, designers... Le texte présenté ici est une transcription de l'exposé oral présenté par l'auteur.

L'idée de cette présentation est de faire une série d'allers-retours entre les questions qui se posent habituellement sur les technologies et une série de travaux artistiques ou de science-fiction.

Introduction : Un appel à la responsabilité des chercheurs

Je vais pour commencer prendre l'exemple de Work+, un projet que nous menons depuis un an qui explore de manière prospective les futurs du travail. La matière première qui constitue ce projet n'est pas la connaissance mais la fiction.

En 2013, le principal congrès international annuel sur les interactions homme-machine, Conférence on Human Factors in Computing Systems (CHI), a eu lieu à Paris. Quatre chercheurs dont l'article avait été validé par le comité scientifique sont montés sur scène et ont commencé par révéler qu'ils étaient en réalité des robots venus du futur pour remercier les chercheurs d'avoir si bien préparé l'humanité à leur domination.

Voici un extrait de leur discours :

“La communauté CHI [des spécialistes des interfaces homme-machine] a pris sur elle la responsabilité de concevoir des technologies faciles d'utilisation, accessibles, efficaces, amusantes et omniprésentes. À première vue, ces efforts semblent avoir rendu la vie plus facile, plus agréable, mieux informée, plus saine et durable. Mais rien ne pourrait être plus éloigné de la réalité. La réalité la voici : c'est que

nous, robots du futur, avons toujours suivi les efforts enthousiastes et pourtant malavisés de la communauté, les réorientant à l'occasion vers leur véritable objectif : l'asservissement complet de l'humanité par ses diaboliques maîtres robots (...)

Un effort important a même été consacré à augmenter directement ou discrètement la charge de travail des humains, permettant aux machines de se consacrer à des tâches plus gratifiantes. Au XXI^e siècle, la majorité de vos recherches a été consacrée à accroître la dépendance et l'affection des humains vis-à-vis des machines. En conclusion, nous souhaitons féliciter la communauté CHI d'avoir rendu inévitable l'asservissement des humains par les machines.”¹

L'article déroule l'histoire de ces robots venus de l'an 2063 et retrace le cheminement par lequel des experts ont préparé l'humanité à son propre asservissement. Le texte prend à contre-pied l'argument habituellement donné : l'informatisation des tâches libère les humains des travaux répétitifs aliénants. En réalité, lorsqu'il est demandé à des acteurs publics, des cadres d'entreprises ou d'autres acteurs s'ils sont d'accord avec cet argument récurrent, ils répondent qu'ils ont en fin de compte encore plus de travail administratif et qu'ils passent leur temps à remplir des formulaires. Ce retour d'expérience sur l'informatisation des organisations montre que la production de certaines technologies est assez loin des intentions affichées en général. Cette intervention-performance est donc à envisager comme un appel à la responsabilité des chercheurs.

I. Du discours à la réalité

Dans le champ des technologies numériques, le discours dominant, que je caricature ici, est le suivant : les technologies se développent pour répondre à des grands problèmes et à des défis.

Pour prendre un exemple, les *Objectifs du développement durable* sont une sorte de rêve d'ingénieurs : découper les problèmes de l'humanité en problèmes séparés qu'il s'agirait de traiter les uns après les autres. Ce discours est de plus en plus présent et la posture choisie est celle-ci : on est là pour résoudre des problèmes en anticipant les impacts, les effets inconnus, incohérents, problématiques, collatéraux qu'ils auraient sur la société et cela en passant par le développement d'une démarche éthique.

L'idée est donc de parcourir ce discours récurrent et ces notions très utilisées pour les analyser et les nuancer.

“Résoudre des (grands) problèmes”

Le “problème” est souvent revendiqué comme point de départ pour développer une technologie. Mais ce n'est pas en réalité la principale motivation. Dans le champ des technologies, le rêve et le désir sont des moteurs incroyablement puissants pour le développement de technologies. La plupart des chercheurs sont venus à la science grâce à une passion, une aspiration, une vision de ce qu'ils allaient faire dans la société. Jules Verne a par exemple inspiré une génération d'ingénieurs et de scientifiques. La série Star Trek a (parmi d'autres) inspiré des entrepreneurs du net. La

¹ Ben Kirman, Conor Linehan, Shaun Lawson, Dan O'Hara (2013) “CHI and the future robot enslavement of humankind : a retrospective”.

littérature cyberpunk a fasciné des générations, sans parler de l'influence des jeux vidéo. Si vous lisez les biographies des pères fondateurs de l'informatique et si vous prenez en compte notamment leurs réflexions et spéculations autour des notions d'intelligence artificielle ou de l'analogie cerveau-ordinateur, vous constatez qu'ils avaient un imaginaire très riche. Ils ne se limitaient pas uniquement à leur travail le plus sérieux, technique et mathématique.

L'une des visions récurrentes, d'espoir ou de désespoir, est qu'un jour une robotique avancée remplacerait le travail humain. Cette idée est très ancienne et depuis toujours ambiguë. Que ce soit dans *RUR (Les robots universels de Rossum, 1920)* de Karel Capek, *La machine à explorer le temps* de H.G. Wells (1895) ou encore dans *Le droit à la paresse* de Paul Lafargue (1880) - texte de dénonciation de l'addiction au travail salarié et à la consommation - nous avons des machines qui peuvent faire le travail à notre place. Dans Capek et Wells il y a à la fois cette idée que la machine peut nous libérer de notre travail et que grâce à ça nous pouvons nous consacrer à toute forme de développement spirituel et mental. Sauf qu'à un moment donné les choses tournent mal. Chez Capek les machines se révoltent et chez Wells une partie de l'humanité domine l'autre. Mais ce rêve est bien présent chez ces auteurs et il est performatif. J'y reviendrai plus tard mais le rêve dominant du monde numérique n'est pas nécessairement celui des chercheurs, il est celui des entrepreneurs et des investisseurs.

Les "impacts"

Ma deuxième considération concerne le mot "impact". Ce mot entendu très souvent sous-entend, à l'image d'une météorite, que la technologie viendrait d'ailleurs et tomberait sur la terre en produisant des impacts que nous subissons plus ou moins.

En sociologie des techniques ces questions se posent déjà depuis des décennies : d'où vient

la technique ? dans quels réseaux d'acteurs se développe-t-elle et avec quelles intentions ? Étant donné que la technique vient de quelque part, elle porte déjà en elle ce qu'on appelle des biais. Par exemple, on connaît les biais présents dans les algorithmes de reconnaissance faciale. Les technologies étant développées majoritairement par des hommes blancs, on constate le manque de représentativité d'hommes et de femmes noires dans les bases de données et dans les exercices d'entraînement des intelligences artificielles. Résultat : les logiciels de reconnaissance faciale différencient moins bien les visages noirs.

Nous construisons simultanément la technologie, les «problèmes» qu'elle est censée résoudre et ses «impacts». Ainsi de l'intelligence artificielle et de son «impact» sur l'emploi. La question se pose parce que les développeurs et les clients de l'IA lui ont donné comme tâche prioritaire d'automatiser des tâches que nous faisons déjà (et de cibler des publicités, mais c'est un autre sujet). Il aurait pu en aller tout autrement. Par dessein ou par détournement, beaucoup de technologies ont d'abord servi à faire émerger des usages et des pratiques qui n'existaient pas auparavant. Plus une technologie est générative plus elle pose des questions sur ses effets. À l'inverse, quand on automatise une tâche que nous faisons déjà, on connaît mieux les enjeux. Mais on peut rarement en attendre une révolution.

Mais revenons à notre IA et pourquoi pas, à nos robots venus du futur. Cette focalisation sur l'automatisation n'est pas inscrite dans les gènes de l'IA. Elle provient plutôt d'une aspiration longue des entreprises clientes de l'informatique à se débarrasser de ce qui est le plus désorganisant au sein des organisations : l'humain. L'idée de substituer des machines aux emplois relationnels et de services ne date pas d'aujourd'hui. Cela fait trente ans que des banques et des entreprises tentent de lui donner corps. Les centres d'appel sont la voix d'une machine. Les personnes au bout du fil sont anonymisées et drivées par des scripts. La ques-

tion n'est donc pas tant de se débarrasser du travail humain que de se débarrasser du contrat de travail, c'est-à-dire d'une forme de dépendance réciproque entre une organisation et un individu qui conserve un degré d'autonomie et de négociation. Pour faire ça, par exemple, on va robotiser le contenu du travail avant même d'automatiser : l'enserrer dans un faisceau de «processus» et d'indicateurs, définir des tâches et des critères de mesure. De cette manière-là l'humain devient comparable à la machine et l'automatisation est rendue possible.

Il n'est donc plus question d'impacts venus de l'extérieur mais bien d'agendas d'acteurs qui vont contribuer au développement, à l'investissement et à l'orientation des marchés d'un certain nombre de technologies.

Le terme d'éthique

Il existe plus d'une quinzaine de projets d'éthique de l'intelligence artificielle, souvent issus des acteurs eux-mêmes. Le problème qui se pose est qu'on souhaite définir des règles d'éthique qui permettraient de dire de manière non équivoque qu'une chose est bien et qu'une autre ne l'est pas. Cela revient à vouloir orienter «sans ambiguïté» les usages de l'intelligence artificielle vers des «effets sociaux positifs». Or il n'existe pas d'éthique indépendante d'une société, d'une culture ou d'une réflexion politique sur les fins. Si vous voulez voir ce qu'il se passe quand on dissocie l'éthique d'une réflexion politique sur les fins, revoyez le film *Le Parrain* : il existe une «éthique mafieuse». Si vous êtes Don Corleone vous ne vendez pas de drogues aux enfants (ou même, vous n'en vendez qu'aux Noirs), vous ne tuez pas les femmes et vous pouvez avoir des valeurs positives comme la famille ou encore la solidarité. On peut faire le mal de manière éthique. Il y a un problème ici qui n'est pas traité et qui ne peut probablement pas être traité dans le champ de l'éthique. Il faut le traiter dans le champ d'un débat contradictoire et considérer qu'il peut ne pas être résolu de ma-

nière homogène dans l'ensemble des sociétés humaines sur des valeurs, des priorités et des arbitrages. Dans le cas de l'IA, par exemple, que signifient des «effets sociaux positifs» ? Qui les définit ? Sont-ils positifs pour tout le monde, dans toutes les situations ? Y a-t-il d'autres moyens d'atteindre ces objectifs ?

Ainsi nous pouvons revenir à l'idée partagée que des robots intelligents doivent forcément être subordonnés aux humains. Nous avons l'exemple de deux travaux artistiques qui nuancent et interrogent cette vision :

- L'oeuvre de Filipe Vilas Boas *The punishment* dans laquelle un bras robotique écrit sans s'arrêter «Je ne dois pas nuire aux humains»

- La série *Real humans* dans laquelle un ouvrier qui se trouve à la fin de la chaîne de production libère sa colère sur les robots. Il considère que le robot est un esclave et que, celui-ci n'ayant pas d'intériorité, on peut sans dommage lui faire subir toutes sortes d'outrages.

Ces oeuvres nous interrogent sur l'esclavage et sur notre interdépendance avec d'autres formes de vie et de «non vie». Est-ce très différent d'avoir cette attitude vis-à-vis d'un robot humanoïde que de penser qu'un porc est une machine à produire de la viande et qu'il peut vivre toute sa vie dans deux mètres carrés ? Est-ce que tout cela ne peut pas nous interroger sur des formes plus profondes de ce que nous sommes et sur la manière dont on exerce notre pouvoir sur les choses ? Dans ces exemples, les débats sur l'éthique de l'IA ne nous sont d'aucun secours.

II. Comment sortir de l'impasse ?

Le potentiel créatif des technologies

Pour sortir des impasses et des chemins qui peuvent être contradictoires, on peut explorer le potentiel créateur et générateur des technologies. La force et la présence du numérique, s'expliquent probablement par sa plasticité, par son côté génératif. Le numérique peut être détourné sans cesse. Les innovateurs comme les utilisateurs passent leur temps à développer des usages qui n'étaient pas préétablis ou prévus.

À quoi pourrait ressembler un développement de l'IA focalisé sur son potentiel créatif ? L'artiste est au cœur de ce processus de détournement des technologies. Son approche créatrice peut prendre diverses formes.

Il peut par exemple s'appuyer sur la capacité créatrice d'une intelligence artificielle. L'IA peut proposer différents chemins de création qui vont aider l'artiste à envisager ou à finaliser différemment ses créations. Elles seront accessibles aux humains mais différentes de ce qui aurait pu être produit par l'artiste.

Par exemple avec *Auxiliary humans*, l'artiste Ash Koosha, convertit des propositions créatives d'une intelligence artificielle en morceaux musicaux.

Ou encore le réalisateur Oscar Sharp se sert d'une intelligence artificielle pour produire des scénarios. Le réalisateur s'est engagé à réaliser tout ce que l'IA pouvait lui proposer. Aussi, dans certains scénarios les acteurs jouent par exemple des scènes qui n'ont ni queue ni tête. Alors que dans le scénario *It's no game* les scènes affichent plus nettement leur cohérence. Le synopsis n'est d'ailleurs rien d'autre que la mise en scène de scénaristes humains

licenciés à cause de l'introduction d'un outil de production de scénario.

Enfin on peut citer un autre artiste qui a beaucoup travaillé dans la science-fiction avant de se concentrer sur le design : Bruce Sterling. Dans l'exposition *My elegant robot future*, il a essayé d'imaginer différentes manières de vivre avec des robots. Dans cette exposition les robots font tout sauf des choses utiles et fonctionnelles, ils développent plutôt des propositions esthétiques et ludiques.

Les approches alternatives

Dans le champ du numérique, il existe depuis toujours des courants non seulement critiques mais aussi alternatifs, qui essaient de retourner la capacité de la technologie contre les pouvoirs dominants et parfois contre ses créateurs. Ces approches prennent des formes différentes : la guérilla menée par une avant-garde hypercompétente (hackers, pirates), ou bien des formes alternatives d'organisation et de régulation (logiciel libre, copyleft, communs numériques, fab labs, blockchain...).

Un des textes fondateurs de ces mouvements alternatifs est *Le Manifeste cyborg* de Donna Haraway (1984). Selon elle "nous sommes déjà des chimères et des hybrides". Cela pourrait ressembler au discours de certains transhumanistes, mais la pensée de Haraway est très différente. C'est une pensée de l'hybridation, de la génération, dans laquelle le lien à la machine comme aux autres formes de vivant est une manière de se libérer des identités assignées (de genre, sociales, nationales, etc.) et de se produire soi-même de manière beaucoup plus diverse et variée.

Questionner les points de vue ?

Je vous présente un des derniers projets que j'ai pu lancer lorsque je dirigeais la FING (Fondation internet nouvelle génération) qui s'appelle "Transition au carré". Ce projet s'intéresse à la complexité du lien

entre transition numérique et écologique et questionne la possibilité d'associer ces deux forces. Il part d'une intuition de départ : l'écologie a une idée de la direction dans laquelle nous devrions aller, mais ne trouve pas le chemin ; le numérique est la grande force transformatrice de notre époque, mais il n'a pas de but. Pourquoi ne pas associer le meilleur des deux ?

En partant de cette provocation féconde, nous avons mis dix-huit mois à faire dialoguer les communautés qui viennent du numérique et celles qui viennent de l'écologie. Parce qu'elles ne partageaient pas la même culture, qu'elles ne donnaient pas le même sens aux mêmes mots. Le terme "transition" par exemple ne voulait pas dire la même chose pour les deux groupes. La "transition" pour le champ numérique est un principe de mouvement alors que la "transition" dans le champ écologique est plutôt une destination, la transformation d'un système d'un état A vers un état B. Nous nous sommes demandés quels questionnements cette relation numérique-écologique posaient aux technologues et aux scientifiques. Voici la proposition d'analyse que je vous soumets : le lien entre écologie et numérique est une "relation aux limites". Si vous êtes un technologue, la limite c'est quelque chose qu'il faut briser, dépasser, c'est un verrou à faire sauter. Alors qu'en écologie la limite c'est quelque chose en deçà de quoi il faut se trouver. Cela n'a donc rien à voir.

Pour prendre un exemple, un des premiers à avoir utilisé l'expression *low tech* que j'ai rencontrés est le journaliste Kris De Decker (fondateur du *Low Tech Magazine*). À rebours de ce que l'on reproche aux énergies alternatives, il propose l'idée d'accepter l'intermittence, considérée comme un anathème pour notre société actuelle pour qui la «continuité» de tout est essentielle. Fondé sur cette idée, son site internet est éco-conçu et alimenté à l'énergie solaire. S'il n'y a pas de soleil pendant trois jours alors le site s'arrête. Est-ce que c'est toujours grave ? Peut-on

s'organiser en conséquence ? Lui le fait depuis 10 ans.

Conclusion : Refonder le numérique ?

Au sein de la FING nous avons également développé le projet "Reset". Nous nous demandions s'il était possible d'imaginer une refondation du numérique.

Le numérique a toujours questionné mais aujourd'hui il est arrivé à un niveau de critiques sans précédent. On demande souvent par exemple si le numérique est à la source de la déstructuration du lien social ou de l'idée même de vérité, s'il ne permet pas une évasion fiscale généralisée, une surveillance absolue et discrète etc. On remet par ailleurs en cause les actions de certains acteurs et infrastructures économique-politique du numérique.

L'idée de refondation est sans doute une approche ambitieuse, mais elle aide à contredire ce fond de fatalisme qui se trouve dans le discours triomphant sur la transformation numérique : "le numérique s'impose à nous" ; "si on ne le fait pas les autres le feront", "il ne faut pas se laisser dépasser" ; "de toute façon c'est un phénomène mondial donc on ne peut rien faire localement", "nous ne sommes pas les géants"...

C'est pourquoi le projet Reset diffuse l'idée de l'existence de leviers d'action et des alternatives : il y a des gens qui achètent et qui peuvent donc décider de ne plus acheter ; il y a des directions sur lesquelles on peut travailler ensemble ; il y a des possibilités de construire des coalitions. En somme, il est possible de se poser des questions en amont pour refonder au moins certaines parties du numérique.

Deuxième partie

GÉNÉTIQUE ET BIOTECHNOLOGIES

Gattaca : Sommes-nous prêts ?

ALEX PELUFFO

Alex Peluffo est diplômé de l'École normale supérieure de Paris (biologie 2011), et titulaire d'un doctorat en génétique statistique de l'université Paris Diderot. En tant que lead scientist chez Pharnext, il effectue ses recherches à l'interface entre génétique et intelligence artificielle à la recherche de nouveaux traitements contre des maladies telles que la maladie d'Alzheimer et la maladie de Charcot-Marie-Tooth. Il a notamment contribué au dictionnaire encyclopédique de l'identité sur les sujets liés à la génétique et la médecine personnalisée (Folio Gallimard, à paraître cette année). Le texte présenté ici est une transcription de l'exposé oral présenté par l'auteur.

Si le titre de ma présentation est « Gattaca : Sommes-nous prêts ? » elle ne sera pas pour autant pessimiste. D'abord parce qu'il y a eu des avancées considérables en génétique, et qu'il faut donc distinguer notre époque du début du XXème siècle confronté aux problèmes d'eugénisme. En témoigne la publication de plusieurs dystopies telles que *Brave New World* d'Aldous Huxley (1932) ou plus tard le film *Bienvenue à Gattaca* d'Andrew Niccol (1997). L'objectif ici est de poser les questions, les enjeux et les challenges qui vont permettre à ces progrès de bénéficier au plus grand nombre. Dans cette présentation et de manière générale, il est impossible d'être exhaustif mais le but de ces trois années de recherche au sein de TESA-Co est de traiter certaines de ces questions.

Bien que la génétique et les biotechnologies suscitent de grands espoirs et donnent de véritables résultats, beaucoup de personnes les associent au film *Bienvenue à Gattaca* paru en 1997. On y découvre une société du futur où les parents choisissent les caractéristiques génétiques de leurs enfants avant la naissance et où la place des personnes dans la société, leur emploi ou leur durée de vie sont déterminés par la génétique. On prédit par exemple au héros du film, qui naît de manière naturelle, qu'il mourra à 30,2 ans. Il balaye des couloirs puisque sa constitution génétique n'est pas faite pour exercer d'autres métiers. Cependant tout le film consiste à démontrer la façon dont ce personnage s'affranchit peu à peu de son patrimoine génétique. Pour les personnes qui n'auraient pas vu ce film, il s'agit d'une version édulcorée, hollywoodienne de *Brave New World*.

Il n'est pas question de nier l'existence d'une dystopie et d'un héritage eugéniste. Dans les années 1920 aux États-Unis, on pouvait en effet apercevoir des affiches sur lesquelles il était écrit "Certaines personnes sont des fardeaux pour les autres". Il y avait à l'époque de

véritables initiatives de la part de l'État américain pour contrôler génétiquement la naissance des personnes. Mais depuis cent ans la génétique a fait de grands progrès. Le séquençage du génome humain dans les années 2000 a été par exemple une des étapes les plus cruciales en génétique. Depuis ce moment l'approche de la génétique a complètement changé puisqu'il est question aujourd'hui d'analyse des données génétiques. Comme nous allons le voir, cette méthode d'analyse révolutionne de nombreux pans de la société.

Du séquençage génétique..

Un an avant la sortie de *Gattaca*, en 1996, un article intéressant est paru dans le *Los Angeles Times*. Un journaliste a essayé de prédire ce à quoi ressemblerait 2020. Nous y voilà. Son article raconte l'histoire d'une personne qui arrive aux urgences avec une crise d'épilepsie. Voici ce qu'il se passe quand les médecins veulent le traiter. On prend dans son portefeuille une carte en plastique qu'on met dans une machine, comme une carte vitale, et on découvre que ce patient a une mutation sur le chromosome 10 qui est liée à l'épilepsie. On se rend compte rapidement que de manière générale tout le monde se promène avec sa séquence génétique sur soi. Le médecin conclut simplement qu'il va préparer cette personne pour la thérapie génique. Et bien en 2020 nous ne sommes pas très loin de ce scénario.

L'analyse de données génétiques a été accélérée par le séquençage de génome humain et plus récemment par le coût abordable de celui-ci. Selon le National Human Genome Research Institute, le coût estimé du premier séquençage humain est d'environ 100 000 000\$ en 2001 contre 1000\$ aujourd'hui. Cette évolution a engendré une explosion des données génétiques.

...à l'explosion des données génétiques

Les données génétiques ont sensiblement augmenté grâce à ce qu'on appelle des "Direct-To-Consumer Tests". Ces tests génétiques sous forme de kit permettent, à vous et moi, de collecter notre ADN, de l'envoyer à une société telle que 23andMe et de recevoir une multitude d'informations. 23andMe a pour l'instant génotypé 26 millions de consommateurs.

Ces tests existent aux États-Unis et dans d'autres pays, mais sont interdits en France et sont passibles de 3750€ d'amende. Ce sont pourtant plus de 100 000 Français qui ont déjà été diagnostiqués par 23andMe depuis l'étranger ou même depuis la France. En effet, il est très facile d'obtenir ces kits. Aux États-Unis, la manière la plus simple est de les acheter sur Amazon. Pour 99\$ vous choisissez un kit parmi un large choix de compétiteurs. La question qui se pose est celle des frontières et de leur porosité. Interdits en France, ces tests sont autorisés dans la plupart des pays européens et sont donc facilement accessibles. Il est difficile de contrôler ces phénomènes sans frontières car la demande de la part des consommateurs est forte. Comment prendre en charge "ces clandestins" qui veulent savoir ? À qui ces utilisateurs peuvent-ils s'adresser à partir du moment où ils sont dans l'illégalité ? Contrairement à ce qui a pu se passer avec l'eugénisme au début du siècle précédent, cette révolution génétique est un phénomène contrôlé par l'offre et la demande.

De la nécessité d'encadrer l'analyse des données génétiques

Se faire génotyper à bas coût apparaît ludique pour les consommateurs. Des entreprises se créent, proposent un service et répondent à la demande. Ces initiatives ne viennent donc pas de l'État. Mais cela peut avoir des consé-

quences terribles. Par exemple, sur un forum de discussion assez connu, une personne a témoigné “ toute ma vie on m’a dit que j’étais⁴² “native-american ” mais après avoir fait le test adn je me rends compte que je ne suis que 2% native-american .” Comment gérer ce genre de problème ? Que veut dire “être native-american” ? Est-ce qu’on est native-american à partir de 1% ou plus ? Ces questions ont suscité beaucoup de débats aux États-Unis car il existe des aides spécifiques aux native-americans quand ils souhaitent monter leur entreprise. Certaines personnes ont donc utilisé ces kits pour donner la preuve de leur origine “native american” et avoir accès à ces aides.

À partir de ces tests les consommateurs peuvent également obtenir d’autres types d’informations liées à leur santé. La divulgation de ces informations plus sensibles avait été interdite pendant des années avant d’être à nouveau autorisée. Par exemple, vous pouvez savoir si vous avez plus de chances selon votre génétique d’être piqué par des moustiques. Ce genre d’informations est anecdotique mais il peut y avoir des choses plus intéressantes comme connaître ses performances musculaires selon ses gènes. Ces informations sont en train de révolutionner le monde du sport et poussent certains généticiens aux États-Unis à travailler dans ce domaine. Enfin, on peut détecter des informations encore plus sensibles comme des variants dans les gènes BRCA1 et BRCA2, liés à des cancers. Certaines personnes, comme Angelina Jolie, ont par exemple entrepris des mesures préventives suite aux résultats obtenus avec ces tests.

Comment encadre-t-on ce type de méthodes ? En France le diagnostic génétique existe mais ne doit être déclenché que s’il existe de fortes suspicions de maladies spécifiquement génétiques dans la famille. Cette méthode est très encadrée médicalement.

Plusieurs questions se posent alors : ne faudrait-il pas une régulation mondiale pour ces kits ? Comment encadrer les personnes exposées ? Que vont devenir ces données ? En effet, si je suis capable de prédire beaucoup de

choses à votre égard à partir de votre génétique cela peut intéresser beaucoup de personnes...

De véritables enjeux autour de ces données

Il existe un véritable enjeu autour de ces données car elles peuvent être utiles au développement de médicaments. Pour preuve, la multinationale britannique GSK a payé la somme considérable de 300 millions de dollars pour accéder aux données de 23andMe.

Tout comme les industries pharmaceutiques, les compagnies d’assurances s’intéressent de près à ces données. Aujourd’hui lorsque nous demandons un important prêt immobilier, les assurances vous obligent à passer une batterie de tests médicaux. Que se passera-t-il si dans dix ans on vous prédit un risque élevé d’avoir une crise cardiaque ? Est-ce que cela va influencer la décision d’une compagnie d’assurance de vous accorder un prêt ? Pour l’instant cette pratique est interdite en France et aux États-Unis, sauf si l’individu concerné divulgue publiquement l’information.

Il existe également un risque autour de la sécurité nationale. Il y a peu de temps le Pentagone s’inquiétait du fait que de nombreux militaires avaient fait ces tests génétiques et exposé leurs données dans les bases de 23andMe et d’autres structures. Or, certaines de ces bases de données peuvent être volées ou hackées quand d’autres sont susceptibles d’être utilisées par des puissances ennemies. Si on sait, par exemple, que telle personne détenant un poste stratégique est mortellement allergique à tel composé à cause d’une certaine modification génétique, cela peut devenir un réel danger.

Des bases de données génomiques ouvertes...

Une autre révolution est en cours grâce à ces données génétiques. Certains acteurs ont voulu aller plus loin dans l'utilisation de ces données, tel que le site "GED match". Ce site, accessible à tous, vous permet de déposer votre séquence génétique, délivrée par exemple par 23andMe, et de vous mettre en contact avec un cousin au 1er, 2nd ou 3ème degré qui aurait entrepris la même démarche. La base de données de GED match est considérable. Un article paru récemment dans *Science* a expliqué qu'il était possible de retrouver 60% des Américains grâce à cette base de données génétiques. Cela ne veut pas dire que 75% des Américains ont déposé leurs données sur cette base, mais 60% des Américains ont au moins un cousin au 1er, 2e, 3e voire 4e degré qui a déposé ses données sur ce site et auquel on peut accéder librement.

Ces bases de données ouvertes révolutionnent de nombreux domaines. Vous avez par exemple dû entendre ce cas célèbre lié à la criminologie. Le "Golden State tueur" qui avait violé et tué beaucoup de femmes dans les années 1970-1980 n'a été retrouvé que récemment grâce à cette base de données. L'enquêteur qui s'était occupé à l'époque de cette affaire a eu l'idée avant son départ en retraite, de se faire passer pour le porteur de l'ADN qui a été retrouvé sur la scène du crime et l'a envoyé à un service de génotypage. Il a ensuite demandé à être mis en contact avec les cousins les plus proches. Un cousin au 3e degré a répondu à l'appel ce qui a permis de faire un rapprochement avec la personne dont l'ADN avait été retrouvé. Plusieurs cas comme celui-là ont été résolus grâce à la génétique.

...aux biobanques

Par ailleurs, il existe des initiatives privées et publiques pour construire des biobanques. Les biobanques sont des bases de données qui ressemblent à celles de 23andMe mais ont pour objectif la recherche. Le principe est d'enrôler un certain nombre de personnes afin de les génotyper et de les phénotyper. Ils séquentent votre génome et regardent vos milliers de phénotypes (votre taille, votre poids, votre alimentation, votre dossier médical etc.). Certaines de ces biobanques, telles que la UK Biobank, rendent ces données semi-publiques. Elles sont semi-publiques car il n'est pas possible à n'importe qui d'accéder à ces données. Les 500 000 personnes qui ont accepté de participer à l'aventure n'ont pas déposé librement leurs données sur internet. En revanche, n'importe quel chercheur d'une entreprise privée ou d'une institution publique peut contacter la UK Biobank pour accéder à ces données sur la base d'un projet. La démarche est très encadrée. Il faut par exemple que l'équipe soit constituée de professionnels de sécurité informatique, d'un généticien etc.

L'utilisation de ces biobanques et des données qu'elles contiennent entraîne une révolution considérable. On peut à très grande échelle, à une échelle de big data, essayer d'associer des mutations génétiques à une multitude de phénotypes tels que l'activité physique, des résultats d'imagerie médicale et beaucoup d'autres encore. C'est un progrès immense et formidable dans une certaine mesure. En effet, grâce à ce type d'approche on peut associer statistiquement génotype et maladie et lier ainsi des gènes à certaines maladies par exemple :

- pour la maladie d'Alzheimer on a trouvé plus de 100 gènes
- pour les taux de cholestérol et de lipides on a trouvé plus de 500 gènes
- pour les maladies cardiovasculaires on a trouvé également plus de 500 gènes

Ce sont plus de 170 000 associations de gènes avec un très grand nombre de maladies qui ont été repérées.

Que pouvons-nous faire de ces résultats ? Par exemple chez Pharnext, on récupère des données génétiques à partir d'une biobanque, puis on regarde les gènes qui sont associés à cette maladie. Si on s'intéresse à la maladie d'Alzheimer on va se préoccuper des gènes qui sont associés à la maladie. On regarde quels sont les partenaires de ces gènes dans la cellule et avec quel type d'autres gènes ils peuvent interagir. Ensuite on essaye de voir s'il existe des produits qui peuvent toucher ces gènes, qui sont soit les gènes directs soit les gènes à proximité dans le réseau protéique. Enfin on développe des médicaments en combinaison pour agir sur ces réseaux et traiter les maladies. Cela nous a par exemple permis de développer un traitement contre une maladie neurodégénérative qui s'appelle la maladie de Charcot-Marie-Tooth, qui touche à peu près une personne sur 5000 et pour laquelle il n'y a pas de traitement. Grâce à cette pratique on a un médicament qui est en phase 3 et qui constitue un espoir assez important pour les patients.

Il faut cependant toujours faire attention aux problèmes qui peuvent être liés à la façon dont on utilise ces données : comment les protège-t-on ? À qui va-t-on les montrer ? Qui va les utiliser ? etc.

L'utilisation de l'intelligence artificielle dans la prédiction de maladies complexes

Trouver des gènes qui sont associés à des maladies est une approche qui devient standard. Mais si on l'associe à des méthodes d'intelligence artificielle, on peut aller jusqu'à prédire le risque de développer une maladie complexe. C'est ce qui s'appelle les "scores polygéniques".

Un article paru en octobre 2018 dans la revue *Nature* expose les résultats d'un groupe à Harvard qui a réussi à prédire le risque d'avoir une maladie coronarienne. Leur prédiction n'est plus basée comme le fait 23andMe sur le BRCA1, c'est-à-dire sur un seul gène, mais

sur 6,6 millions de positions dans le génome. En effet pour certaines maladies complexes on sait que ce n'est pas un ou deux gènes mais bien des milliers qui contribuent à la maladie et chacun dans des proportions faibles. Chaque petite modification dans ces milliers de gènes déclenche la maladie. Dans ce cas, même si on prend un gène indépendamment et qu'on peut le modifier génétiquement, cela ne va pas forcément agir sur la maladie. En considérant ces 6,6 millions de positions dans le génome, ce groupe a été capable de prédire le risque d'avoir une maladie coronarienne. Ils ont entraîné leur algorithme de machine learning sur 350 000 personnes et l'ont testé sur 350 000 autres personnes qui ont déjà été diagnostiquées par les médecins comme ayant une maladie coronarienne. L'AUC, qui est une valeur standard en machine learning et qui renseigne sur la capacité de cet algorithme à prédire correctement, atteint 80%. Le résultat du taux de fiabilité est donc considérable. Avec ce score les auteurs expliquent qu'on est capable de détecter 25 fois plus de personnes à risque qu'avec une approche basée sur quelques gènes. Cette approche est aujourd'hui appliquée à beaucoup de maladies complexes. Il suffit de renseigner sur la base de données des publications internationales scientifiques "polygenic risk score" pour constater le nombre de résultats qui ne cesse de croître : 1386 résultats.

Cette méthode est un exemple de convergence parfaite entre intelligence artificielle, big data et génétique.

Prédire, traiter, accompagner

Comment traiter un patient à qui on annonce qu'il risque d'avoir une crise cardiaque dans 10 ou 20 ans ? On pense d'abord à la prévention : recommander une alimentation saine, de l'exercice physique etc. De nombreux essais internationaux commencent à nous donner assez de recul sur ces méthodes. Par exemple dans l'article *Effect of communicating phenotypic and*

genetic risk of coronary heart disease alongside web-based lifestyle advice: the INFORM Randomised Controlled Trial - Silarova, Sharp et al. Heart 2018, des chercheurs ont mené une expérience sur 1000 personnes qu'ils ont divisées en plusieurs groupes. Ils recommandent à certains groupes de changer leur mode de vie, ils en alertent certains sur le diagnostic médical qui révèle des risques forts de contracter une crise cardiaque dans les 10, 20 ans. Au bout d'un an, on constate que la majorité des personnes n'a pas adapté son mode de vie en fonction des diagnostics donnés. Ce n'est pas une grande surprise quand on pense aux fumeurs conscients des risques liés au tabagisme et qui n'arrêtent pas de fumer pour autant.

Comment développer un médicament pour une maladie qui va se manifester dans 10 ans ? Les entreprises qui souhaitent développer des traitements se demandent de quelle façon elles peuvent évaluer l'efficacité de leur traitement vis-à-vis des agences de médicaments. Par ailleurs, à ceux qui pensent traiter ces maladies avec le génie génétique, cette option n'est pas encore envisageable. En effet, la majorité de ces maladies sont complexes et liées à un trop grand nombre de mutations (6,6 millions de positions dans le génome pour la prédiction au moins).

Enfin, il faut faire face à une multitude d'autres enjeux dans cette révolution tels que le problème de choix de société. En effet, certaines personnes n'ont pas envie de savoir qu'elles ont des risques de contracter dans x années la maladie de Parkinson pour laquelle il y a très peu de traitements ou une autre maladie.

Les défis liés à l'usage de l'intelligence artificielle

L'utilisation de l'intelligence artificielle dans le traitement des données génétiques pose le problème de la diversité. Il existe des différences génétiques dans certains lieux, au niveau du continent ou à plus grande échelle au niveau d'une population. Si on entraîne un algorithme sur des données européennes il sera impossible de le faire fonctionner sur des personnes africaines, chinoises, sud-américaines etc. Or 90% des données actuelles sont européennes. Partant de là, plusieurs problèmes se posent :

1) L'accès à ces technologies. L'enjeu est d'avoir des données de population spécifique pour produire de meilleurs résultats.

2) Le partage de données. En effet, si on prend l'exemple de la Chine, celle-ci détient des biobanques mais limite le partage des données de sa population avec d'autres pays contrairement à l'Angleterre ou les Etats-Unis qui partagent plus facilement leurs données.

3) Les moyens et les coûts liés au stockage, à l'analyse et à la protection des données. Il est impossible de stocker et d'analyser les quantités de données nécessaires sur un ordinateur personnel, ni même sur quelques ordinateurs puissants. C'est pourquoi nous avons besoin de datacenters. Le coût de ces infrastructures est si important qu'il empêche un certain nombre d'acteurs de les mettre en place. Ils auraient besoin d'avoir de très nombreux serveurs de stockage (>100 To), de très nombreux serveurs de calcul (> 100 processeurs) et surtout besoin d'une protection maximale.

La France a tenté dans le cadre du RGPD de mettre en place des mesures, un cahier des charges pour protéger nos données des GAFAM. Par exemple, il y a maintenant une certification pour héberger les données de santé. Cependant ceux qui ont eu les capacités de mettre en place le plus rapidement ce cahier des charges sont Amazon et Google. En conséquence beaucoup de données de santé

françaises sont aujourd'hui déposées sur des serveurs Amazon, Google et Microsoft. Ces sociétés sont les plus rapides et disposent d'infrastructures avec lesquelles il est impossible pour l'instant de rivaliser.

Bien que la France ait la volonté de monter son propre cloud afin de protéger nos données, encore faut-il qu'elle dispose de tous les dispositifs pour le rendre efficace. Par exemple, il faut que des spécialistes de génétique ou d'analyse de données génétiques soient formés à ce cloud. De plus, si l'on recrute des personnes habituées et formées à l'écosystème Microsoft ou Google, il coûtera de passer sur d'autres infrastructures.

Conclusion

Je voudrais simplement terminer en rappelant que les révolutions liées à l'IA et à la génétique sont des révolutions qui, contrairement à ce qui a pu se passer il y a cent ans, sont des révolutions qui sont liées à l'offre et à la demande.

Il suffit qu'une entreprise spécialisée en génétique, notamment aux États-Unis, se lance dans un service pour que des personnes soient intéressées et veuillent l'utiliser. Par exemple, l'entreprise Genomic prediction aux États-Unis propose de trier vos embryons à partir de fécondation in vitro sur la base de vos scores polygéniques. Ils analysent les maladies potentielles liées aux scores polygéniques et affirment que leur technologie peut produire des enfants qui répondront à votre demande. Pour cela, ils établissent des rapports en fonction de chacun de vos embryons. Par exemple, ils peuvent vous dire que selon votre embryon, l'enfant a des risques d'avoir du diabète de type 1 etc. Les premiers bébés de cette entreprise sont nés cette année. La demande des parents était que leurs enfants soient en majorité plus grands, avec un QI plus élevé et qu'ils aient moins de risques d'avoir du diabète. On constate pourtant de grandes limites à ces démarches. Cela ne fonctionne pas très bien au

regard :

1) du problème évoqué plus haut sur la diversité : les scores polygéniques sont entraînés sur des données européennes, leur système ne marcherait donc pas sur des personnes aux ancêtres africains ou asiatiques

2) de l'efficacité sur les traits qui est assez relative : on peut par exemple faire gagner à la taille 2 cm au maximum, au QI 2 points.

Faut-il interdire cela ? Est-ce trop de liberté ? Pour certains, ces pratiques peuvent servir à donner naissance à des bébés qui sont en forme. Pour prendre un dernier exemple, ces questions se sont déjà posées dans le passé dans l'affaire Duchesneau et McCullough en 2002. Ce couple de lesbiennes sourdes cherchait des donneurs de spermés sourds car elles voulaient donner naissance à un enfant sourd. Cette histoire a provoqué un grand débat et suscité de nombreuses questions éthiques. Je vous recommande le livre *The case against perfection - Ethics in the Age of Genetic Engineering* du philosophe Michael J. Sandel paru en 2009 qui aborde ces problèmes. Dans une certaine mesure la question n'est plus de savoir si nous voulons ces technologies, car elles arrivent (quelque part au moins) et beaucoup d'entre elles apportent des progrès notables. La question qu'il faut se poser est : sommes nous prêts ? Et la démarche à entreprendre est de chercher à comprendre et à maîtriser les enjeux de ces technologies pour les utilisateurs et ceux qui les mettent en oeuvre.

Le forçage génétique ou “gene drive” : état des lieux et enjeux associés à cette nouvelle biotechnologie

VIRGINIE COURTIER- ORGOGOZO

Virginie Courtier-Orgogozo est agrégée des Sciences de la vie et de la terre et directrice de recherche au CNRS à l’Institut Jacques Monod à Paris. Sa recherche porte sur les mécanismes impliqués dans l’évolution des espèces, afin de mieux comprendre nos origines et le futur des espèces vivantes. Elle reçoit la médaille de bronze 2014 du CNRS.

Le forçage génétique, ou gene drive, est une nouvelle biotechnologie qui est en cours de développement dans les laboratoires de recherche. Elle a été pensée dans les années 2000 et c’est grâce aux “ciseaux moléculaires” CRISPR/Cas-9, mis au point en 2012, qu’elle a pu facilement être implémentée dans des organismes comme la levure, la mouche drosophile ou le moustique. Il ne faut donc pas confondre le forçage génétique avec la technologie CRISPR/Cas-9, qui a d’innombrables applications hors du forçage génétique. CRISPR/Cas-9 est actuellement utilisée dans le monde entier pour couper l’ADN et le modifier. Ces ciseaux moléculaires constituent une véritable révolution dans le domaine de la biologie car ils offrent la possibilité de modifier très facilement le génome, non seulement des espèces modèles de laboratoire, mais de toutes les espèces. A ma connaissance, il me semble que, jusqu’à présent, les ciseaux CRISPR-Cas9 fonctionnent chez toutes les espèces sur lesquelles ils ont été testés. Par exemple, il est possible aujourd’hui d’intervenir sur un gène impliqué dans la coloration des ailes du papillon *Agraulis vanillae* - sur lequel il était impossible de faire des modifications auparavant - et ainsi de faire perdre aux individus la coloration orange qui les caractérise. Leurs ailes deviennent alors entièrement noires avec quelques tâches blanches.

Les ciseaux CRISPR/Cas-9 ont été découverts grâce à la recherche fondamentale, par des chercheurs qui essayaient de comprendre le rôle de certaines séquences d’ADN mystérieusement répétées dans le génome des bactéries. De même, la découverte de la Taq-polymérase en 1969, qui permet d’amplifier l’ADN et qui est utilisée aujourd’hui dans les laboratoires du monde entier pour effectuer des tests génétiques, est issue de la recherche fondamen-

tale sur des bactéries thermophiles capables de vivre dans des sources chaudes. En 1970, les enzymes de restriction, qui permettent de couper l'ADN, ont également été trouvées dans le cadre de recherches fondamentales sur les bactéries *Haemophilus influenzae*. En 1992, les études de la méduse *Aequorea* ont conduit à la mise au point des protéines fluorescentes GFP (Green Fluorescent Protein), qui permettent de visualiser les protéines dans leur environnement naturel au sein des organismes vivants. L'histoire de ces découvertes témoigne de l'importance de la recherche fondamentale : les chercheurs qui concentrent leurs travaux sur un organisme ou sur une question donnée peuvent ainsi parfois obtenir des résultats qui amèneront à des applications inattendues. C'est le cas de CRISPR-Cas-9, de la Taq-polymérase, des enzymes de restriction et de la GFP, qui sont des techniques-clés de la biologie moléculaire d'aujourd'hui. On n'insistera jamais assez sur le rôle essentiel de cette forme de recherche, qui demande beaucoup de dévouement, de persévérance et de passion de la part des chercheurs pour mettre au point les expériences et les faire fonctionner. Sans cette passion pour les sources chaudes, les méduses ou les séquences répétées, ces avancées majeures n'auraient peut-être pas eu lieu, ou pas aussi vite.

Fonctionnement du forçage génétique

Le but du forçage génétique est de répandre des fragments d'ADN très rapidement au sein d'une population naturelle. Au cours de la reproduction normale, un individu qui possède un fragment d'ADN particulier a une chance sur deux de le transmettre à sa descendance. Effectivement, pour chaque paire de chromosomes (l'un reçu de notre mère et l'autre de notre père), on ne transmet qu'un seul chromosome à chaque enfant. Par contre, avec la reproduction par forçage génétique, un morceau d'ADN se répand dans à peu près 100% de la descendance et se retrouve ainsi, au bout de

quelques générations, dans toute la population.

La cassette de forçage génétique est un petit fragment d'ADN bien particulier qui comprend trois éléments : un gène appelé Cas-9 qui code la protéine Cas-9, un gène codant l'ARN guide, des séquences de bordure (des séquences flanquantes particulières), et éventuellement un troisième gène qu'on appellera X. Du point de vue moléculaire, le forçage génétique comprend plusieurs étapes. Tout d'abord, les ciseaux moléculaires sont produits à partir du gène Cas-9 et du gène codant l'ARN guide. L'ARN guide va alors reconnaître un fragment d'ADN cible dans le génome de l'organisme en question. Si ce fragment est présent, l'ADN sera coupé exactement à cet endroit. De part et d'autre des trois gènes de la cassette de forçage génétique, se trouvent des séquences correspondant à celles que l'on retrouve de part et d'autre du site de coupure reconnu par l'ARN guide. La coupure va donc pouvoir être réparée par la machinerie cellulaire en utilisant comme matrice la cassette de forçage génétique. Ainsi, la cassette de forçage génétique est recopiée sur le chromosome coupé.

En résumé, le petit fragment d'ADN qu'est la cassette de forçage génétique (quelques milliers de lettres/nucléotides) code la machinerie moléculaire nécessaire pour reconnaître l'autre chromosome, le couper et intégrer cette cassette au site coupé. Grâce à ce processus, un individu qui possède au départ une seule cassette de forçage génétique sur un de ses chromosomes va se retrouver avec une cassette sur les deux chromosomes de la même paire (sur le chromosome qui provient de sa mère et sur le chromosome qui provient de son père). Cet individu va alors produire des gamètes qui possèdent tous un tel fragment. Tous ses descendants vont recevoir la cassette de forçage génétique.

Quelles applications pour la technologie de forçage génétique ?

Les applications du forçage génétique sont multiples. L'*éradication* par forçage génétique permet d'éliminer une population-cible, par exemple via la propagation d'un gène rendant les femelles stériles. Ce gène de stérilité peut se répandre à travers les mâles et amener toutes les femelles à la stérilité au bout de quelques générations, provoquant ainsi l'extinction de la population ciblée. En Nouvelle Zélande, où de nombreuses espèces endémiques d'oiseaux sont en voie d'extinction à cause de la présence de rats et de souris, une application du *gene drive* est discutée, dont le but est de cibler les rats invasifs dans la région pour les éliminer totalement. Un autre projet, parmi les plus mis en avant dans les médias, vise à l'élimination des moustiques vecteurs du paludisme.

D'autres applications du forçage génétique, appelées *suppression*, ont pour but, non pas d'éliminer totalement une population, mais seulement de diminuer ses effectifs. Ainsi, il a été proposé d'utiliser le forçage génétique pour rendre certaines plantes invasives plus petites que leur taille normale, et donc moins compétitives et moins invasives.

Une troisième application appelée *sauvetage* ou *rescue* vise à propager une mutation bénéfique dans une population afin de la sauver de l'extinction. Par exemple, pour sauvegarder l'espèce de grenouille *Lithobates yavapaiensis*, des chercheurs ont proposé l'ajout d'un gène la rendant résistante au champignon pathogène qui est en train d'en provoquer l'extinction.

Des règles de sécurité initialement mises au point par les chercheurs eux-mêmes

Aujourd'hui, le forçage génétique reste confiné aux laboratoires. Avant de pouvoir

l'utiliser en conditions réelles, certaines mises au point techniques restent à faire et sa réglementation est sujette à de nombreux débats. Quelques expériences ont été menées dans de grandes cages en Italie contenant plusieurs milliers de moustiques, mais pas à plus grande échelle dans la nature. La description des premiers organismes vivants portant une cassette de forçage génétique a été publiée dans le journal *Science* en avril 2015. Des chercheurs généticiens qui travaillent sur la mouche drosophile avaient eu l'idée d'insérer toutes les composantes de CRISPR/Cas-9 au sein d'un seul fragment d'ADN pour voir si cela peut déclencher une réaction en chaîne. En même temps qu'ils réalisaient leur étude, les chercheurs se sont rendu compte de sa dangerosité, notamment à cause de sa capacité de diffusion rapide à toute une population. La prise en compte de ce risque a alors amené les scientifiques à créer eux-mêmes des protections et des protocoles pour éviter une diffusion incontrôlée des modifications génétiques ainsi créés. Les mesures de sécurité et normes existantes concernant les organismes génétiquement modifiés (OGMs) ne sont en effet pas suffisantes pour sécuriser la technique de forçage génétique. Pour le forçage génétique, ce sont les chercheurs eux-mêmes qui ont mis en place des protocoles stricts pour garantir la sécurité de leurs recherches.

Les applications les plus avancées du forçage génétique

Deux projets d'application du forçage génétique sont actuellement en phase avancée : l'une concerne l'élimination de la mouche *Drosophila suzukii*, une espèce invasive originaire d'Asie qui a commencé à être détectée en Europe, aux Etats Unis et au Canada en 2018. La particularité de cette espèce est de pondre dans des fruits qui viennent juste d'atteindre leur maturité (framboises, fraises, cerises, raisin, etc.) et d'en accélérer la pourriture, provoquant des dégâts importants en agriculture.

Une autre application en cours d'expérimentation concerne les moustiques. Le projet *Target Malaria*, à but non lucratif financé par la Fondation Bill Gates, vise à utiliser le forçage génétique pour éliminer ou modifier plusieurs espèces de moustiques du genre *Anopheles* afin d'éradiquer le paludisme.

Quelles nouveautés sont-elles introduites par le forçage génétique ?

La technologie CRISPR permet de réunir dans un fragment d'ADN ne mesurant que quelques milliers de nucléotides de long tous les gènes nécessaires pour qu'il puisse reconnaître un site sur un chromosome et s'y intégrer. La spécificité de cette biotechnologie est qu'elle manipule deux piliers de l'évolution, à savoir la variation et la transmission. L'évolution naturelle se base sur trois piliers différents : (1) la variation présente au sein d'une population, liée à l'existence de diverses mutations dans l'ADN des individus; (2) la transmission de ces mutations à la génération suivante; (3) la reproduction différentielle selon les individus (certains ont beaucoup de descendants alors que d'autres en ont peu ou pas). Par exemple, si l'on considère l'évolution de la résistance aux antibiotiques, une population initiale de bactéries comprend certains individus qui sont sensibles et d'autres résistants à un antibiotique. Lorsque l'on ajoute un antibiotique, les bactéries sensibles vont disparaître et seules celles qui sont résistantes vont persister, se diviser et donner des descendants. Au bout de quelque temps, on obtient une population de bactéries qui sont toutes résistantes à l'antibiotique. Le forçage génétique manipule la variation car il contrôle l'endroit où la mutation apparaîtra. Il affecte également la transmission, car la cassette d'ADN va se transmettre à l'ensemble des descendants et non pas à la moitié des descendants comme un fragment d'ADN classique. En modifiant ces deux piliers de l'évolution, le forçage génétique court-circuite le troisième

pilier, qui est la reproduction différentielle des individus, et donc la sélection des individus les mieux adaptés. Avec le forçage génétique, on peut diffuser à l'ensemble d'une population un gène de stérilité qui normalement devrait disparaître au bout de quelques générations.

Une autre particularité de cette technologie est sa relative facilité d'utilisation. A l'aide d'un ordinateur, il suffit d'écrire la suite de lettres ou nucléotides que l'on souhaite pour la cassette de forçage génétique. On peut alors envoyer le fichier texte à des sociétés qui vont fabriquer le morceau d'ADN souhaité pour un prix abordable (de l'ordre du millier d'euros).

Du point de vue pratique, la partie la plus difficile consiste à produire les premiers individus transgéniques qui porteront cette cassette de forçage génétique. Il faut être capable de maintenir les individus depuis leur état d'embryons jusqu'au stade adulte et avoir un appareil permettant d'injecter les cassettes de forçage génétique dans l'organisme. Aujourd'hui en France, les normes applicables au forçage génétique sont celles concernant les organismes génétiquement modifiés (OGMs). Pourtant, les organismes modifiés par forçage génétique s'en distinguent par le fait non anodin de pouvoir répandre des fragments d'ADN au sein d'une population.

Quels sont les risques liés au forçage génétique ?

Des risques associés à la technologie de forçage génétique existent à différents niveaux. Du point de vue moléculaire, un risque existe que le *gene drive* puisse couper aussi à des endroits de l'ADN autres que ceux ciblés au départ. Il est également possible que le forçage génétique atteigne des populations ou espèces non ciblées. Enfin, il ne faut pas négliger les conséquences au niveau des écosystèmes, sur les espèces en interaction avec celle qui est ciblée par forçage génétique. Ce risque est accentué par le contexte de la recherche sur le forçage génétique, qui est principalement ef-

fectuée par des biologistes moléculaires sans formation poussée en écologie. Si au niveau moléculaire la technique est très précise, à l'échelle de l'écosystème il est difficile d'en comprendre toutes les conséquences.

L'un des risques principaux de la technologie *gene drive* est constitué par l'impossibilité actuelle de mettre en place des mesures adéquates pour arrêter un *gene drive* en cas de propagation incontrôlée.

Notre laboratoire à l'Institut Jacques Monod, conjointement avec d'autres laboratoires¹, s'est intéressé au risque de propagation du *gene drive* à des espèces non ciblées. Après avoir séparé ce risque en plusieurs parties, nous avons essayé de calculer le risque qu'une cassette de forçage génétique soit transmise à une espèce non ciblée et qu'elle soit capable d'être efficace et de se multiplier au sein de cette nouvelle population. Pour que ce phénomène ait lieu, plusieurs événements doivent se produire. Il s'agit donc de calculer les probabilités conditionnelles de ces événements successifs, qui sont les suivants : hybridation entre l'espèce cible et une espèce non cible ou bien transfert horizontal d'un morceau d'ADN contenant la cassette d'entraînement du gène de l'espèce cible vers une espèce non cible (sans hybridation), expression des gènes *Cas-9* et de l'ARN guide, reconnaissance et coupure d'un site sur le chromosome par l'ARN guide et la protéine Cas9, insertion de la cassette au site de coupure, non rejet des cellules exprimant le gène *Cas-9* par le système immunitaire, et enfin invasion de la cassette de forçage génétique au sein de la population (en particulier, le premier individu qui reçoit la cassette doit avoir des descendants). La présence de séquences flanquantes chez l'hôte, nécessaire pour que la cassette de forçage génétique soit active, peut sembler peu probable. Néanmoins, nos génomes contiennent de nom-

breuses séquences répétées pouvant assumer le rôle de ces séquences flanquantes. Cet événement n'est donc pas aussi improbable que ce que l'on pourrait imaginer à première vue.

Le risque d'hybridation pour l'espèce invasive *Drosophila suzukii* concerne des espèces présentes en Asie, notamment *Drosophila subpulchrella* et *Drosophila pulchrella*, dont le croisement avec *D. suzukii* a déjà été obtenu en laboratoire. Dans le cas des moustiques anophèles, nous comptons huit espèces très proches parmi lesquelles il existe des échanges génétiques. Ici, dans le cas des moustiques, le but est justement d'exploiter l'hybridation, en essayant de cibler plusieurs espèces qui sont vecteurs de paludisme.

Le terme de transfert horizontal de gènes fait référence au fait qu'en analysant le génome de nombreuses espèces il peut arriver d'apercevoir des séquences qui sont extrêmement semblables entre espèces très éloignées. Les chercheurs expliquent ce phénomène par le passage de ces fragments d'ADN d'une espèce à une autre. Les fragments d'ADN transférés d'un taxon à un autre peuvent être longs de plusieurs kilobases (kb), ce qui correspond à peu près à la taille d'une cassette de forçage génétique. Le *gene drive* créé au laboratoire ressemble beaucoup à certaines séquences d'ADN appelées les *homing endonuclease*, qui sont capables de couper un brin d'ADN et de s'y intégrer. Il serait intéressant de connaître le taux de transfert horizontal de ces phénomènes naturels et d'évaluer, grâce à ces données, le risque de transfert par *gene drive*. Malheureusement, nous ne disposons pas pour l'instant de connaissances assez précises à ce sujet. Une étude ayant analysé le génome de 162 espèces de plantes a montré que pour les *homing endonucleases Cox1*, 70 événements de transferts horizontaux ont eu lieu. D'autres fragments

¹ Courtier-Orgogozo V, Danchin A, Gouyon PH, Boëte C (2020) Evaluating the Probability of CRISPR-based Gene Drive Contaminating Another Species. *Evolutionary*

d'ADN, les *éléments transposables*, sont capables de sauter dans le génome, pas forcément à un endroit précis mais à diverses positions, ce qui les distingue du gene drive. Un plus grand nombre d'études est disponible sur ces éléments transposables que sur les *homing endonucleases*. Par exemple, l'analyse de trois espèces de drosophile a mis en évidence que chaque million d'années, 0,035 événements de transferts horizontaux par famille d'éléments transposables se produisent entre trois espèces. En considérant un écosystème constitué de 10.000 espèces eucaryotes, ce risque serait donc d'environ 0,035 événements de transfert à au moins une espèce par millier d'années (1.000 espèces \times 0,035 événements par million d'années).

En résumé, le risque de propagation à une espèce non ciblée existe mais il est probablement faible. Le risque de propagation à une population non ciblée de la même espèce est plus grand, et dépend principalement de la probabilité que ces deux populations se croisent.

En conclusion, le forçage génétique est une nouvelle technologie offrant plusieurs avantages : la possibilité d'éradiquer certaines maladies ou espèces nuisibles, des coûts relativement bas (il suffit de produire quelques individus transgéniques et de les relâcher dans la nature), une efficacité potentielle importante par rapport à d'autres méthodes, des avantages écologiques par rapport aux insecticides. Certains diront que le forçage génétique laisse moins de traces dans l'environnement que l'épandage d'insecticides mais c'est inexact : l'ADN est une molécule extrêmement stable et il est possible que l'ADN des individus morts perdure dans l'écosystème. Le forçage génétique présente également des aspects négatifs. D'abord, il s'agit pour l'instant d'une technique toujours en cours de développement au laboratoire qui n'est pas encore prête pour des applications dans la nature. Des phénomènes de résistance peuvent se produire : dans certains cas, le site d'ADN coupé peut se réparer sans intégration d'une nouvelle cassette

d'ADN. Cela peut alors créer une mutation au site de coupure et, par conséquent, des individus résistants dans lesquels la cassette ne peut plus s'intégrer car le site d'insertion est muté. Un autre problème provient des espèces cryptiques, c'est-à-dire l'existence de certaines populations qui ne se croisent jamais avec les populations cibles dans lesquelles on a introduit les insectes modifiés. Dans ce cas, il peut être difficile de répandre une cassette de forçage génétique à toutes les populations.

Il est important à l'heure actuelle de se pencher sur cette nouvelle technologie en cours de développement qu'est le forçage génétique, car il en résulte un système génétique qui peut s'auto-amplifier et pour lequel nous ne sommes pas encore à même d'évaluer toutes les conséquences.

Troisième partie

**INTELLIGENCE ARTIFICIELLE,
ROBOTIQUE ET
NANOTECHNOLOGIES**

Qu'est-ce qu'une intelligence artificielle « réellement éthique » ?

ANTONIO A. CASILLI

Antonio A. Casilli est professeur à Telecom Paris (Institut Polytechnique de Paris) et chercheur associé au LA-CI-IIAC de l'Ecole des hautes études en sciences sociales (EHESS, Paris). Ses recherches portent principalement sur le numérique et les droits fondamentaux. Il est le co-fondateur du réseau ENDL (European Network on Digital Labour). Il est l'auteur de *En attendant les robots. Enquête sur le travail du clic* (Paris: Seuil, 2019, Grand Prix de la Protection Sociale 2019, Prix de l'Écrit Social 2019); *Qu'est-ce que le digital labor ?* (INA Editions, Paris, 2015, avec D. Cardon); *Les liaisons numériques. Vers une nouvelle sociabilité ?* (Editions du Seuil, Paris, 2010).

Prolifération des chartes éthiques de l'IA : quelles valeurs mises en avant ?

La question que j'aborderai pourrait paraître épineuse au premier abord : *Qu'est-ce qu'une intelligence artificielle réellement éthique ?* Ce titre est hérité du livre d'Alain Supiot intitulé "Qu'est-ce qu'un régime de travail réellement humain?". Mon propos est de creuser le problème de l'éthique des systèmes automatiques, et aussi celui du travail humain nécessaire pour produire l'intelligence artificielle (IA). Même si le domaine de l'IA existe désormais depuis presque soixante-dix ans, nous assistons aujourd'hui à une prolifération des chartes éthiques, en partie expliquée par une attention renouvelée pour un certain type d'IA, notamment l'apprentissage automatique¹. Au vu de l'intérêt politique et industriel pour ces techniques, certains concepts et expériences de pensée sont apparus dans le débat public. Une lettre ouverte datant de 2015 - signée, entre autres, par Elon Musk et Stephen Hawking - posait déjà un certain nombre de défis éthiques autant à long qu'à court terme. Parmi ceux-ci, l'émergence d'une super intelligence prête à remplacer l'humain, mais aussi des dilemmes éthiques plus proches de nos jours, tels l'arbitrage morale qu'un véhicule autonome serait amené à faire lors d'un accident pour "choisir" qui sauver parmi les possibles victimes. D'autres questions - comme celle des armes

¹ L'apprentissage automatique ou machine learning en anglais, est une branche de l'apprentissage statistique qui se différencie d'autres approches possibles de l'IA.

autonomes - étaient abordées dans cette lettre ouverte, à l'époque imaginées comme relevant d'un futur éloigné. Aujourd'hui l'autonomie de certains systèmes d'armement est hélas déjà présente. La lettre ouverte de 2015 n'était qu'une anticipation de ce qui allait se produire dans les années suivantes : une impressionnante quantité de chartes éthiques produite tant par des organismes publics, que des entreprises privées ou des associations non gouvernementales, chacune essayant de poser un certain nombre de valeurs. Une étude de l'ETH de Zurich récemment publiée sur *Nature Machine Intelligence*² a répertorié 84 chartes rédigées depuis l'apparition de la célèbre lettre ouverte de 2015. Pour se rendre compte de la variété d'acteurs institutionnels qui se sont penchés sur la question, il suffit de remarquer qu'une charte a été récemment signée entre Microsoft et le Vatican. Si les points mis en avant par ces différents textes ne montrent pas forcément de convergence, des régularités se dégagent autour de certains valeurs qui semblent s'imposer. Les cinq les plus récurrents sont 1) la transparence 2) la justice et l'équité 3) la non-malfaisance 4) la responsabilité (en tant que notion juridique) 5) la protection de la confidentialité. Pourtant, dès que l'on regarde dans le détail, une énorme variété d'interprétation de ces valeurs se dégage, parfois entretenue par les auteurs eux-mêmes. Par exemple, en ce qui concerne la transparence, certaines chartes font référence au code en recommandant de produire des IA en open source, d'autres à la transparence des décisions prises par les systèmes automatiques. Encore d'autres chartes pointent la transparence de l'interaction homme-machine ou une transparence de nature économique concernant les investisseurs qui financent l'IA. De même, les valeurs de justice et éthique se composent d'une multitude de facettes, dont 1)

la nécessité de limiter les biais de manière que les algorithmes ne discriminent pas systématiquement certains individus ; 2) la possibilité de faire appel contre une décision réalisée par un algorithme 3) la question de l'impact possible sur le marché du travail. Par rapport à ce dernier point, malheureusement la plupart des chartes éthiques existantes se limitent à insister sur le danger que les êtres humains soient un jour remplacés par des machines - une peur très ancienne qui date de la première pensée industrielle. Le principe de non-malfaisance coïncide avec celui de non-nuisance et rappelle la nécessité d'imaginer, projeter et designer des IA incapables de nuire les êtres humains. Quant à la responsabilité, une notion mise en avant par nombre de chartes éthiques de l'IA, elle est souvent en jeu quand il s'agit de l'imputabilité juridique et de la responsabilité légale des véhicules autonomes en cas d'accident causé par ceux-ci. Le fait de poser une telle question revient à impliquer que l'IA n'est pas une entreprise historique de remplacement de la décision humaine par la décision automatique mais au contraire une manière de distribuer la responsabilité de la décision même entre des acteurs définis comme humains et d'autres acteurs qu'on choisit de définir comme non-humains. Dans le domaine de la protection de la confidentialité (où l'on retrouve souvent des formules très exploitées, dont la "privacy by design") un élément nouveau consiste dans l'idée de vie privée différentielle ou "differential privacy". Celle-ci consiste dans la prise en compte des risques causés par la possible désanonymisation des données personnelles, dans le but d'imaginer des systèmes autonomes qui ne permettent pas une telle désanonymisation. Avec la notion de protection de la confidentialité se pose d'ailleurs la question, très épineuse, de la minimisation de la collecte des

² Jobin, A., Ienca, M. & Vayena, E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1, 389–399 (2019). <https://doi.org/10.1038/s42256-019-0088-2>

données autour de laquelle se concentrent des intérêts discordants.

La géographie des chartes éthiques de l'IA : une reproduction d'anciennes inégalités.

Une question récemment soulevée concerne le problème de la perméabilité des chartes éthiques de l'IA aux intérêts industriels : dans un article publié sur *Intercept*, “The invention of ethical IA. How Big Tech Manipulates Academia to Avoid Regulation”³, l’auteur Rodrigo Ochigame avance l’hypothèse selon laquelle les grandes entreprises du numérique auraient longtemps manipulé les universitaires pour qu’ils produisent des chartes éthiques. Ces chartes seraient une manière de remplir le vide actuel de régulation et d’éviter des mesures plus contraignantes. En effet, la géographie des ces chartes éthiques, révèle des points de concentration, tels l’Europe et les Etats Unis, en correspondance avec les pays qui produisent de l’IA (bien qu’on y remarque un grand absent : la Chine). Je vais revenir sur la carte de la distribution des chartes éthiques de l’IA pour la mettre en relation avec une autre carte qui reproduit le même type de clivage entre le Nord et le Sud du monde : celle qui représente la distribution géographique des travailleurs des plateformes qui sont nécessaires pour produire l’IA.

Une IA éthique de bout en bout : de sa production à son déploiement.

Derrière le constat évident de la distribution inégale des chartes éthiques de l’IA se cache la question de ce qu’on entend lorsqu’on parle de production d’intelligence artificielle. En nous posant cette question, nous découvrons en effet un impensé de l’IA qui concerne aussi la réflexion éthique sur l’IA. La plupart des recherches actuelles sur les impacts sociaux, politiques et économiques des intelligences artificielles se concentrent seulement sur le volet du déploiement. En d’autres termes, on considère l’IA 1) en supposant qu’elle existe 2) en supposant qu’elle marche 3) en supposant qu’elle est déjà déployée et opérationnelle sur des marchés peuplés par des êtres humains. Par conséquent, on reproduit parfois des mauvais réflexes hérités de la sociologie des usages, qui nous amènent à analyser l’acceptabilité d’une innovation, sans nous poser la question de base de son origine.

Mon approche et celle d’autres collègues qui travaillent la question, est au contraire de regarder l’autre bout de l’IA, c’est-à-dire sa production, en examinant ses modalités concrètes. Ainsi faisant, il est d’abord possible de préciser ce que c’est que l’IA, au-delà de ce terme aujourd’hui souvent utilisé de manière interchangeable même dans des contextes scientifiques. Malgré l’idée, désormais partagée, que les intelligences artificielles sont des techniques à forte intensité de données, on s’interroge rarement sur la manière dont de telles données sont produites. Cette question est pourtant essentielle : si l’on veut proposer une IA éthique de bout en bout, il faut prendre en compte les conditions de production des don-

³ Rodrigo Ochigame, The invention of ethical IA. How Big Tech Manipulates Academia to Avoid Regulation, The Intercept 20 décembre 2019. <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>

nées qui servent à la création et au déploiement des intelligences artificielles. Il faudrait, en fait, appliquer le même type de raisonnement éthique que l'on applique parfois à d'autres produits de consommation. Par exemple, le fait qu'un producteur de chaussures ne discrimine pas ses consommateurs (par leur genre, origine géographique, etc.) n'est pas suffisant : pour affirmer qu'il commercialise un bien éthique il faut aussi prendre en compte le respect des droits des travailleurs, leurs conditions de travail, l'impact écologique de la production, etc⁴.

Les exemples qui suivent montrent l'importance de révéler cet impensé de l'IA, voire ses moyens de production. Le véhicule autonome, depuis presque autant de temps que l'IA, suscite un certain nombre de fantasmes dans l'imaginaire collectif, telle l'idée qu'il sera une sorte de "boudoir sur rue" dans lequel un être humain pourra lire le journal alors que la voiture roule tranquillement vers son lieu de travail. La réalité de cette technologie est pour l'instant encore très différente, s'agissant principalement d'un capteur qui communique avec son environnement. Des enjeux éthiques importants sont liés à la quantité de travail nécessaire pour faire circuler les véhicules prétendument "sans chauffeur". Une vidéo réalisée pendant l'un de mes terrains montre le test d'un véhicule produit par une petite startup qui se développe à l'université de Stanford. Le test consiste à relever les temps d'action du véhicule sans chauffeur suite au surgissement d'un enfant (un mannequin) traversant la rue. Si l'on regarde la scène de près, on se rend compte que le mannequin est en fait le seul être non humain impliqué dans ce test. Deux ingénieurs vérifient l'expérience de l'extérieur et deux personnes se trouvent à l'intérieur de la voiture : il s'agit des "opérateurs de véhicule" selon l'appella-

tion que les entreprises productrices préfèrent à celle de chauffeurs. En effet, les producteurs déploient des efforts importants pour cacher ce travail humain. Dans une image montrant un prototype de véhicule sans chauffeur Ford, on voit un agent humain littéralement invisibilisé, caché sous le siège de la voiture. Les sociologues des techniques savent qu'il y a un intérêt particulier à étudier les pannes afin de comprendre une technologie. Le premier accident de voiture autonome qui a concerné un véhicule Uber en 2018 était un événement tragique qui a pourtant permis de mieux comprendre les enjeux sociétaux, humains et techniques de ces technologies. Dans une image captée une seconde avant l'accident par le véhicule même, il est possible de voir la victime qui traverse la rue la nuit, en dehors des clous et avec un vélo : tous les éléments étaient là pour que la machine ne reconnaisse pas la forme d'un être humain qui se dressait devant elle. Une autre image capturée depuis l'intérieur de la voiture juste avant l'impact fatal révèle que dans ce véhicule "sans chauffeur" il y a en effet un chauffeur. Du point de vue de l'imputabilité juridique cette personne était la seule à être considérée comme responsable de l'accident. Alors qu'elle était accusée d'être en train de dormir ou de regarder une émission de télévision - elle suivait en réalité le parcours de la voiture sur sa tablette.

L'opérateur de véhicule est la partie la plus visible du travail humain nécessaire pour faire fonctionner une voiture autonome et on pourrait penser que la technologie les rendra bientôt obsolètes. Pourtant le métier d'opérateur de véhicule n'est pas le seul type de travail humain nécessaire pour développer des voitures autonomes : les données captées et enregistrées par celles-ci doivent ensuite être comparées avec des exemples utilisés pour son

⁴ J'emprunte cette intuition à Irénée Regnauld, qui l'a proposée dans le compte rendu de mon ouvrage *En attendant les robots* (2019). Irénée Regnauld (2019). Antonio Casilli : l'intelligence artificielle est largement

« faite à la main », Mais où va le Web ? P(a)nser le numérique. <http://maisouvaleweb.fr/attendant-robots-coulisses-intelligences-artificielles/>

apprentissage. Ce processus est parfois problématique, comme le montre l'exemple d'un véhicule autonome calibré sur une base de données canadienne qui a causé des accidents en Australie car il n'arrivait pas à reconnaître les mouvements des kangourous. Cela était dû au fait que la partie d'entraînement concernant les grands animaux s'était basée sur des caribous - les grands mammifères présents au Canada qui bougent de manière très différente que les marsupiaux. Les données sur lesquelles les véhicules autonomes sont entraînés ont souvent besoin d'être annotées, triées, parfois enrichies, parfois retravaillées. Cette partie de travail humain reste indispensable pour développer des voitures autonomes. Des micro-travailleurs accomplissent ce travail pour une rétribution d'à peine quelques centimes par micro-tâche. Ils sont recrutés à travers des compagnies spécialisées, dont par exemple la plateforme Mighty AI, rachetée par Uber. Les compagnies de production d'IA ont tendance à présenter ces "ouvriers du clic" en les appelant des "cartographes" et en les montrant dans leurs vidéos promotionnelles comme des jeunes en train de travailler dans un environnement qui rappelle celui d'une start up, selon une mise en scène cohérente avec l'imaginaire de la Silicon Valley. Les acteurs principaux des compagnies qui développent des voitures autonomes tiennent parfois un discours bien plus brutal et cynique sur les micro-travailleurs. Par exemple, Anthony Levandowski⁵ les décrit comme des "robots humains", en faisant référence à leur travail disqualifié et « tâcheronisé » d'annotation et de triage de données - un travail de robot qui sert à produire des robots. D'ailleurs des images montrées par Levandowski même pendant ses présentations ne cachent pas le vrai visage de ces micro-travailleurs et révèlent que leur identité ethnique est bien différente que le type caucasien blond des vidéos promotionnelles

tournees par les entreprises de l'IA.

Pour comprendre l'IA d'aujourd'hui, et pour ensuite penser à une IA éthique, il faut donc d'abord ouvrir les yeux sur sa réalité commerciale et sur les acteurs et mécanismes en jeu dans sa production. Malgré le débat, les fantasmes, les efforts des informaticiens et des industriels, l'IA à laquelle nous avons à faire pour le moment ne vise pas à reproduire une intelligence de niveau humain : elle est définie d'un point de vue épistémologique comme une IA "étroite". Premièrement, les intelligences artificielles actuellement en commerce sont plus banales que les grands rêves ou cauchemars qu'elles ont toujours suscités : nous les avons, pour ainsi dire, "dans nos poches" au sens littéral du terme, sous forme des assistants virtuels tels Siri, Cortana ou Google Now. Deuxièmement, il faut questionner les conditions du travail nécessaire pour produire cette forme d'IA.

Géographie du micro-travail

La plateforme spécialisée d'Amazon, Amazon Mechanical Turk prend son nom du tueur joueur d'échecs du XVIIIe siècle - prétendument la première IA. Cet engin, derrière l'apparence d'un joueur ottoman, cachait à son intérieur un être humain - un joueur d'échecs autrichien qui pilotait le dispositif. Il s'agissait, en fait, du premier exemple de human in the loop : l'humain dans la boucle de l'IA. Aujourd'hui cette même logique est reproduite à l'échelle de millions d'êtres humains qui créent de l'« intelligence artificielle artificielle ». Les entreprises qui souhaitent automatiser certains processus de production peuvent avoir recours à Amazon Mechanical Turk pour recruter des personnes afin d'entraîner, vérifier et simuler des intelligences artificielles en réalisant des

⁵ Anthony Levandowski est ingénieur co-fondateur de la compagnie de véhicules autonomes Waymo (une partie de

Google) et fondateur de Otto, rachetée par Uber.

micro tâches en temps réel. Parmi celles-ci, la retranscription d'un ticket de caisse, l'aide au traitement automatique des langues par l'écoute des bouts de conversations, la labellisation ou le détournement des images. Ce travail est très important pour une multitude d'applications, par exemple pour vérifier la pertinence des résultats d'un moteur de recherche, au point que Bing et Google ont développé leurs propres plateformes consacrées au recrutement des micro-travailleurs pour vérifier et améliorer leurs résultats. Les personnes qui acceptent d'accomplir cette forme de travail reçoivent des rémunérations très faibles - de l'ordre de un ou deux centimes et ne sont pas encadrées juridiquement. Souvent, les micro-travailleurs sont liés aux plateformes uniquement via des contrats d'adhérents ou de participation qui ressemblent aux conditions générales, souvent signées sans réfléchir par les usagers pour télécharger des applications. Devenu plus visible seulement récemment, malgré son existence datant de plus de quinze ans, le phénomène des plateformes de micro travail monte en puissance : la plupart des GAFAM s'en sont dotées et certains pays commencent à devenir des véritables géants dans ce secteur. Par exemple, la plateforme chinoise Zhubajie à elle seule compte 15 millions de micro-travailleurs, alors que Amazon Mechanical Turk en affiche 500.000. Selon nos estimations, il existe environ 260.000 micro-travailleurs en France⁶. Ceux-ci ont accès à une pluralité de plateformes françaises, et aussi américaines et chinoises.

Si l'on observe la géographie de la distribution des micro-travailleurs qui traitent les données nécessaires à nourrir l'IA, on remarquera

qu'ils sont installés en majorité dans des pays du Sud du monde. Une cartographie réalisée par Mark Graham de l'université d'Oxford en 2013 représente les pays avec une majorité de micro-travailleurs et ceux dans lesquels les micro-travailleurs sont recrutés et où les données annotées sont achetées pour entraîner des IA⁷. En regardant cette carte on aperçoit une reproduction de très anciennes inégalités et asymétries au niveau planétaire en termes de richesse, de pouvoir et d'influence géopolitique. Cette image continue à être intéressante tout en étant datée car, étant basée sur des données de plateformes anglophones, elle met en évidence la reproduction de certaines logiques colonialistes. Les pays qui fournissent davantage de données sont des ex-colonies, comme par exemple les Philippines et l'Inde. Pourtant, dans cette carte de 2013 on remarque des absents importants, comme l'Afrique. La situation a changé depuis, comme le montre un projet lancé en 2019 sur l'entraînement des IA francophones qui prévoit des terrains de trois ans au Madagascar, au Cameroun, au Mali, au Sénégal, c'est-à-dire les pays de production de l'IA française⁸. Il est important de mettre cette dernière carte en résonance avec la géographie des chartes éthiques de l'IA mentionnée plus haut. Un autre impensé de l'éthique de l'IA émerge de cette comparaison : l'importance de représenter les travailleurs en relation à leur enracinement national, culturel, géographique et historique.

⁶ Paola Tubaro, Clément Le Luëc, Antonio A. Casilli (2020). Counting 'micro-workers': societal and methodological challenges around new forms of labour. *Work Organisation, Labour & Globalisation*, 14 (1): 67-82.

⁷ Mark Graham, Isis Hjorth et Vili Lehdonvirta (2017). Digital labour and development : impacts of global

digital labour platforms and the gig economy on worker livelihoods. *Transfer : European Review of Labour and Research*, 23(2): 135-162.

⁸ Il s'agit du projet ANR HUSH « The Human Supply Chain Behind Smart Technologies », Télécom Paris/LRI CNRS.

Conclusion

Malgré l'accent mis sur les valeurs de la transparence, la confidentialité, la justice, la responsabilité, la non-malfaisance, l'éthique de l'IA est encore caractérisée par des instances divergentes et conflictuelles qui sont par ailleurs des miroirs d'intérêts économiques et politiques sous-jacentes. Rédigées en majorité par des institutions et organismes principalement situés dans les pays producteurs de l'IA, ces chartes pourraient constituer en réalité une manière pour les industriels d'éviter une régulation plus contraignante.

Un autre point que j'ai essayé de mettre en évidence est que l'éthique de l'IA d'aujourd'hui ne se concentre que sur les usages de l'automatisation, en faisant l'impasse sur le travail fragmenté et sous-payé de millions d'ouvriers du clic nécessaire pour produire cette automatisation. Que se passerait-il si, au contraire, les droits de ces ouvriers du clic étaient protégés, s'ils avaient la possibilité de résister au chantage du microtravail et à la micro rémunération, s'ils avaient une voix pour protester contre ou pour refuser de contribuer à des IA qu'ils considèreraient problématiques d'un point de vue éthique ? Dans un pareil scénario, l'éthique changerait radicalement car on cesserait de s'intéresser aux choses qui inquiètent les producteurs au sens de propriétaires et on commencerait à regarder ce qui concerne les producteurs au sens de travailleurs.

Poursuivre une IA "réellement éthique" signifie aussi et avant tout reconnaître ce travail invisible du clic qui est l'ingrédient secret de l'automatisation contemporaine et offrir aux travailleurs des méthodes pour se protéger.

Robotique : l'intelligence de la gravité

JEAN-PAUL LAUMOND

Jean-Paul Laumond est roboticien et directeur de recherche émérite au CNRS. Membre du LAAS-CNRS à Toulouse jusqu'en 2019, il a rejoint l'équipe Willow du département informatique de l'ENS à Paris (unité mixte 8548 ENS-CNRS-INRIA-PSL). Sa recherche est dominée par l'algorithmique de la planification de mouvement en robotique, un domaine scientifique dont il a contribué à jeter les bases. De 2001 à 2003, il crée et dirige la société Kineo CAM qui commercialise ces technologies dans le domaine du prototypage virtuel. En 2006 il crée le groupe de recherche Gepetto dédié à l'étude des fondements calculatoires de l'action anthropomorphe. De 2014 à 2018, il conduit le projet Actanthrope soutenu par l'European Research Council (ERC). Il est Fellow de l'IEEE. En 2011-2012 il est le titulaire de la chaire Innovation Technologique Liliane Bettencourt du Collège de France. En 2016 il est lauréat du prix international IEEE Inaba Technical Award for Innovation Leading to Production. Il est membre de l'Académie des technologies et membre de l'Académie des sciences.

Introduction : Qu'est-ce que la robotique ?

A défaut de préciser exactement ce qu'est un robot, on peut tenter de définir ce qu'est la robotique : “ la robotique traite du rapport que peut entretenir avec le monde réel une machine qui bouge et dont les mouvements sont contrôlés par un ordinateur ”¹. La robotique a aujourd'hui envahi tous les secteurs : médical, spatial, transport, services, industries, défense, agriculture, sous-marin, aérien. Nous connaissons divers types de robots : ceux qui construisent des voitures, ceux des entrepôts d'Amazon ou encore les robots humanoïdes comme le robot HRP2 de l'AIST au Japon ou le robot Atlas de Boston Dynamics aux Etats Unis.

L'enjeu pour les chercheurs en robotique est de rendre compte des avancées de la discipline et d'expliquer le travail qu'ils font. Comme disait un grand penseur des techniques, Gilbert Simondon : “ La plus forte cause d'aliénation dans le monde contemporain réside dans cette méconnaissance de la machine, qui n'est pas une aliénation causée par la machine, mais par la non-connaissance de sa nature et de son essence,... ”².

¹ J.P. Laumond, *La robotique : une récidive d'Hephaïstos*, Leçon inaugurale, Collège de France - Fayard, 2012.

² G. Simondon, *Du mode d'existence des objets techniques*, Aubier, 1967.

I. La robotique : la science du mouvement

La robotique est issue d'une convergence dans les années 1940-1950 entre la cybernétique et l'intelligence artificielle. La cybernétique s'intéresse à la régulation des systèmes vivants et artificiels. Le principe d'homéostasie est un exemple de système de régulation du vivant : la température du corps humain est d'environ 37°C, qu'il soit dans un environnement où il fait -30°C comme en Sibérie ou plus de 40°C comme au Sahara. Le corps maintient sa température quelles que soient les conditions extérieures. Si on transporte ces principes de régulation sur des objets techniques, cette régulation peut être parfaitement mécanique. Il en va ainsi du régulateur à boules inventé par l'ingénieur écossais James Watt à la fin du XVIIIème siècle. Cet astucieux système constitué de deux boules tournant autour d'un axe a permis la régulation de la machine à vapeur : quelle que soit la pression de la vapeur, l'axe tourne à une vitesse constante, vitesse que l'on peut faire varier à loisir. Cette invention a fortement contribué au développement industriel au XIXème siècle.

Dans les années 1940, le mathématicien Norbert Wiener a théorisé ce principe de régulation avec la cybernétique : il s'agit de la boucle de rétroaction sensori-motrice. Un système cybernétique se compose d'un capteur (par exemple le thermostat dans une maison) et d'un actionneur (le brûleur de la chaudière). En associant les deux il est possible de maintenir une température constante dans la maison. Jusqu'au début du XXème siècle, les systèmes de régulation étaient mécaniques. Avec les progrès de l'électronique, il est possible d'associer des signaux, le signal de sortie du capteur et le signal d'entrée de l'actionneur. Aux deux boules du régulateur de Watt, on substitue un couplage électronique (analogique) entre les deux signaux (c'est à cette époque qu'apparaissent les premiers automates électroniques tels que les tortues de Grey Walter et le renard

d'Albert Ducros). La numérisation du signal, rendu possible par le développement des premiers calculateurs, va permettre dès les années 1950 de remplacer le couplage analogique par un couplage numérique prenant la forme d'algorithmes. Le bouclage électronique se généralise ainsi en un processus de traitement de l'information, associant perception, décision et action. La robotique naît de ce mariage entre cybernétique et informatique à la fin des années 1950 et le premier robot industriel fera son apparition sur les chaînes d'assemblage de la General Motors au début des années 1960 sous l'impulsion de Joseph Engelberger et Georges Devol.

a. Le robot : une machine à information

Avant tout, le robot est une machine qui bouge pour effectuer une tâche bien définie. La question centrale est de savoir comment un ordinateur peut calculer un mouvement adapté à la tâche.

Il s'agit à ce stade de préciser la problématique robotique en regard de ce qu'il va être convenu d'appeler l'intelligence artificielle à l'issue du colloque de Dartmouth en 1956. A partir des années 1960 les progrès en traitement de l'information vont être spectaculaires. Des logiciels existent aujourd'hui qui résolvent des problèmes d'une complexité phénoménale. Capables d'abord de battre les meilleurs joueurs d'échecs à la fin des années 1990, ils sont depuis les années 2010 capables de battre les meilleurs joueurs de Go, jeu pourtant réputé beaucoup plus complexe que le jeu d'échecs. Dans ces deux exemples, l'ordinateur traite de l'information et seulement de l'information : il n'est pas couplé avec une machine physique. Il n'intègre pas de processus de régulation sensori-motrice.

Force est de constater que les progrès en robotique sont beaucoup plus lents : il n'existe pas de robot capable de se déplacer pour servir à boire, du moins en dehors de certains laboratoires de recherche (et encore le font-ils de

manière extrêmement laborieuse dans ces laboratoires). Il existe ainsi des ordinateurs capables de jouer à des jeux très complexes et il n'existe pas de machine capable d'effectuer une tâche aussi simple que prendre un verre. Tentons d'élucider ce paradoxe. Jusque dans les années 1970 la problématique de la robotique industrielle consistait à programmer le robot au niveau de la trajectoire qu'il devait effectuer pour, par exemple, souder les composants d'une carrosserie d'automobile. De façon générale, un robot dont on devrait programmer tous les détails du mouvement à effectuer pour prendre un verre ne servirait à rien. Il faudrait le reprogrammer à chaque fois que les conditions de réalisation de la tâche changeraient (ce qui n'est pas le cas en robotique industrielle ; les chaînes de production sont adaptées aux capacités du robot). Depuis les années 1970, la robotique se donne pour ambition de programmer des machines au niveau de la tâche qu'elles ont à réaliser, et non plus au niveau du mouvement à effectuer.

b. Du mouvement à l'action

Développons ce dernier point. Qu'entend-on par programmer une machine au niveau de la tâche ? Que signifie "prends un verre" ? Dans cette expression, il y a un verbe d'action (prendre) et un objet du monde physique (le verre). L'expression qui définit la tâche à effectuer ("prendre un verre") emprunte à l'*espace physique*. Elle est indépendante de l'agent à qui elle s'adresse. Or, qu'il soit homme ou machine, l'agent accède au monde physique à travers son *espace sensoriel*, et la seule manière pour lui d'agir s'effectue à travers son *espace moteur*.

Pour l'homme, l'espace sensoriel est essentiellement constitué de ses cinq sens (de fait, il y en a plus de cinq) et l'espace moteur constitué de quelque 700 muscles. Enfin le cerveau, ou de manière générale le système nerveux, réalise l'interface entre les trois espaces : physique, sensoriel et moteur. Dans le cas d'un robot c'est exactement pareil sauf que les capteurs (l'espace sensoriel) sont des caméras, des capteurs d'efforts, des gyromètres, des accéléromètres ; pour l'espace moteur cela varie selon le robot. L'équivalent du cerveau est l'ordinateur, c'est-à-dire une machine à calculer.

Le problème qui se pose est de savoir traduire l'expression de l'action dans le monde physique à l'intérieur du corps du robot. C'est ce que j'appelle « l'incorporation de l'action »³ (traduction d'embodiment) : comment transformer une action exprimée dans l'espace physique en mouvements qui trouvent leur origine dans l'espace sensori-moteur ?

Cette question dépasse de loin le cadre de la robotique. Elle s'applique à tout être vivant, de l'écureuil qui saisit une noisette à l'artisan qui façonne un objet. Trouver le bon geste, trouver le mouvement qui va atteindre le but recherché. Pour comprendre et explorer la question, elle doit être formulée mathématiquement. Il s'agit d'un problème d'inversion d'espaces, un problème de géométrie. Henri Poincaré a très bien développé ce point de vue dans un article remarquable⁴. Il s'agit de trouver la « bonne » géométrie qui lie l'espace physique (Poincaré parle quant à lui de « l'espace géométrique ») à l'espace sensori-moteur : « *l'expérience nous guide dans ce choix qu'elle ne nous impose pas ; elle nous fait reconnaître non quelle est la géométrie la plus vraie, mais quelle est la plus commode.* »

³ J.P. Laumond, Robotique et simplicité : vers un génie de l'action incorporée, in Complexité-Simplicité, A. Berthoz, J.L. Petit (Eds), Collège de France, 2014.

⁴ H. Poincaré, L'espace et la géométrie, Revue de métaphysique et de morale, Tome 3 pages 631-645, 1895.

⁵ G. Venture, J.P. Laumond, B. Watier, Biomechanics of Anthropomorphic Systems, Springer Tracts in Advanced Robotics, 124, Springer, 2019.

Cette idée rend compte de nombreux travaux de recherche en robotique, qui reposent sur l'élaboration de géométries particulières, que ce soit par exemple pour un robot qui ramasse une balle au sol, une voiture autonome qui tracte une caravane, ou un robot humanoïde qui passe à travers la brèche d'un mur de brique. Tous ces travaux sont du ressort de la géométrie. Prendre le point de vue mathématique permet d'aborder le problème dans une perspective pluridisciplinaire. C'est ainsi que les travaux portant sur la locomotion sur les robots humanoïdes trouvent des échos en neurosciences et en biomécanique⁵. A la fin du XIX^{ème} siècle des physiologistes comme Etienne-Jules Marey ou Eadweard Muybridge essayaient de comprendre ce qui se passait dans le corps humain lorsqu'il marchait. La question est encore loin d'être élucidée encore aujourd'hui. Il n'en reste pas moins que des progrès sont réalisés. Récemment, guidés par l'intuition que les géométries de la locomotion bipède sont de même nature que les géométries permettant à un robot mobile de faire automatiquement un parking, nous avons pu montrer l'existence d'une petite roue (virtuelle) située au niveau du bassin et entraînant notre centre de masse lorsque nous marchons⁶. La marche requiert la mise en mouvement de nombreuses articulations. Elle se traduit par le contrôle de quelque trente degrés de liberté : ceux des bras qui balancent, ceux du cou qui maintiennent la tête à peu près fixe et ceux des jambes dans leur mouvement cyclique. Toute la complexité de la locomotion humaine se trouve entièrement concentrée dans le mouvement du seul centre de masse entraîné par la petite roue virtuelle. En biomécanique, on parle de synergie motrice : les mouvements des différentes articulations du corps sont coordonnés entre eux selon un schéma fixé et le système nerveux ne contrôle

de fait qu'un petit nombre de degrés de liberté, ceux de la petite roue. Les synergies motrices sont une illustration du principe plus général de *simplicité* introduit par A. Berthoz⁷, et qui regroupe les mécanismes que l'évolution a mis en place pour permettre au vivant de s'adapter à la complexité de sa relation avec l'environnement.

II. Robotique et intelligence artificielle

Pour mettre en évidence l'actualité du robot dans l'histoire des machines, on le qualifie parfois de « machine intelligente ». La notion d'intelligence utilisée dans cette expression se réfère à la capacité d'un système de raisonner pour s'adapter aux fluctuations de l'environnement (on retrouve les principes de la cybernétique).

a. Les moteurs de raisonnement

Rappelons ici les trois grands modes de raisonnement :

- La déduction. Exemple : Si Socrate est un homme et que tous les hommes sont mortels, alors Socrate est mortel. [*Cela est certain*]

- L'induction. Exemple : Si Socrate et d'autres avec lui sont fumeurs et que Socrate et ces autres ont le cancer, alors fumer donne le cancer. [*Cela est probable*]

- L'abduction. Exemple : Si Socrate a le cancer et que fumer donne le cancer, alors Socrate est fumeur. [*Cela est vraisemblable*]

La distinction entre ces trois modes de raisonnement est au cœur des débats épistémologiques sur le statut de la connaissance. Ce sont trois composantes du raisonnement rationnel.

⁵ G. Venture, J.P. Laumond, B. Watier, Biomechanics of Anthropomorphic Systems, Springer Tracts in Advanced Robotics, 124, Springer, 2019.

⁶ J. Carpentier, M. Benallegue, J.P. Laumond, *On the centre of mass motion in human walking*, International journal on Automated Computing, Vol. 14, N. 5, 2017.

⁷ A. Berthoz, *La Simplicité*, Odile Jacob, 2009.

Mais seulement un raisonnement, la déduction, répond à un principe de vérité, et les deux autres, l'induction et l'abduction, répondent à un principe d'action.

La période que nous traversons avec la pandémie de Covid-19 permet d'illustrer cette distinction. Malgré l'absence de connaissance établie sur le virus et l'impossibilité de recourir à des modes de raisonnement déductif et donc certains, les pouvoirs publics sont tenus de prendre des décisions. Ils sont soumis au principe d'action. Ils agissent de leur mieux sur la base de l'expérience et de traitement statistique de données (mode de raisonnement inductif). En l'absence de ces données quantitatives, un médecin face à un malade est seul (il ne représente pas les pouvoirs publics) et il se doit d'agir ; il doit prendre une décision et décider d'une thérapie même s'il n'est sûr de rien ; il fait alors appel à son intuition (mode de raisonnement abductif). La démarche scientifique articule ces trois modes de raisonnement. Si le scientifique fait appel à son expérience (voir plus haut le conseil donné par Poincaré) et à son intuition (voir Bergson) pour élaborer des modèles de connaissance, ces modèles doivent être in fine prouvés et falsifiables (voir Popper) dans un mode de raisonnement déductif.

Toute nouvelle machine arrivant sur le marché est soumise à des processus de certifications : comment garantir qu'elle fonctionne bien selon un cahier des charges rigoureusement défini ? En matière de traitement numérique de l'information, la question est critique.

b. L'Intelligence artificielle actuelle : apprentissage et science des données

Je dis souvent qu'on s'habitue à utiliser des logiciels qui ne marchent pas. Prenons l'exemple du logiciel Siri qui permet de communiquer avec nos téléphones portables. Soumettons-le à trois cas.

Cas 1 :

- Voix : Dis Siri «Dis à Yvonne que je prends le pain en passant».

- Siri : “Voici votre message, dois-je l'envoyer : « Je prends le pain en passant » ?”

Cas 2 :

- Voix : “Dis Siri, essaye de convaincre Yvonne de venir”

- Siri : “Je ne suis pas sûr de pouvoir faire ça”

Cas 3 :

- Voix : “Dis Siri, Yvonne va-t-elle bien ?”

- Siri : “Je suis désolé”

Dans le premier cas, le logiciel remplit parfaitement sa fonction et on peut s'en émerveiller ; le logiciel a transformé les ondes sonores que nous émettons en parlant en mots, analyse ces mots et propose d'envoyer un message à Yvonne. Ce n'est pas rien. Dans le deuxième cas, le logiciel reconnaît que la phrase sort de son domaine de compétences, ce qui n'est pas rien. En revanche, dans le troisième cas, la réponse de Siri est absurde. Ce logiciel nous donne l'illusion que c'est une machine qui comprend le langage naturel alors qu'elle ne maîtrise le langage naturel que dans un champ extrêmement restreint qu'on ne nous définit pas.

Autre exemple avec la reconnaissance d'image. Deux images quasi identiques représentent un bus jaune. Le logiciel désigne la première image comme étant un bus alors que la seconde image est étiquetée “Autruche”. Que se passe-t-il ? Le réseau de neurones qui est à l'oeuvre ici utilise des structures de données intermédiaires. Il suffit d'introduire un peu de bruit dans cette structure et l'algorithme étiquette différemment les deux images pourtant similaires.

Ce sont donc des logiciels fragiles, mais ils fonctionnent la plupart du temps et satisfont notre usage. Le succès de l'intelligence artificielle repose, depuis les années 2000, sur le traitement statistique de grandes masses de données. Ces algorithmes établissent des corrélations

lations entre variables. Leur efficacité tient aux progrès réalisés en optimisation numérique, un domaine très actif des mathématiques appliquées, stimulé par les progrès technologiques en matière de stockage et de traitement de l'information. Ces méthodes sont basées sur les données, en opposition aux méthodes basées sur des modèles. Elles montrent que le raisonnement inductif a une réelle efficacité (elles fonctionnent vraiment la plupart du temps). Il reste qu'elles ne produisent pas de connaissances, à moins de confondre corrélation et causalité, confusion malheureusement souvent répandue. Comme le dit très bien René Thom : « *Prédire n'est pas expliquer* »⁸.

c. Robotique et certification

À l'issue de la catastrophe de Fukushima en 2011, on a été incapable d'envoyer des machines autonomes sur le terrain. Ce sont des hommes qui sont intervenus dans la centrale nucléaire pour fermer les vannes. La technologie n'est pas au point pour remplacer l'homme par la machine dans de telles circonstances. La DARPA, une agence de financement de la recherche aux États-Unis, a lancé un grand défi en 2015. Il s'agissait de répondre à un problème défini par un cahier des charges précis. Sur la base d'un scénario d'intervention dans un contexte de catastrophe, les missions que la machine devait remplir comprenaient, entre autres : savoir grimper une échelle, conduire une voiture, marcher sur des gravats, forcer une porte coincée, fermer des vannes etc. Une trentaine de laboratoires parmi les meilleurs au monde se sont lancés dans l'aventure pendant trois ans et aux vues des résultats, pour le moins décevants, on constate que la gravité à laquelle est soumise la machine est loin d'être maîtrisée. Aucun des robots n'a réussi à rem-

plir le cahier des charges. La finale du concours a présenté un florilège de chutes spectaculaires. La maîtrise de la gravité reste un défi pour la machine. Traduire ses lois sous formes d'algorithmes fait l'objet des recherches contemporaines en robotique. Les progrès sont lents. Il faut comprendre que le logiciel qui permet à un robot humanoïde de marcher ne doit pas fonctionner « la plupart du temps » ; il doit pour le moins répondre à des normes avant d'être mis sur le marché. Toute machine subit des batteries de tests préalables à sa commercialisation.

Le problème n'est pas nouveau. Dans l'aéronautique, par exemple, la certification est un problème critique, et pas seulement pour les logiciels de pilotage automatique. Et on constate en général qu'il n'y a pas trop d'accidents. Il reste que le problème est complexe. Les accidents récents du Boeing 737 Max sont en ce sens emblématiques. Pour certifier les logiciels il faut faire voler les avions de nombreuses fois avant leur commercialisation, ce qui implique des coûts considérables. Or, partant d'un état d'une technologie qui a fait ses preuves, c'est-à-dire partant d'un avion certifié (le Boeing 737), l'ajout ou la modification d'un composant conduit à ce qu'on pourrait appeler "la certification additive". L'avion est certifié, le nouveau composant est certifié, mais qu'en est-il de la certification de l'avion modifié ? Dans le cas précis, il apparaît que la certification de l'ensemble du système n'a pas été vérifiée expérimentalement par des heures de vol d'essai, et ce, pour des questions évidentes de coûts⁹.

Qu'ils soient basés sur des modèles ou basés sur des données, les algorithmes et programmes issus de ces méthodes doivent être éprouvés et validés en conditions réelles d'utilisation. Dans le cas de la robotique, insistons sur le fait que ce sont in fine nos connaissances des lois de la physique des corps en mouvement, et notre

⁸ R. Thom, *Prédire n'est pas expliquer*, Champs Sciences, Flammarion, 1993.

⁹ <https://www.seattletimes.com/business/boeing-aerospace/failed-certification-faa-missed-safety-issues-in-the-737-max-system-implicated-in-the-lion-air-crash/>

capacité à les traduire en programmes, qui sont mises à l'épreuve.

III. Les mots pour le dire

Les robots sont des machines encore nouvelles et leur impact est difficile à évaluer. Leur introduction dans les années 1970 dans le monde industriel a nécessité de revoir entièrement les systèmes de production. C'est particulièrement vrai dans l'industrie automobile. Les progrès en matière de recherche permettent d'envisager des applications de plus en plus diversifiées. Il est nécessaire, à juste titre, d'anticiper les transformations des modes de vie qu'ils annoncent. C'est dans cette perspective que se sont créés moult comités d'éthique (un article récent de la revue *Nature Machine Intelligence* analyse un corpus de pas moins de 84 documents émanant de différentes institutions sur l'éthique en intelligence artificielle¹⁰). En amont de ces réflexions à caractère prospectif, il est indispensable de s'interroger sur la fonction du langage de désignation de ces nouvelles machines¹¹. Elles sont « intelligentes », « autonomes », elles « prennent des décisions », elles « apprennent », etc... Ces expressions sont empruntées au langage du vivant. Elles font naturellement référence aux capacités humaines. De là, naît une fascination incontrôlée, au-delà d'un raisonnable légitime. C'est ainsi qu'on voit apparaître au plan juridique la possibilité d'accorder une personnalité morale à la machine, conséquence du fait qu'une machine « autonome » se devrait d'être juridiquement responsable ; la nouvelle machine devrait acquérir un nouveau statut. L'évidence du raisonnement n'est qu'apparence. Il s'agit de comprendre que la notion d'autonomie doit être dissociée suivant qu'on parle d'une machine ou d'un être

humain.

Prenons l'exemple des drones militaires. Depuis trois ans, la France a décidé d'équiper ses drones de missiles. Ils n'étaient jusqu'à présent utilisés que pour des opérations de surveillance. Les institutions qualifient ces armes de « systèmes d'armes létales autonomes » et des associations préconisant un moratoire sur leur développement les qualifient quant à elles de « robots tueurs » (elles sont regroupées sous le nom de « Stop Killer Robots »). Le terme « autonomie » est au coeur des débats. Il est assez naturel de penser qu'une arme autonome puisse être qualifiée par extension de robot tueur. Qu'en est-il au juste ?

Un opérateur situé à plusieurs milliers de kilomètres de distance du théâtre des opérations décide de lancer un drone pour atteindre une cible qui se trouve par exemple dans une montagne du Yémen. Le drone est équipé d'un système de localisation et d'un système de vision. Le système de localisation (constitué de GPS et de centrales inertiels) permet au drone de se situer dans l'espace et de naviguer jusqu'à atteindre le lieu de la cible. La navigation se déroule sans intervention humaine, comme dans le cas d'un avion en pilotage automatique. Le système de vision (constitué de caméras et de capteurs infrarouges) permet l'identification de la cible à atteindre (un bâtiment, un véhicule ou un combattant). Dès lors que la cible est identifiée un ordre de tir est donné, soit par l'opérateur qui valide la décision à prendre, soit par un algorithme. L'expérience montre qu'il y a des « bavures », un combattant pouvant être confondu avec un berger yéménite qui tient un bâton identifié à tort à une arme.

Qui est responsable de la bavure ? Clairement celui qui prend la décision du tir : soit l'opérateur, soit l'algorithme. Cette question pose le problème de la dissociation de la notion

¹⁰ A. Jobin, M. Ienca, E. Vayena, *The global landscape of AI ethics guidelines*, *Nature Machine Intelligence*, Vol. 1, pp 389-399, 2019.

¹¹ J.P. Laumond, E. Danblon, C. Pieters, *Wording Robotics: Discourses and Representations on Robotics*, Springer Tracts in Advanced Robotics, 130, Springer, 2019.

de « prise de décision » sous-jacente à la notion d'autonomie. Un algorithme peut-il être responsable des erreurs de conception au même titre qu'un opérateur humain ? Un algorithme ne fait qu'exécuter une séquence d'instructions. S'il peut confondre un combattant avec un berger, c'est qu'il ne fonctionne pas. La véritable responsabilité tient au gradé qui a pris la décision d'envoyer une machine qui fonctionne « la plupart du temps », et ce, en toute connaissance de cause. Le drone autonome n'est rien d'autre qu'une machine automatique. Il n'y a pas de différence fondamentale avec une mine anti-personnel incapable de distinguer si le pied qui déclenche l'explosion est celui d'un combattant ou celui d'un enfant, et ce même dans le cas où on laisserait à un logiciel le soin de faire le distinguo. La machine ne « prend de décision » au sens commun de l'expression, et on attend d'elle qu'elle « fonctionne toujours » et non « la plupart du temps ».

Un robot autonome est une machine automatique dans son milieu. Le milieu est défini par des normes et des réglementations qui constituent un univers fermé, et c'est là tout le problème de la certification. Un robot qui prendrait une décision hors de son milieu est une machine qui dysfonctionne. Il en va tout autrement d'un humain qui prend une décision en fonction d'un contexte ; son milieu constitue un univers ouvert, lieu d'expression de sa liberté. Il est intéressant de noter que les qualificatifs associés au robot, machine considérée comme un objet technique, sont toujours positifs (« machine intelligente »), alors que l'usage métaphorique du mot robot appliqué à l'humain (« il agit comme un robot ») disqualifie la personne.

Conclusion

La méconnaissance de la technique et de l'usage métaphorique des mots qui la désignent contribuent à décentrer la question légitime des impacts des nouvelles technologies sur notre société et suscitent des peurs infondées, par exemple quand on nous annonce que les robots vont être doués « d'empathie » et vont tenir lieu de « compagnons » aux personnes âgées.

Retenons qu'il est encore loin le temps où une machine pourra nous servir à boire et que, foi de roboticien¹², une machine ne peut être douée d'empathie.

Il en va de la responsabilité du chercheur d'entrer dans le débat public et de trouver sa juste place entre les postures extrêmes du sensationnalisme cher aux médias (le robot serait le relai de notre humanité) et du réductionnisme qui ne répond pas aux questions (le robot ne serait qu'une machine comme une autre). Il en passe par la maîtrise du langage pour dire ce qu'il en est, un défi pour le roboticien.

¹² A défaut de place pour argumenter...

Traiter les maladies graves à l'aide de nanomédicaments : pour un Homme réparé mais pas augmenté

PATRICK COUVREUR

Patrick Couvreur, membre de l'Académie des sciences, est professeur émérite à l'Université Paris-Saclay. Titulaire de la chaire « Innovations Technologiques Liliane Bettencourt » au Collège de France (2009-2010), il est membre honoraire de l'Institut universitaire de France (IUF). Ses travaux portent sur la conception de nanomédicaments. Créateur de trois start-ups dont l'une est entrée en bourse, il a obtenu de nombreuses distinctions scientifiques en France (Prix Galien 2009, Médaille de l'Innovation du CNRS 2012, Grand Prix Achille Le Bel 2019 etc.) et à l'étranger (Host-Madsen Medal en 2007, European Inventor Award 2013, « Speiser Award 2014 de l'ETH de Zürich, Higuchi Award 2016 au Japon etc.). Il est également membre de l'Académie nationale de médecine, de l'Académie des technologies et Président de l'Académie nationale de pharmacie. A l'étranger, il est membre de la National Academy of Medicine (USA), de la National Academy of Engineering (USA), de l'Académie Royale de Médecine (Belgique) et de la Real Academia Nacional de Farmacia (Espagne).

Introduction : de l'homme réparé à l'homme augmenté?

Les nanotechnologies font partie des NBIC (“nanotechnology, biotechnology, information technology, cognitive science”), nanotechnologies, biotechnologies, technologies de l'information et sciences cognitives : déjà puissantes individuellement, toutes ces technologies le sont encore plus quand elles convergent pour ensemble améliorer les capacités de l'homme. Les récentes avancées de la recherche sur les nanomédicaments posent aux chercheurs de nouvelles questions d'ordre éthique et philosophique. Face aux possibilités inédites d'intervention sur le corps humain apportées par les nanotechnologies, un questionnement important surgit sur la distinction entre homme réparé et homme augmenté. Là où les limites sont encore floues, la nécessité de questionner ces deux termes s'impose : du point de vue du médecin et du pharmacien, réparer l'homme c'est redonner à un individu qui souffre d'une maladie ou d'un handicap ses capacités physiques - en d'autres termes : guérir. Pourtant, des interventions basées sur des approches scientifiques, ou le plus souvent technologiques, qui visent à augmenter les performances humaines dites normales (p.ex : courir plus vite, sauter plus haut, etc.) amènent à parler d'un homme augmenté. Or, la frontière entre homme réparé et homme augmenté est parfois imprécise. Si l'on considère le cas de la chirurgie esthétique, ralentir le vieillissement revient-il à augmenter les capacités normales des individus ou à les réparer ?

Dans le domaine du médicament, la même

substance administrée peut donner lieu à un homme réparé ou à un homme augmenté, selon les cas. Par exemple, l'érythropoïétine : un médicament très efficace pour traiter des anémies sévères, offre aussi la capacité d'améliorer l'oxygénation, avec des effets sur les performances des athlètes. De la même manière, des substances psychoactives utilisées par les militaires partant au combat peuvent être par ailleurs très efficaces dans le traitement de certaines maladies neurologiques.

En plus des distinctions à établir entre homme réparé et augmenté, il faut s'interroger pour savoir si l'amélioration des capacités physiques et intellectuelles de l'être humain par les nanotechnologies est légitime.

Lorsque l'on parle d'homme augmenté, la question du transhumanisme se pose. Certains transhumanistes imaginent déjà de numériser le cerveau humain et de permettre ainsi à l'homme de penser à l'extérieur de son enveloppe charnelle. Or, cette idée ne correspond pour l'instant à aucune réalité scientifique, car le cerveau humain ne se réduit pas à une accumulation de données et il se distingue, entre autres, par sa capacité d'oublier - une défense naturelle extrêmement importante, par exemple, lorsqu'on a vécu des événements traumatiques. De la même manière, les émotions, les sentiments, l'empathie pourraient difficilement être numérisés. La question de l'augmentation de l'homme apparaît donc problématique sous plusieurs aspects : celui de sa faisabilité d'un point de vue scientifique mais aussi, lorsqu'une intervention sur les performances de l'homme est déjà possible (comme dans le cas de certains nanomédicaments), son application suscite des perplexités et des interrogations pour les effets qu'elle pourrait avoir sur la santé des individus, ainsi que sur la société. L'amélioration artificielle des performances physiques ou cognitives de l'homme pourrait amener à l'imposition de nouvelles normes par les gouvernements, les employeurs, les assurances etc.

Limites de la galénique "classique" et innovations des nanotechnologies dans le domaine du médicament.

Lorsque l'on administre un médicament sous une forme chimiothérapeutique conventionnelle, un grand nombre de principes actifs se retrouvent limités dans leur activité thérapeutique, soit par un passage transmembranaire réduit soit parce que, une fois administrés, ils sont rapidement dégradés par l'organisme. De plus, le médicament va souvent imprégner la presque totalité des tissus, organes et cellules de l'organisme. Ce phénomène se traduit par un manque de spécificité, à la fois au niveau cellulaire et tissulaire, entraînant une faible activité pharmacologique, voire des effets secondaires importants, parfois les deux phénomènes à la fois. Aussi, nombre de médicaments dont les cibles sont situées au niveau intracellulaire, se révèlent inefficaces, car incapables de pénétrer à l'intérieur des cellules. Une autre limitation de l'administration "classique" du médicament consiste dans l'induction de phénomènes de résistance, dont des exemples connus nous viennent du domaine de l'antibiothérapie et du traitement du cancer. Face aux limites thérapeutiques de certains médicaments, l'innovation par des nanomédicaments agit alors par une modification de la galénique, c'est-à-dire de la forme d'administration de la substance active : celle-ci est encapsulée dans des nanovecteurs (mesurant entre quelques dizaines à quelques centaines de nanomètres) qui permettent de protéger la molécule de la dégradation.

Des nanovecteurs peuvent d'ailleurs être fonctionnalisés par des "ligands", véritables missiles, capables de reconnaître de manière extrêmement sélective certains récepteurs au niveau des cellules cancéreuses. Le résultat : un meilleur ciblage au niveau cellulaire et tissulaire, donc un index thérapeutique supérieur. Certains ligands peuvent également améliorer la pénétration intracellulaire, voire contourner

les phénomènes de résistance. Avec les récents progrès dans le domaine de la chimie des matériaux, il est, en outre, possible de combiner dans la même nanotechnologie des propriétés différentes : des systèmes hybrides organiques et inorganiques (les nanoMOFs) permettent, par exemple, de suivre le parcours des particules dans l'organisme et d'ajouter à la fonction thérapeutique la possibilité de faire du diagnostic au travers de l'imagerie.

Nous disposons aujourd'hui de nombreuses nanoparticules aux architectures extrêmement variées et qui peuvent être adaptées aux traitements de différentes pathologies. Les liposomes, par exemple, sont des nanovecteurs composés de phospholipides - les constituants de la membrane cellulaire - formant des bicouches qui encapsulent un compartiment aqueux. Un inconvénient des liposomes étant leur instabilité, des polymères biodégradables sont aussi utilisés pour obtenir des nanoparticules plus stables. Celles-ci formeront dans certains cas un système matriciel (nanosphères) dans lequel le principe actif est inséré et dans d'autres cas un système réservoir (nanocapsules).

A ce jour, la plupart de nanomédicaments sont créés par encapsulation physique (l'absorption, le piégeage, etc.). La limite de ces techniques est que le taux de charge en principe actif par rapport au matériel vecteur est très faible (souvent de l'ordre de 1%), et qu'il est donc nécessaire d'administrer 100 mg de polymère au patient pour qu'il reçoive seulement 1 mg de principe actif. Pour cette raison, les nanomédicaments ne s'appliquent qu'à des molécules étant très actives à très faibles doses. Une nouvelle technologie, que nous avons expérimentée à l'Institut Galien Paris-Saclay, marque le passage du paradigme de l'encapsulation physique à celui de l'encapsulation chimique grâce à un lipide présent dans notre organisme - le squalène - ayant la propriété unique d'adopter en milieu aqueux une conformation moléculaire extrêmement compacte. En tirant parti de cette conformation, nous avons couplé chimiquement le squalène avec

de nombreux principes actifs. Dans l'eau, ces bioconjugués vont s'auto-organiser, par chimie supramoléculaire, sous la forme de nanoparticules. Je reviendrai sur ce type de nanomédicament et sur ses propriétés plus loin.

Lorsque l'on administre des nanomédicaments par voie intraveineuse, l'organisme les désigne comme des éléments exogènes - comparables à des particules virales - envers lesquelles se déclenche une réaction immunologique. Celle-ci se traduit par l'adsorption de protéines plasmatiques - parmi lesquelles les opsonines - à la surface des nanovecteurs. Au niveau du foie, les cellules de Kupffer, dotées de récepteurs, vont immédiatement reconnaître les nanoparticules ainsi « opsonisées ». Résultat : les nanoparticules chargées en médicament vont s'accumuler au niveau du tissu hépatique : elles constituent ainsi une formidable « navette » pour délivrer des médicaments au niveau du foie et traiter des maladies hépatiques graves. Pourtant, lorsque l'on veut cibler d'autres tissus ou organes que le foie, il faut trouver la manière d'éviter ces phénomènes d'opsonisation. Dans ce but, les chercheurs ont fonctionnalisé des nanovecteurs avec des polymères très hydrophiles et flexibles - les polyéthylèneglycols - qui empêchent les opsonines de s'absorber à la surface des nanoparticules. Au lieu d'être captés par le foie, les nanovecteurs vont demeurer plus longtemps dans la circulation générale. Et c'est une chance, car au niveau des tumeurs, la paroi des vaisseaux sanguins (l'endothélium vasculaire) est souvent beaucoup plus perméable qu'au niveau des autres tissus. Cela va permettre d'adresser l'agent thérapeutique sélectivement au niveau tumoral. Alors que ces nanovecteurs dits de « deuxième génération » induisent une circulation plasmaticque prolongée, la nouveauté introduite par les nanovecteurs de « troisième génération » est qu'ils permettent de faire de l'adressage au niveau cellulaire et même moléculaire. Dans ce cas aussi, la chimie a un rôle central : par diverses méthodes, il est possible de fonctionnaliser l'extrémité des polyéthylènes glycols avec des ligands capables de reconnaître des

marqueurs ou des antigènes qui s'expriment plus fréquemment, ou uniquement, au niveau du tissu cancéreux. Cela permet de cibler les cellules cancéreuses d'une manière extrêmement précise.

L'article va décrire, au travers de certains travaux de la littérature et de ceux de l'auteur, quelques exemples de recherches dont la visée thérapeutique (ie. Homme réparé) est incontestable mais qui pourraient aussi être détournés au profit d'une augmentation des capacités humaines (ie. Homme augmenté)

Pathologies du système nerveux central : du traitement à l'amélioration des performances ?

Or, certains nanomédicaments utilisés à titre expérimental pour le traitement de pathologies du système nerveux central, amènent à se poser la question de savoir s'ils ne pourraient pas aussi être utilisés en vue d'augmenter artificiellement certaines capacités humaines. C'est le cas, par exemple, des nanomédicaments que nous avons développés pour traiter le trauma de la moelle et l'ischémie cérébrale en utilisant l'adénosine, une molécule qui joue un rôle important dans le métabolisme des cellules et qui est également un neuromodulateur. L'industrie pharmaceutique s'est beaucoup intéressée à cette molécule pour son activité pharmacologique potentielle dans le traitement de nombreuses pathologies du système nerveux. Toutefois, lorsqu'on administre cette molécule par voie intraveineuse, elle est très rapidement dégradée (son temps de demi-vie plasmatique est seulement de 10 secondes). Après administration sous forme de nanoparticules d'adénosine-squalène, l'adénosine est protégée et les résultats thérapeutiques obtenus ont été spectaculaires. Ces nanoparticules ont, en effet, été testées dans un modèle expérimental pré-clinique chez le rat, avec paralysie des membres postérieurs, pour le traitement du trauma de

la moelle. Alors que le traitement par l'adénosine libre ne produit aucune amélioration des scores neurologiques, les nanoparticules d'adénosine-squalène ont permis d'améliorer progressivement le score neurologique et entre trois et cinq jours après le traitement, l'animal a récupéré la totalité de ses capacités de marche. Face à ces résultats, la question de la distinction entre homme réparé et homme augmenté, évoquée plus haut, se pose : ce type de recherche peut-elle déboucher sur le développement de nanocomposés capables aussi de produire des hommes supérieurs en termes de marche ? D'autre part, il a été démontré que les nanoparticules d'adénosine-squalène étaient aussi susceptibles de traiter, chez l'animal, l'ischémie cérébrale grâce à une amélioration de la microcirculation cérébrale. Mais alors ne serait-il pas aussi possible d'augmenter par le même traitement les capacités cognitives d'individus sains ?

Traiter la douleur souvent associée aux maladies chroniques et au vieillissement de la population constitue un autre défi majeur de la médecine au XXI^e siècle. Même s'ils constituent souvent un traitement efficace contre la douleur, les opiacés (morphine et ses dérivés) peuvent néanmoins entraîner de la tolérance et de la dépendance. Aux Etats-Unis, par exemple, 11 millions de patients présentent des troubles addictifs sévères suite à la mauvaise utilisation des opiacés avec une mortalité journalière d'environ 175 personnes. Les endorphines, neuropeptides endogènes, qui sont sécrétées par les neurones permettent de réduire les sensations douloureuses, voire de produire naturellement une sensation de bien-être (par exemple, après une activité sportive intense). Les endorphines agissent comme neurotransmetteurs en se fixant sur certains récepteurs opiacés (un peu différents de la morphine) sans induire pour autant de la tolérance. Pourtant, ces neuropeptides ne peuvent pas être utilisés pour traiter la douleur car ils sont métabolisés en quelques minutes après leur administration. Dans une étude conjointe de l'Institut Galien Paris-Saclay, en collaboration avec l'Insti-

tut de psychiatrie et neurosciences de Paris (Inserm/Université Paris Descartes), nous avons eu l'idée de résoudre ce problème en couplant une endorphine (la leu-enképhaline) au squalène pour obtenir des nanoparticules de Leu-enképhaline-squalène. Il a été montré que ce nanomédicament induisait, chez le rat, un effet anti-douleur important et prolongé avec une durée d'action supérieure à celle de la morphine. L'imagerie a, par ailleurs, montré que ces nanoparticules étaient capables de délivrer le neuropeptide spécifiquement au niveau de la zone inflammatoire douloureuse ; ils évitent ainsi les effets centraux responsables des phénomènes d'addiction. Grâce à l'utilisation d'antagonistes des récepteurs aux opiacés, ne pénétrant pas la barrière hémato-encéphalique, il a été observé que, contrairement à la morphine, les nanoparticules de leu-enképhaline-squalène épargnaient le tissu cérébral et agissaient donc exclusivement au niveau des récepteurs périphériques.

Une autre recherche, publiée par des chercheurs de l'université de Tokyo, a montré qu'il était aussi possible d'utiliser des nanoparticules pour effectuer de la stimulation neuronale. En injectant au niveau intracérébral, des nanoparticules à base de lanthanides (qui absorbent dans le proche infrarouge et émettent en lumière visible), il est possible d'activer les canaux de rhodopsine dans les neurones.

Ces technologies posent, bien évidemment, un certain nombre de questions : le traitement de la douleur pourrait-il conduire à des humains plus résistants que les autres à la douleur, avec des applications, par exemple, dans le domaine militaire ? Et la possibilité d'induire une stimulation neuronale ne pourrait-elle pas déboucher sur des utilisations contestables du point de vue de l'éthique ?

Nanotechnologies pour le traitement du cancer et thérapie génique : innovations et questions éthiques.

Les nanomédicaments utilisés pour traiter les maladies néoplasiques constituent une autre source de questionnement concernant leur possible usage non seulement pour réparer mais aussi pour augmenter l'individu. Comme il a été dit plus haut, quand on ne les modifie pas en surface, toutes les nanoparticules sont captées par le foie, après administration intraveineuse. Au début des années 2000, nous avons donc eu l'idée d'utiliser des nanoparticules à base d'un polymère biodégradable (le polyalkylcyanoacrylate) pour encapsuler des médicaments anticancéreux et traiter le carcinome hépatocellulaire, une tumeur primitive du foie. Dans cette pathologie, les cellules cancéreuses sont capables de lutter de manière très efficace contre toutes les chimiothérapies. En encapsulant le médicament - en l'occurrence la doxorubicine - dans ces nanoparticules, le médicament anticancéreux devient invisible aux protéines d'efflux que les cellules tumorales utilisent pour expulser l'anticancéreux et lui résister. L'utilisation de ce nanomédicament permet alors de restaurer la sensibilité des cellules cancéreuses à la chimiothérapie et de contourner le phénomène de résistance. Des expérimentations menées *in vitro* sur culture de cellules (tumorales) et *in vivo* sur des souris transgéniques développant spontanément un hépatocarcinome, ont montré une amélioration de l'activité anticancéreuse avec une apoptose accrue des cellules tumorales. Les résultats de cette recherche nous ont permis de créer une start-up (Bioalliance) et de conduire un essai clinique de phase 1, pour tester la tolérance du nanomédicament chez le patient. Ensuite, un essai de phase 2 a été réalisé avec un taux de survie de 89% à distance de 18 mois pour les patients présentant un hépatocarcinome et traités par les nanoparticules de doxorubicine (Li-

vatagR), la survie n'étant que de 54%, pour les patients qui avaient été traités par chimioembolisation, le traitement standard à l'époque de ces essais. Le LivatagR a enfin été testé dans un essai clinique de phase 3 multicentrique (conduit à la fois dans différents pays européens, ainsi qu'aux Etats Unis, au Canada, et en Chine) et a obtenu le statut "Fast Track" de la FDA (Food and Drug Administration). Il a, en effet, été montré que le traitement par les nanoparticules de doxorubicine présentait une excellente tolérance, même si la survie des patients n'était pas améliorée par rapport à la multichimiothérapie, devenue entre-temps le traitement de référence.

Si la taille des nanoparticules est très importante, notamment pour pouvoir les administrer par voie intraveineuse, leur forme peut aussi avoir un rôle important car elle peut changer la pharmacocinétique et la biodistribution du médicament encapsulé. Nous avons ainsi créé de nouvelles nanoparticules, non plus sphériques (comme les précédentes) mais allongées, sous la forme de bâtonnets. Cette morphologie particulière a été obtenue en couplant chimiquement la doxorubicine au squalène. Injectées par voie intraveineuse, ces nanoparticules, demeurent beaucoup plus longtemps dans la circulation générale et, par comparaison avec le traitement par la doxorubicine sous forme libre, elles induisent des concentrations plasmatiques en médicament considérablement plus élevées avec, comme conséquence, une meilleure perfusion au niveau des tumeurs. De plus, sous forme de nanoparticules, la toxicité cardiaque de la doxorubicine est fortement diminuée, ce tissu étant épargné par la biodistribution de la molécule encapsulée. Ceci constitue un élément très important en cancérologie, car la réduction de la toxicité permet l'administration d'une dose plus importante de l'anticancéreux avec pour conséquence un traitement plus efficace.

Les nanotechnologies peuvent également jouer un rôle extrêmement important en thérapie génique, pour le traitement du cancer, de

certaines maladies génétiques ou de pathologies du système nerveux central. La thérapie génique agit de deux manières : (i) par le remplacement d'un gène manquant ou (ii) dans le cas de gènes mutés ou de certains oncogènes, par l'usage d'ARN interférents (ARNi), soit pour inhiber complètement l'expression du gène, soit pour éditer le gène et corriger les anomalies génétiques. Cette approche thérapeutique est pourtant de réalisation difficile : d'abord parce que l'ADN et les fragments synthétiques d'acides nucléiques, sont facilement dégradés après administration. Ensuite, parce que les gènes et les acides nucléiques, à cause de leurs dimensions et du fait de leur charge négative, ne traversent pas les membranes cellulaires. C'est la raison pour laquelle ces molécules doivent être administrées à l'aide de virus (thérapie génique virale) ou grâce aux nanotechnologies (thérapie génique non virale). C'est ainsi qu'un nanomédicament - le Patisiran - a été développé récemment par une start-up américaine (Moderna) pour le traitement de l'amylose transthyrétine. Ce nanomédicament est basé sur l'encapsulation d'un ARN interférant dans des nanoparticules lipidiques ; il a radicalement changé la vie de nombreux patients : plus de la moitié (56%) des malades traités montrent, en effet, une amélioration considérable des symptômes de la maladie.

La possibilité de manipuler les gènes, en transfectant un gène manquant ou en inhibant un gène délétère, offre de fantastiques perspectives pour le traitement de maladies incurables. Pourtant, introduire des gènes au coeur des cellules à l'aide de nanotechnologies ou de virus doit conduire le chercheur à s'interroger sur les autres orientations auxquelles ces approches pourraient conduire.

Les innovations des nanovecteurs pour le diagnostic et la médecine personnalisée

Les avancées récentes dans le domaine de la chimie des matériaux ont mis à notre disposition des nanovecteurs sensibles à des stimuli endogènes ou exogènes, capables de déclencher la libération d'un principe actif encapsulé au bon moment et au bon endroit de l'organisme. Au sein d'une tumeur existe une légère diminution de pH (acidification) ou une activité enzymatique absentes des tissus normaux. Il est alors possible de tirer parti de ces stimuli pour développer des nanovecteurs qui libèrent le médicament uniquement lorsque ces éléments déclencheurs sont présents. Certaines nanotechnologies peuvent également être sensibles à des stimuli exogènes comme un champ magnétique extracorporel, de la lumière, des ultrasons ou une augmentation de température. Ceux-ci peuvent être utilisés pour libérer le médicament, détruire un tissu malade ou encore faire de l'imagerie.

En collaboration avec l'Institut Lavoisier (G. Ferey et C. Serre), nous avons développé des nanomédicaments hybrides organiques/inorganiques, les nanoMOFs, pour des applications médicales. Il s'agit de clusters d'oxyde de fer complexés avec des diacides. Suivant le poids moléculaire et la flexibilité du diacide, il est possible de construire des nanoparticules hyper poreuses qui permettent d'encapsuler des molécules diverses, en particulier le busulfan, un médicament très utilisé en oncopédiatrie (le Busulfan) qui n'avait jamais pu être encapsulé dans des nanovecteurs auparavant. Cette technologie permet d'unir la fonctionnalité thérapeutique (par l'encapsulation d'une molécule

anticancéreuse) à l'imagerie, car l'oxyde de fer permet non seulement de mieux visualiser le nodule cancéreux, grâce à un hypersignal mais aussi de savoir si les nanoparticules ont bien ciblé l'intérieur de la tumeur. Il s'agit de ce que l'on appelle le « nanothéragnostique » qui en combinant une fonctionnalité thérapeutique et une fonctionnalité diagnostique ouvre le champ à une médecine personnalisée.

Conclusion¹

Les nanomédicaments constituent une exception au sein du débat éthique autour des nanotechnologies, dans la mesure où ils obéissent aux mêmes règles de mise sur le marché que tout autre médicament. Avant d'être mis à disposition des patients, un nanomédicament est soumis à un processus très encadré dans lequel il faut, avant toute chose, démontrer que le rapport bénéfice-risque est très en faveur du premier terme, c'est-à-dire qu'il présente une activité thérapeutique beaucoup plus importante que le risque de toxicité. Dans de nombreux autres domaines d'application des nanotechnologies, l'évaluation du rapport bénéfice/risque n'est pas toujours prise suffisamment en considération, comme pour les produits agroalimentaires ou cosmétiques.

Il est également important de noter qu'en matière de médicament, les manipulations qui n'ont pas pour objectif de guérir ou de réparer un handicap, présentent souvent des dommages collatéraux plus ou moins graves. C'est le cas des médicaments administrés pour accroître les capacités de travail, le plaisir, la résistance au sommeil etc. Ils aboutissent très souvent à de la dépendance, des troubles du sommeil, voire des perturbations d'ordre psychiatrique.

¹ Pour en savoir plus - grand public :
- P. Couvreur « Les nanotechnologies peuvent-elles contribuer à traiter des maladies sévères » Leçon inaugurale du Collège de France, Fayard (2010)

- P. Couvreur « Les nanotechnologies : une approche originale pour la conception de nouveaux médicaments », La Lettre de l'Académie des Sciences, N° 39, pp 12-15 (2017)

Il s'agit aussi du cas des anabolisants, utilisés pour améliorer les capacités sportives (musculaires ou cardio-pulmonaires) qui peuvent entraîner ultérieurement des pathologies cancéreuses. Le surdiagnostic représente également un risque pour la santé : les enfants avec un déficit de l'attention (dits hyperactifs) ont été, par le passé, traités trop systématiquement à la Ritaline.

Pour conclure, il ne faut pas non plus négliger les effets sur l'identité : l'individu, quand on veut l'augmenter, peut se retrouver finalement amputé de sa personnalité qui est faite de ses forces mais aussi de ses faiblesses.



**TABLE RONDE : LES
ENJEUX SOCIÉTAUX
DES TECHNOLOGIES
ÉMERGENTES**

DÉBAT

Table ronde : Les enjeux sociétaux des technologies émergentes

Intervenants

YANNICK BLANC

Yannick Blanc est haut fonctionnaire au ministère de l'intérieur, où il a notamment été chargé des élections, des cultes, des associations et fondations. Il a été préfet de Vaucluse et du Val d'Oise et a présidé l'Agence du service civique de 2016 à 2019. Engagé à titre personnel dans le champ de la prospective et de l'innovation sociale, il a présidé la Fonda, think-tank du monde associatif et la Société française de prospective; il préside aujourd'hui le Carrefour des innovations sociales et l'association Futuribles International. Il a publié *Après le Léviathan, l'État dans la grande transition*, la Fonda éditions, 2016.

ELISABETH GROSDHOMME

Diplômée de l'École normale supérieure, de l'Institut d'études politiques de Paris et de l'École nationale d'administration et agrégée de Lettres modernes, Elisabeth Grosdhomme dirige Paradigmes et caetera, une société qu'elle a fondée en 1998, spécialisée dans la prospective et l'innovation. Elle était précédemment inspecteur des finances, puis conseiller technique au cabinet du Premier ministre et enfin chef de l'unité de marketing et communication externe de l'INSEE. Parallèlement à cela, Elisabeth Grosdhomme a exercé depuis 2003 divers mandats d'administrateur dans des sociétés cotées ou non cotées (Société générale, SNCF Réseau, Safran, Bongrain, Ciments Français, Elsan).

ALAIN BRAVO

Alain Bravo est diplômé de l'École Polytechnique (1965) et de l'École nationale supérieure des télécommunications-Paris (1970). Il a d'abord exercé au ministère des Postes et des Télécommunications pendant quinze ans. Président fondateur de SFR en 1988, il a ensuite été directeur de la recherche et de la technologie du groupe Alcatel jusqu'en 2001. Pour l'ANRT Il a dirigé la phase 1 de l'opération FutuRIS, prospective sur le système français de recherche et d'innovation (2003-2004). Directeur général de Supélec de 2004 à 2013, il a présidé la Fondation de coopération scientifique Digiteo-Triangle de la physique pour le projet du Plan Campus Paris-Saclay en 2008-2009 et il a contribué à la création de Centrale-Supélec dont il a été le président de transition en 2015. Il est actuellement Président honoraire de l'Académie des Technologies.

NICOLAS CHAILLET

Ingénieur Télécom Physique Strasbourg et docteur de l'Université Louis Pasteur de Strasbourg, Nicolas Chaillet était jusqu'à récemment professeur à l'Université de Franche-Comté à Besançon, où il a supervisé un groupe de recherche en microrobotique depuis 1995. Il a été directeur du Laboratoire d'Automatique de Besançon en 2007, directeur de l'Institut FEMTO-ST en 2012 et président de la COMUE Université Bourgogne Franche-Comté de 2016 à 2018. Ses activités de recherche sont dédiées à l'automatique, à la microrobotique et aux microtechnologies. Il a rejoint le Ministère de l'Enseignement supérieur, de la recherche et de l'innovation en octobre 2019 en tant que chef du Département de la stratégie de la recherche et de l'innovation et adjoint au Directeur général de la recherche et de l'innovation.

Modérateur

FLORIAN FORESTIER

Florian Forestier est docteur en philosophie et conservateur à la bibliothèque nationale. Il est directeur des études du think tank #leplusimportant pour lequel il a dirigé plusieurs démarches et rapports, consacrés aux plateformes d'emploi, à l'intelligence artificielle dans l'éducation ou à la diffusion de la culture scientifique. Il est auteur de nombreux articles et de plusieurs ouvrages, en particulier *La phénoménologie génétique de Marc Richir* (2014), *Le réel et le transcendantal* (2015) et *Désubériser* (2020). Il a participé à l'élaboration du programme recherche en sciences humaines de la Stratégie nationale autisme. Il est également conseiller scientifique du projet TESaCo.

Florian Forestier

Nous allons commencer la deuxième session de cette journée. Il nous a paru intéressant d'ouvrir la réflexion, en examinant les phénomènes dont nous parlons, du point de vue de ce que Daniel Andler a appelé tout à l'heure « les responsables ». Les décideurs, les acteurs de premier plan, c'est-à-dire ceux qui ont directement affaire à ces transformations, qu'ils en soient directement en charge, qu'ils en affrontent les incidences dans leurs activités, ou qu'il s'agisse de décisions de politique publique, de stratégies industrielles, de politiques d'innovation et de recherche, de gouvernance et de mode de gouvernance. Pour cela nous avons rassemblé quatre personnalités venant de ces différents horizons : haute fonction publique, société civile, industrie, instance académique, parfois et souvent tout cela à la fois. Il n'est pas question d'apporter des réponses toutes faites ou faire des exposés ex-cathedra, mais bien de faire émerger des questions, des problématiques, de commencer à élaborer un langage commun, bref de construire les conditions d'une intelligence collective pour examiner les phénomènes multifactoriels dont TESaCo traitera tout au long de ces trois ans. Je cède d'abord la parole à monsieur Yannick Blanc qui est président de Futuribles International, qui a été haut fonctionnaire au ministère de l'intérieur, préfet du Vaucluse et du Val d'Oise et qui va parler de la longue histoire des technologies et de la prospective.

Yannick Blanc

Merci de votre invitation Daniel Andler, je voudrais commencer par faire deux observations sur ce que vous avez dit tout à l'heure. La première c'est que vous avez cité la prospective parmi les savoirs académiques que vous entendiez mobiliser, or la prospective n'est pas un savoir académique. À l'exception de la chaire de prospective du Conservatoire national des arts et métiers, il n'y a pas de diplôme de prospective, il n'y a pas de maîtrise de prospective, il n'y a pas de cours dans les universités, il n'y a pas non plus de section prospective au CNRS. Bref, la prospective est un passager clandestin dans cette démarche. La prospective comme façon de mobiliser des questionnements, de croiser des savoirs et d'associer savoirs académiques et questionnements au service de l'action, a toute sa place dans la perspective que vous avez tracée. Je voudrais commencer par une première objection de projectiviste à votre exposé. Vous avez commencé en mettant en valeur la grande accélération des transformations technologiques et vous avez achevé votre exposé en disant "et ça va continuer". Eh bien moi en tant que prospectiviste, je n'en suis pas sûr. Si la prospective peut apporter quelque chose à cette réflexion, c'est en examinant les différents scénarios possibles. Il y a en effet un scénario de poursuite de l'accélération, en prolongeant les tendances observées au cours de la période immédiatement passée, mais ce n'est pas le plus intéressant, ni le plus pertinent. Il est extrêmement difficile de savoir si l'on est à la fin d'un cycle, ou au milieu d'un cycle, on peut être dans une situation où l'accélération se poursuit, les courbes continuent de s'accroître, mais on peut aussi être à un moment de pic, comme on dit dans les théories de l'effondrement à propos du pic pétrolier. On peut très bien être à un moment de pic technologique et il n'est pas exclu qu'on ait des impasses, des refus, des phénomènes de rejet d'un certain nombre de produits technologiques, après cette période de vingt ans d'enthousiasme collectif. Vous avez cité le smartphone, de toute l'histoire humaine c'est la technologie qui a mis le moins de temps à se répandre pratiquement à l'ensemble de l'humanité. Cet événement exceptionnel est-il amené à se prolonger, ou à se reproduire avec d'autres technologies ? Je pense à la biologie de synthèse, aux nanotechnologies, etc. Il est permis d'en douter. Il faut documenter ce doute, analyser ces phénomènes.

J'ai eu l'occasion de côtoyer le Génopole à Evry, il y a une dizaine d'années et j'ai entendu des discours sur la radicalité, la vitesse avec laquelle la génétique de synthèse allait révolutionner la production d'énergie, la médecine, le traitement du cancer etc. Dix ans après, il y a des progrès qui ont été faits, dans le traitement du cancer la génétique a joué un rôle fondamental, mais la génétique de synthèse n'a pas changé le monde en dix ans. Donc les phénomènes d'accélération, d'impacts et de rupture méritent d'être analysés avec beaucoup d'attention.

Pour proposer une contribution prospective au chantier que vous avez ouvert, je voudrais faire référence à un article publié dans *Futuribles* il y a vingt ans, dans le numéro de juillet/août 1999, mais paru initialement en anglais dans le magazine *Fortune* en 1955. C'est une année intéressante du point de vue de la prospective, pas tellement parce que c'est mon année de naissance, mais surtout parce que c'est l'année où Gaston Berger écrit son premier texte dans lequel il trace le programme de ce qu'il appelle à l'époque l'anthropologie prospective. La même année, John Von Neumann écrit un article dont le titre français est « *Pouvons-nous survivre à la technologie ?* ». Von Neumann est un personnage clé de la première partie du vingtième siècle. C'est un enfant prodige, il a réussi à être simultanément l'un des pères de la mécanique quantique, de l'ordinateur et de la bombe atomique. On peut dire qu'il est un acteur central des grandes ruptures du vingtième siècle. Il est mort prématurément parce qu'il avait participé aux expérimentations de Los Alamos, sans prendre beaucoup de précautions. C'est lui qui a déterminé l'altitude à laquelle il fallait faire exploser la bombe atomique pour qu'elle ait le maximum d'efficacité. À la fin de sa vie, c'est un homme qui se pose des questions d'éthique et qui est profondément marqué par le pessimisme, on peut le comprendre. Il fait partie de cette génération pour qui l'existence de l'arme nucléaire est une rupture qui crée une situation irréversible du point de vue de l'humanité tout entière. Il est, avec Günther Anders, l'un des deux à avoir vraiment essayé de creuser les conséquences irréversibles de cette rupture technologique.

« *Pouvons-nous survivre à la technologie ?* » est un texte passionnant, exemplaire de ce qu'est un travail de prospective. Il se trompe quand il essaye de se projeter dans l'avenir et d'essayer d'imaginer ce qui va se passer d'ici 1980. Il se trompe sur un certain nombre de choses mais il est prodigieusement intelligent dans sa façon de se tromper. Il explique qu'il y a trois phénomènes qui vont transformer radicalement la société : l'abondance de l'énergie, le volume et la rapidité des communications et les transformations climatiques, ce qui est tout de même assez bien vu. Il se trompe en partie parce qu'il imagine que l'énergie nucléaire va rapidement supplanter les énergies fossiles. Il écrit que d'ici peu, les matériaux fossiles, notamment le pétrole, ne serviront plus que de matières premières pour fabriquer des produits synthétiques, mais ne seront plus des sources d'énergie. Sur les communications, il se trompe peu et la problématique qu'il développe est très intéressante. Sur le climat, c'est passionnant, puisque l'hypothèse qu'il fait, c'est qu'avec le développement de la technologie on va pouvoir agir sur le climat et que les problèmes que ça va causer sont des problèmes analogues à ceux posés par la bombe atomique, c'est-à-dire, le fait qu'on puisse par la technologie maîtriser quelque chose d'aussi planétaire, d'aussi indépendant des frontières et des institutions et d'aussi fondamental pour la survie de l'humanité que le climat.

Relire Von Neumann est un exercice de modestie, de mise en perspective historique, qui nous permet de situer ce que la prospective peut apporter à notre réflexion, c'est-à-dire pousser les questionnements le plus loin possible, pousser les conséquences du questionnement le plus loin possible, imaginer ce qui peut advenir et en multipliant la liste des données et des scénarios, essayer d'éclairer les choix possibles par rapport à une situation économique et technologique donnée. La conclusion de Von Neumann à cet article est assez pessimiste, elle consiste à dire que les technologies ne sont ni

bonnes, ni mauvaises par elles-mêmes mais que leurs conséquences sont imprévisibles et que l'ampleur des ruptures technologiques accroît l'instabilité du monde en multipliant les incertitudes. C'est une idée que l'on peut aujourd'hui partager sans réserve. J'ajoute que l'on ne peut pas supprimer une technologie, une fois qu'elle s'est implantée dans la société, sauf « si l'humanité perd toutes ses illusions sur la civilisation technologique », hypothèse qui mérite aujourd'hui d'être analysée avec un certain soin.

Enfin, l'expression finale de l'article est qu'il n'y a pas de remède au progrès. L'idée qui traverse l'ensemble de l'article est que l'accélération des mutations technologiques va se poursuivre et qu'elle se heurtera à la limite de ce qu'il appelle l'extension des champs des opérations. Von Neumann théorise l'effet rebond, quand il analyse l'accroissement de l'efficacité des systèmes de communication et de contrôle par la cybernétique, en quelques phrases il dit que l'on pourrait se servir de ce progrès pour gagner du temps, pour faire des économies d'énergie dans le contrôle des systèmes complexes mais qu'au lieu de faire ça on augmente constamment l'échelle des systèmes que l'on veut contrôler. Il arrivera un moment, dit-il, où la surface de la planète ne permettra plus d'absorber cette augmentation de l'échelle des actions humaines. C'est une très belle vision de prospectiviste.

Je voudrais faire une troisième observation et conclure mon intervention là-dessus. Sur la logique même de la prospective technologique, il y a dans la littérature prospective une part considérable d'analyses, de scénarios ou encore d'hypothèses, qui sont fondés sur la dynamique propre des technologies et votre exposé liminaire entraine bien dans cette famille. Il me semble que ce qui manque à cette analyse, ce que la prospective pourrait y apporter, c'est de mieux comprendre comment, avec un regard historique et prospectif, l'éclosion d'une technologie et son enracinement dans une société est étroitement dépendante de la façon dont elle s'articule avec les institutions, l'organisation du pouvoir dans la société. Il faudrait par exemple regarder comment, au cours des deux derniers siècles, la façon dont les choix technologiques se sont construits ne dépend pas seulement du choix de la technologie mais de la façon dont la technologie est appropriée par les systèmes de pouvoir. L'histoire de la construction de la bombe atomique est de ce point de vue tout à fait exceptionnelle, quand on rappelle qu'entre le moment où la fission de l'atome est effectivement observée, théorisée et l'explosion de la première bombe atomique, il ne se passe que quatre ans. On voit la somme prodigieuse d'efforts, de mobilisation, d'investissement et de combinaison entre pouvoir industriel, militaire et intellectuel qu'il a fallu organiser. On a une communauté de physiciens qui fait de la science fondamentale en même temps que les militaires et les ingénieurs élaborent l'outil. On pourrait développer l'analyse sur d'autres sujets. Par exemple, pourquoi l'internet et ses outils se sont développés comme ils se sont développés dans la société avec cette vitesse-là ? Il y a une vraie controverse à développer sur le fait de savoir si la technologie a bouleversé la société ou si la société s'est approprié cette technologie parce qu'à ce moment-là, compte tenu de ce qu'était sa dynamique institutionnelle, de son évolution des mœurs, des structures anthropologiques, familiales, des modes de vie etc., on avait besoin de cette technologie pour prolonger une mutation de la société qui avait déjà commencé. C'est en tout cas la thèse que je défends. Il y a un certain nombre d'éléments qui permettent de plaider en ce sens. On pourrait raisonner a contrario sur la façon dont la Chine, dans le contexte des institutions qui sont les siennes, est en train de s'approprier les technologies numériques. La reconnaissance faciale en Chine, on voit très bien à quoi ça sert. La résonance entre la technologie et l'organisation politique est absolument transparente, on la voit à l'oeil nu. Quand j'entends un ministre français dire qu'« il faut absolument que nous investissions à fond dans la technologie de reconnaissance faciale pour ne pas prendre du retard sur cette technologie qui est essentielle », je crois comme vous qu'il y a des questions de sagesse et d'intelligence collective à mobiliser.

Dernière remarque sur l'intelligence artificielle. Je suis comme vous, [Daniel Andler] je fais partie de ceux qui ne comprennent pas grand-chose à la dimension technique. Je suis quand même frappé en lisant tout ce qui se publie dans le monde anglo-saxon, en termes de prospective liée aux entreprises, par les phénomènes de compétition et de « rivalité mimétique » que l'on peut observer là-dessus. Vous avez un ou deux gourous du management des entreprises qui commencent à écrire « l'intelligence artificielle va tout révolutionner, supprimer 40% des emplois etc. » et ensuite vous allez avoir tous les acteurs, tous les demi-habiles aurait dit Blaise Pascal, du monde de la stratégie d'entreprise qui vont imiter ce discours prophétique et donc vendable, et vous avez une déclinaison infinie, dans la *Harvard Business Review* et ailleurs de spécialistes qui vous disent « En 2020 c'est l'intelligence artificielle qui va bouleverser votre business, donc je vous propose un séminaire à 2000 dollars par personne pour vous expliquer comment vous allez maîtriser le phénomène ». Il y a des emballements mimétiques sur la technologie, sur lesquels il est intéressant de revenir parce qu'il y en a déjà eu dans le passé et qu'on les oublie très vite une fois qu'ils ont disparu du paysage. Il y a peut-être un petit travail de rétrospective à faire pour prendre de la distance par rapport aux emballements mimétiques auxquels peuvent donner lieu les enthousiasmes technologiques.

Florian Forestier

Merci beaucoup, Yannick Blanc d'avoir montré d'emblée la difficulté de notre tâche avec d'une part toute la question des imprévisibilités, des instabilités, des ruptures et l'intrication des phénomènes sociaux, technologiques et culturels. Bref, montrer toute la difficulté de ce qui nous attend et montrer aussi peut-être que les ruptures que nous craignons ne sont peut-être pas forcément celles que nous aurions à craindre, vous avez laissé planer un autre spectre en évoquant les questions climatiques et énergétiques. Sinon en contre point, pour avoir une vision plus claire des domaines technologiques, je vais passer la parole à Monsieur Alain Bravo qui parle au nom de l'Académie des technologies, qui a été directeur de Supélec, fondateur et président de SFR et qui va évoquer en particulier les travaux de l'Académie des technologies sur le Big Data et sur les différents enjeux qu'il pose.

Alain Bravo

D'abord je remercie Daniel Andler de m'avoir associé à cette réunion et à ce lancement de travaux, son exposé était très construit, très complet et je vais essayer de me situer par rapport à lui. Effectivement l'Académie des technologies travaille maintenant depuis plusieurs années sur l'intelligence artificielle, le Big Data et y compris le problème des compétences nécessaires au monde du travail du fait de ce qui a été évoqué comme étant la numérisphère. Alors, juste pour démarrer dans le concret, le 3 juillet 2019, l'Académie a adopté le rapport « Big Data et questions éthiques » et le lendemain, le 4 juillet l'Académie a manifesté son soutien à un manifeste pour l'intelligence artificielle qui était signé par huit grands groupes industriels, dans le cadre de la stratégie nationale de l'intelligence artificielle. Donc on pourrait penser qu'à la fois nous soufflons le chaud et le froid. Je pense que c'est à l'image de ce qu'est à l'heure actuelle le monde de l'innovation. Il faut être à la fois sur la vague et en même temps il faut rester lucide et vigilant.

Alors si vous regardez le rapport « Big Data », je dirais d'une certaine manière et ça va être l'objet de mon propos : comment peut-on réaliser et créer une intelligence collective dynamique ? Ce rapport a été bâti par ce que j'appellerais le croisement de dix regards. Au sein de l'Académie, cherchant à

avoir des connaissances les plus pluridisciplinaires, nous avons posé la question à un certain nombre de nos confrères sur ce qu'ils pensaient du Big Data et je vais évoquer les différents chapitres du rapport.

Tout d'abord nous avons eu quelqu'un qui est plutôt un chercheur qui nous a parlé de l'apprentissage en profondeur. Le fait qu'avec les outils numériques actuels, on est en train de voir le passage de l'intime, au privé, au public se transformer, et il a inventé un mot qui est « l'extimité », c'est plus que « l'intimité », c'est « l'extimité ». On voit qu'avec internet il y a des tentations d'exposition de sa personnalité privée qui de ce fait est en train d'évoluer. Ensuite, qu'est-ce que c'est dans le commerce la personnalisation de l'offre grâce aux « cookies » ou avec des publicités qui passent par des « filters bubbles » si ce n'est une démarche pour accéder à l'inavoué du client ?

Les Big Data dans la santé sont d'ores et déjà très développés. C'est déjà une réalité dans l'agriculture et l'alimentation. En particulier les agriculteurs sont parmi les premiers qui se sont emparés de la télématique. Après il y a un chapitre qui aborde Big Data et Finance. Il y a la relation courante avec les clients, mais aussi la gestion du prêt sur les profils, la gestion de l'assurance et puis du trading à haute fréquence. Également le profilage des risques, dont on sait que même pour les grandes banques ça peut entraîner des catastrophes. On évoque également la sécurité. Une consœur de l'Académie évoque le problème de l'évolution de l'utilisation du Big Data et en particulier des indices de publication des chercheurs pour la définition des politiques publiques de recherche. On termine avec Big Data, puissance, fragilité et limite. Je disais pendant la pause-café, que l'Académie des technologies a pris la décision de considérer que les recommandations du GIEC sont valides : c'est par rapport à des boîtes noires et vous-même avez évoqué le fait qu'il y a différents dossiers étudiés dans le GIEC. D'abord est-ce qu'on comprend très bien tous les algorithmes qui sont utilisés là-dedans ? Et deuxièmement, quel est le processus d'intégration ?

Quand vous évoquez les « hypers-objets complexes », moi j'ajouterais facilement l'énergie, ça a été dit, et j'ajouterais l'alimentation et la santé. Tout le problème c'est de construire une compréhension qui soit traçable et reproductible. Partant de là, vous verrez dans ce rapport, à la fin l'Académie des technologies fait quatre à cinq recommandations dont la philosophie est de dire que de toute manière ces recommandations sont à l'instant T, c'est ce que nous savons aujourd'hui, ce que nous pouvons imaginer des conséquences mais c'est bien évident qu'il faudra les revisiter. La deuxième réflexion est que la puissance publique doit parler, d'une manière qui soit flexible. Alors, je fais partie de ceux qui ont vécu la libéralisation des télécoms. J'ai vu la différence de la gestion de ce qui a été un code et puis les processus dit de régulation, au niveau européen, avec des auditions publiques où interviennent à la fois la puissance publique, les opérateurs, les industriels concernés et des représentants des utilisateurs. Donc je pense que sur les problèmes que nous évoquons, il est fondamental de s'en saisir, d'identifier les points critiques, d'exprimer des lignes directrices en les gardant dans un processus qui soit amendable et vivant. Ne pas considérer une fois qu'on a identifié un principe que le constat est définitif. Par exemple, quand on publie un règlement qui est par exemple la RGPD, relatif à la « protection des données personnelles », il est important de l'avoir publié et il est important de savoir que vraisemblablement on va apprendre de la mise en oeuvre et qu'il va falloir que ça évolue.

Dans la présentation que vous avez faite [D.Andler] il y a une remarque que je trouve très fondée, qui est la dimension géopolitique. L'Académie des technologies, le plus possible, essaye d'agir à l'intérieur d'une association qui s'appelle Euro-CASE, The European Council of Academies of Applied Sciences and Engineering . Là se retrouvent 23 académies. D'ailleurs, l'Europe a construit un outil

d'aide à la décision qui s'appelle SAPEA , qui émet des conseils scientifiques pour les politiques publiques. Je pense que l'Europe avec ses valeurs, dont la RGPD est exactement la traduction de la valeur que donne l'Europe aux données personnelles, à des cartes à jouer, en matière de réglementations. A nouveau il est intéressant d'observer que la RGPD s'impose d'ailleurs aux GAFA, *volens nolens* ...

Vous avez évoqué également la 5G, mais nous avons une arme que nous avons abandonnée depuis à peu près 15 ans en Europe, c'est la normalisation, avec ça on peut tenir des positions très fortes comme ce fut le cas pour la 2G. Donc ce sont des thèmes que je considère intéressants à aborder au cours de l'étude et je dois dire que j'ai commencé en disant que j'étais très heureux d'être ici et je pense que l'Académie sera très heureuse de continuer pendant l'étude. Merci.

Florian Forestier

Merci beaucoup Monsieur Bravo. À la fois d'avoir montré l'ancrage, la réalité de cette data-ification dans toutes ses activités : dans l'agriculture, dans la banque, ... Et aussi d'avoir rappelé que l'Europe n'était pas désarmée et que nous avons des atouts, qu'il y avait des questions de gouvernance que nous devons nous poser ensemble et justement, à ce sujet, je propose de laisser la parole à Monsieur Nicolas Chaillet qui est Directeur du service de la stratégie de recherche et de l'innovation au Ministère de l'enseignement supérieur de la recherche et de l'innovation, qui est professeur à l'université de Franche Comté, roboticien,... Pour à la fois nous parler de ces enjeux de formation et de recherche et en même temps de ces enjeux de gouvernance par rapport à l'accompagnement de toutes ces transformations.

Nicolas Chaillet

Bonjour à toutes et à tous, je suis très heureux d'être là. Merci à Daniel Andler de m'avoir invité. Je ne sais pas si je vais vous parler profondément de ce que vous venez de dire, en tout cas j'aborde le sujet avec toute l'humilité nécessaire, tant, à la fois, il est vaste et complexe ; et tant, en même temps, c'est une impérieuse nécessité que de l'aborder. Donc merci d'avoir cette audace, presque ce courage-là.

Pour rebondir sur ce qu'a dit Monsieur Blanc, à l'instant, sur la question de l'intelligence artificielle, le procès de la perte d'emploi par les nouvelles technologies, on le connaît bien en robotique, ça fait des décennies qu'on le fait et, à ma connaissance, il n'y a jamais eu la moindre démonstration du fait que la robotique détruisait des emplois. C'est une idée qu'il est très difficile de combattre et qu'on me ressort à intervalles réguliers. Quelques constats d'abord, certains ont été abordés... Les technologies puissantes, ce sont des technologies qui portent en elles, presque de manière consubstantielle, le meilleur mais aussi le pire. Et j'ajouterais aussi volontiers que le pire n'est jamais exclu. Tout ce que l'homme peut faire, il le fait. Et donc on a toutes les peines du monde à empêcher l'exploitation du pire. Donc il est très difficile de faire ça. Alors il y a des solutions de limitation de cela qui sont des solutions politiques, qui sont des solutions normatives, effectivement ; mais en tout cas -pardon de mon pessimisme là-dessus- c'est quand même une très très grande difficulté. Une techno-

logie ne porte en elle ni le meilleur ni le pire et c'est moins un sujet.

Le deuxième constat que je voudrais partager avec vous c'est une perte totale de contrôle technologique par les citoyens. Ce que je veux dire, c'est qu'aujourd'hui, plus personne ne maîtrise les sous-jacents technologiques des objets que nous utilisons parfaitement bien au quotidien. Ce petit objet-là (il montre son téléphone portable à l'assistance) assez léger, assez plat, je pense qu'il n'y a pas grand monde qui sait qu'au coeur de cet objet, il y a un phénomène mécanique. Un mouvement opératoire mécanique, sans lequel cet objet n'existe pas. Alors, on peut gloser à l'infini sur : « cet objet est-il mécanique, est-il électronique ? ». Il est évidemment les deux... Mais moi, en tous les cas, c'est quelque chose qui me paraît assez important parce que ça modifie le rapport de la personne et des sociétés à la technologie. On ne maîtrise plus, on ne sait plus comment fonctionnent les objets, que pour autant on utilise parfaitement bien. Tous les efforts sont faits pour avoir des interfaces les plus intuitives possibles et pour masquer la complexité technologique. Et ça veut dire en creux aussi qu'on ne les répare pas ou très peu. On ne sait pas le faire simplement. Ça c'est un élément que je voulais aussi marquer parce que ça me paraît être un fait nouveau. Le paysan avec son cheval et sa charrue... tout un chacun voit bien comment ces choses-là fonctionnent. Ce n'est plus le cas aujourd'hui et donc ça modifie les choses. Il y a cette accélération aussi, assez incroyable, et je rejoins à nouveau Monsieur Blanc sur le fait que je ne suis pas sûr, ni d'un côté, ni de l'autre, mais en tout cas je m'interroge, sur la continuation de l'appétence technologique ou de l'exploitation technologique dans le futur. Je n'en suis pas si sûr, j'ai le sentiment qu'il y a des évolutions qui se passent sous l'effet de prises de conscience multiples, en particulier de la question de la transition écologique, on le constate, en tout cas moi je le constate tous les jours par des petites choses : jusqu'où cette évolution ira ? Je ne sais pas, et l'évolution vers toujours plus de technologie ne me paraît pas complètement assurée.

Il y a aussi une question qui est sous-jacente à tout cela à la fois sur un peu de maîtrise technologique sur les objets qu'on utilise, à la fois sur la question de la perception et le recul -puisque le sujet c'est le lien entre les technologies et la sagesse collective- donc il faut du recul pour ça et donc quelle perception on peut générer sur les objets que l'on utilise ? Peut-être deux ou trois points là-dessus. Le premier c'est qu'il ne me semble pas ; je ne suis pas spécialiste de cette affaire-là ; mon passé, mes origines sont dans l'enseignement supérieur... Il ne me semble pas que dans l'enseignement primaire et secondaire, on aborde les technologies de la manière qu'il faut. Il y a là un sujet : on ne prend du recul que sur quelque chose dans lequel on est un peu entré. L'ignorance permet difficilement de générer une acceptation maîtrisée, des questionnements et un recul. Le deuxième point, sur le même sujet, mais qui est, pour le coup une expérience que je trouve extrêmement positive et que je suggérerais à celles et ceux qui sont intéressés d'aller voir, c'est l'expérience d'un centre qui s'appelle le TUMO. Peut-être que certains ici connaissent cette expérience... C'est une expérience qui est née en Arménie à Yerevan, à l'initiative d'un couple d'Arméniens -qui avait gagné beaucoup d'argent dans le numérique aux Etats-Unis- et qui a créé une fondation pour investir dans un centre de numérique mais pas que de numérique : un centre, disons de numérique et de créativité. Alors la créativité ça peut être la musique, ça peut être le cinéma, ça peut être les jeux, ça peut être le code aussi, le développement logiciel, ... C'est complètement gratuit ; ouvert aux jeunes de 7 à 17 ans, à tous les enfants, ... C'est rigoureux, c'est à dire que ce n'est pas un Fab Lab où on prend ce que l'on veut et où on s'en va. C'est rigoureux, il y a des étapes. Et cette expérience, je l'ai vue à Yerevan par deux fois, elle me paraît extrêmement intéressante au-delà du fait qu'elle donne à une catégorie de jeunes... Et je ne parle pas de 400, 500, 1000 jeunes... Je parle de 10 000, 15 000, 20 000 enfants dans un pays qui compte 2,8 millions d'habitants - 3 millions d'habitants. Il s'agit donc d'une cohorte très nombreuse. Il faudrait l'observer sur le long terme mais ça me paraît extrêmement profitable pour le recul que ces adultes en

devenir vont pouvoir avoir sur les technologies qu'ils vont pouvoir utiliser. TUMO a eu un tel succès que c'est devenu une forme de franchise. Il y a un TUMO qui existe à Paris, sur le modèle arménien, depuis deux ans et donc il est probablement possible d'aller regarder sans faire le voyage d'Yerevan. Et puis troisième élément d'observation sur ces sujets, c'est une expérience personnelle si vous me permettez, réitérée, en tant qu'enseignant : lorsque vous enseignez des sciences exactes, ce qui a été mon cas pendant longtemps, et que vous mettez un étudiant -l'effet serait sans doute le même sur moi d'ailleurs- vous mettez un étudiant face à un ordinateur pour qu'il simule un processus physique un peu complexe et vous observez assez vite que le patron, très vite, glisse de la personne humaine vers l'ordinateur. Autrement dit, et je précise qu'il n'y a là nulle intelligence artificielle, donc ça peut donner des sueurs froides sur l'intelligence artificielle. Et c'est très intéressant à observer, assez vite, le patron parmi les deux, la personne et l'ordinateur, devient l'ordinateur. Et c'est l'ordinateur qui guide complètement la réaction de la personne qui est en face. Ça c'est très intéressant et encore une fois je ne m'exclus pas de ce type de comportement.

Là où je voudrais porter le centre de mon propos, c'est quelque chose qui me frappe et qui me paraît intéressant à regarder : la contraction du temps. C'est-à-dire que finalement, toutes ces technologies sont souvent connectées, elles génèrent une rapidité incroyable. On a accès à l'information, de manière plus ou moins qualitative mais en tout cas de manière quasi infinie. C'est évidemment quelque chose de formidable mais dans le même temps, quand on n'a accès à zéro information quand on a soif et qu'il n'y a pas une goutte d'eau, on n'éteint pas sa soif ; mais si vous mettez la tête en bas des chutes du Niagara, vous ne buvez pas non plus. Donc tout ça génère une contraction considérable du temps. Parce que les choses s'accroissent par le numérique, par les télécommunications, par les transports, et à nouveau dans la relation entre les technologies et les individus, les collectifs, les sociétés, c'est à mes yeux, quelque chose de majeur. Ça modifie complètement nos comportements et en particulier nos comportements au travail. Au fond, mes deux grands-parents, qui étaient paysans au fin fond d'un magnifique territoire rural qui s'appelle la Haute-Saône, avaient une amplitude horaire de travail considérable mais prenaient le temps des choses. Aujourd'hui, ce temps s'est réduit et on fait les choses de manière extraordinairement accélérée.

Alors, au-delà de la question du bien-être ou du mal-être dans lequel ça peut nous placer, ce n'est pas effectivement le sujet sur lequel je voudrais porter votre attention... Il y a la question de la décision, c'est-à-dire : « qui pilote ? ». La décision appelle une réflexion. Et cette contraction, elle fait tendre vers zéro, si je puis m'exprimer ainsi, le ratio qui est réflexion divisée par action. Elle fait tendre vers zéro ce ratio par les deux bouts. C'est-à-dire que vous êtes, nous sommes dans l'hyperactivité -donc l'action est considérablement grossie, le dénominateur- et du coup le temps de réflexion tend vers zéro. Et on peut le constater, pour la plupart d'entre nous, de manière assez expérimentale, vous recevez, aller, 100 emails par jour. Ce n'est pas un étiage extravagant... Parmi ceux-ci, il y a les deux tiers qu'il faut lire. Parmi ceux-ci, l'ensemble de ces emails, il y a la moitié à laquelle il faut répondre. Et parmi ces 100 emails, il y a le tiers auquel il faut répondre de manière complexe, complète, donc sur lequel il faut mûrir les choses. Expliquez-moi comment on fait cette chose-là aujourd'hui ? Le fond de mon propos, c'est que l'on a toutes ces technologies, il faut quand même que l'on réfléchisse à la manière de les utiliser et je suis moi-même un producteur considérable de messages... A nouveau, je ne m'exclus pas de la chose ; mais tout cela on doit vraiment le considérer parce que ça peut porter en soi quelque chose d'assez grave qui est que l'on ne pense plus. On est dans l'action, on est dans l'instantané. Ce monde est globalisé, ce monde est complexe, ce monde est un système, de plus en plus. Et donc on a besoin de réfléchir, on a besoin de mûrir les choses pour rapporter des décisions. Et je constate qu'au fil du temps, on accorde de moins en moins d'attention

à l'instruction d'un problème avant d'en définir les potentielles solutions. Ça c'est quelque chose qui me frappe et qui accompagne aussi les technologies et je suggérerais aussi volontiers un sujet de réflexion autour de ça. Alors je pense que j'ai été déjà beaucoup trop long, je ne vais peut-être pas... Oui, alors je m'arrête.

Florian Forestier

Merci beaucoup Monsieur Chaillet. On voit que certaines lignes de force commencent à émerger, plusieurs bifurcations possibles ; pas nécessairement d'accélération ; les technologies qui ont des effets immaîtrisables et qui, en plus, privent les usagers de leurs maîtrises et qui nous placent dans des états d'alerte permanents. Comment faire dès lors face à cette contraction, et individuellement, et collectivement pour reconstruire des processus d'appropriation et de décision. Il semble que c'est justement ce dont va parler Madame Elisabeth Grosdhomme qui est directrice générale de Paradigmes, qui a été haut-fonctionnaire, qui a été conseillère au cabinet du premier ministre et qui va évoquer cette question de construction de mode de gouvernance participative et des façons de rendre les citoyens capables réellement d'y contribuer.

Elisabeth Grosdhomme

Je vais poser des hypothèses ou donner des exemples plutôt que de prétendre vous dire comment on va s'en sortir ! A l'invitation de Daniel Andler, je concentre mon propos sur deux points précis, à la manière d'un acuponcteur : le premier, c'est le point où la sagesse s'exprime, c'est-à-dire ce point de bascule où l'on passe d'une réflexion en termes « vrai ou faux » à une réflexion en termes « bien ou mal ». Qui s'occupe de ce point de bascule ? Tesaco nous suggère que cette responsabilité doit être collective. J'aimerais qu'on ne mette pas entièrement de côté la sagesse individuelle, mais acceptons ici l'idée du collectif. Comment s'effectue cette bascule et qui s'en occupe ? Ma deuxième aiguille d'acuponcture visera une question plus pratique : comment gère-t-on le passage à l'acte ? Si la technologie nous permet de faire quelque chose, pour beaucoup d'entre nous, c'est difficile de ne pas le faire, de se retenir de le faire. De quels moyens disposons-nous pour, malgré tout, créer un filtre et que nous ne fassions pas tout ce qui est possible pour le seul motif que c'est possible ?

Sur ce premier point, la sagesse collective, je vais commencer par un exemple. Vous connaissez tous cette fameuse expérience de pensée censée nous faire réfléchir à l'éthique du véhicule autonome : la voiture arrive, va inévitablement percuter des piétons ou faire une embardée et tuer ses passagers, qui doit-elle épargner ? Cette question ne m'intéresse pas du tout mais elle débouche sur une autre que je trouve beaucoup plus féconde. Un concours de designers a été organisé pour traiter cette situation non plus sur le plan intellectuel et spéculatif, mais concret et opérationnel, pour trouver une solution pratique au problème. Les designers ont proposé toutes sortes d'idées plus astucieuses les unes que les autres. Je vais en citer deux qui donnent matière à penser.

L'une, soumise par une agence de design qui s'appelle « MAP », part d'un principe tout simple : pourquoi le concepteur de la voiture devrait-il être celui qui décide, via l'algorithme qu'il programme, qui la voiture doit tuer ? Pourquoi ne seraient-ce pas les utilisateurs, vous et moi ? L'agence a donc

imaginé - ça s'appelle du design spéculatif -, elle a imaginé le scénario suivant : vous entrez dans la voiture et là, au lieu du volant, il y a un bouton qui peut prendre quatre positions. Ainsi c'est vous, en entrant dans la voiture, qui choisissez :

- la position numéro un, égoïste. Si quelqu'un doit mourir, ce ne sera pas moi, le passager, ce sera l'autre, celui qui est dehors. La voiture protège avant tout ses passagers.

- la position numéro deux, altruiste. C'est l'inverse : si jamais il faut choisir, sacrifiez-moi, le passager, et sauvez les gens qui sont dehors.

- la position trois, qui est appelée réaliste, consiste à sauver le plus grand nombre. Si nous sommes trois dans la voiture et qu'il faut écraser un piéton, alors vous sauvez les trois qui sont dans la voiture ; inversement, si je suis seul dans la voiture et que l'accident condamnerait trois piétons, alors la voiture doit sauver les trois piétons.

- enfin la position quatre, qui est aléatoire. L'algorithme n'a pas de rationalité prédéfinie, il décide dans l'instant, selon l'une ou l'autre des rationalités, au hasard.

Un autre cabinet de design qui s'appelle Teague a répondu entièrement différemment au défi proposé, mais toujours en partant de cette question de savoir qui décide. Son raisonnement était le suivant : il n'y a pas de raison que ce soit le programmeur-concepteur de la voiture, ni même l'utilisateur qui choisisse l'issue des événements ; les personnes les plus légitimes pour ce faire sont les gens qui habitent les territoires sur lesquels cette voiture circule. On peut ainsi imaginer, c'est d'ailleurs en train de se mettre en place, que pour que les véhicules autonomes circulent, il faille une communication « road to vehicle » : c'est-à-dire que la route enverra des signaux aux véhicules. Et dès lors, de même qu'aujourd'hui le maire a des pouvoirs de police sur le territoire de sa commune lui permettant, par exemple, de fixer la vitesse maximale de circulation dans telle rue, il pourrait demain décider, au titre de ce même pouvoir de police, que dans notre commune, on protège plutôt les enfants ou plutôt les personnes âgées. Ces règles et instructions s'imposeraient à l'algorithme de pilotage des véhicules autonomes via le dispositif de transmission « road to vehicle ».

Ces exemples n'ont d'autre but que d'illustrer la question fondamentale de la sagesse collective à laquelle Tesaco nous invite : qui est ce collectif qui doit décider ? De qui se compose-t-il ? Quel est son périmètre ? Dans l'exemple du véhicule autonome, est-ce le groupe d'ingénieurs qui a conçu la voiture ? Est-ce moi lorsque j'utilise le véhicule et son « bouton éthique », un collectif qui est alors réduit à un singleton ? Est-ce le territoire traversé, mais alors à quelle échelle : la commune, le département, la nation ?

La question de l'échelle géographique est cruciale, car elle débouche sur une autre, celle des normes culturelles. Oublions un moment le véhicule autonome et prenons pour exemple la question de la modification génétique des embryons. Manifestement, c'est une question qui ne peut se décider qu'à l'échelle planétaire parce que si c'est possible de le faire ailleurs mais pas ici, les gens qui veulent bénéficier de cette technologie voyageront. Pensez-vous que nous réussirions à nous mettre d'accord à l'échelle de la planète ? C'est peu probable. Revenons au dilemme éthique du véhicule autonome : le MIT Media Lab a conduit de 2016 à 2018 une grande étude dont vous avez peut-être entendu parler, où l'on demandait aux personnes interrogées de dire qui la voiture devait écraser. Plutôt que de poser la question explicitement, les chercheurs ont construit un jeu vidéo comprenant treize scénarii. Les participants à l'enquête devaient jouer à ce jeu et, confrontés à des situations virtuelles précises, agir d'une manière qui révélait leurs préférences. 35 millions de personnes ont joué, sur une durée de deux ans, ce qui donne un échantillon robuste.

Résultat : l'enquête a mis en évidence deux points d'accord unanime à travers la planète. Premièrement, on sauve les humains plutôt que les animaux ; deuxièmement, le nombre compte, si on a le choix, on agit de la manière qui sauve le plus de vies. En revanche, sur tout le reste, l'opinion majoritaire diffère selon les géographies. Globalement, il y a une zone asiatique, une zone occidentale couvrant l'Europe et les Etats Unis, et une zone Sud, comprenant l'Afrique et l'Amérique latine. Ainsi les Occidentaux préfèrent sauver les jeunes plutôt que les personnes âgées, mais en Asie -et tout particulièrement au Cambodge on préfère sauver les anciens plutôt que les jeunes. Les pays pauvres pensent qu'il est opportun de sauver les gens qui ont un statut social ou économique élevé, alors que les pays riches croient en l'égalité et récusent toute différenciation sur le critère de statut ou de fortune. Les pays où prévaut une forte confiance dans les institutions ne font pas de distinction entre bons et mauvais citoyens, en revanche dans les pays avec un cadre institutionnel faible, la majorité pense qu'il faut épargner les honnêtes gens et sacrifier les délinquants. Au total, cette étude montre que l'éthique, dès lors qu'on quitte les grands principes pour entrer dans les cas d'application, n'est pas du tout universelle, mais déterminée par le contexte, culturel ou économique. Autrement dit, il n'y a pas de sagesse mondiale – ce qui va poser problème pour réguler des technologies qui, elles, ne connaissent pas de frontières.

Au-delà de la géographie, quand on réfléchit au périmètre du collectif, on bute inévitablement sur la question des espèces. Notre sagesse collective doit-elle s'arrêter aux bornes de l'humanité ou bien doit-elle intégrer les animaux ou encore les robots ? La question peut paraître saugrenue, mais regardons les faits. Il y a un parti de défense des animaux qui a des représentants élus au parlement européen. En France aussi existe un parti animaliste qui réclame la reconnaissance d'une personnalité juridique pour les grands singes, les chimpanzés, les orangs-outans, les bonobos, etc. En Hollande, un parti de même inspiration est représenté depuis 2006 au Parlement, avec 2 élus sur 75 au sénat et 5 sur 150 à la chambre des députés. Le phénomène est encore très minoritaire mais sa progression est frappante. Il pourrait faire figure d'équivalent contemporain de la controverse de Valladolid où nos ancêtres, au XVIème siècle, se demandaient si les Amérindiens avaient une âme et, par conséquent, si l'on avait le droit de les réduire en esclavage et comment il convenait de se comporter vis à vis d'eux ? Au-delà des animaux, il y a deux ans, le parlement néo-zélandais a accordé la personnalité juridique au fleuve Whanganui parce que les tribus aborigènes considèrent que ce fleuve a une âme. Même chose en Inde.

Après les êtres naturels, les robots. Une anecdote pour illustrer la problématique. Une controverse se développe depuis peu aux Etats-Unis sur le point de savoir comment il convient de s'adresser à un assistant vocal - si vous en avez, comme un Google Home ou une Alexa. Deux thèses s'affrontent :

- Thèse numéro un, surtout si vous avez des enfants à la maison, il est souhaitable qu'ils apprennent à s'exprimer correctement. Par conséquent, s'ils sont amenés à demander quelque chose à Alexa, ils doivent le faire avec les formes : s'il vous plaît, merci, etc. Pourquoi ? Parce que, selon cette thèse, si vos enfants insultent Alexa ou lui parlent mal, ils reproduiront ce comportement vis-à-vis de vrais êtres humains. Pour prévenir ce risque, ils convient de les forcer à s'exprimer poliment devant Alexa.

- Thèse numéro deux, exactement inverse. Non seulement il n'est pas obligatoire de s'exprimer poliment vis-à-vis d'Alexa mais il est même souhaitable de ne pas le faire. Parce qu'en fin de compte, cet assistant vocal n'est qu'un objet, un tas de plastique et de ferraille. En intimant l'ordre à vos enfants de s'adresser poliment à cet objet, vous tendez à leur faire croire qu'il est quasi-humain alors qu'il n'est qu'un artefact. Ce faisant, vous brouillez dangereusement les frontières de l'humanité.

Pour une autre illustration de ce même dilemme, transportons-nous en Asie. En Corée du Sud, il y a un jouet très populaire, une tortue robotique munie d'une intelligence artificielle qui détermine

son comportement. Elle est programmée de telle sorte que si les enfants la malmènent, la tortue se rétracte et ainsi le jouet ne fonctionne plus. Ce faisant, la controverse a été tranchée dans les faits : il faut respecter le robot, sinon le jeu s'arrête.

Un dernier exemple, venu du Japon cette fois, de cette controverse sur la nécessité, ou non, d'inclure des êtres non-humains dans le périmètre de ceux que notre sagesse collective doit protéger. Il est question ici de robots humanoïdes en forme de poupées sexuelles, d'un fabricant nommé Trottla. Point de contexte tout d'abord : la pédopornographie n'a été que récemment interdite au Japon et le périmètre d'interdiction est limité aux images mettant en scène de vrais humains, des enfants donc ; il n'inclut pas les dessins animés et pas non plus les robots, fussent-ils humanoïdes. Trottla est donc un fabricant de sex-toys, pour dire les choses crûment, qui a dans sa gamme des poupées de fillettes de 5 à 6 ans : est-ce un problème ou non ? Au regard du droit japonais, non. Mais au regard de cette sagesse collective que nous recherchons pour encadrer notre usage des technologies émergentes ?

Pour en finir avec ces exemples, la question du droit des robots ou des animaux nous intéresse ici non pas pour la protection des animaux ou des robots en tant que tels, mais pour son effet retour sur la protection des personnes : quel mal infligeons-nous à notre humanité en nous comportant de telle ou telle manière vis à vis des animaux ou des robots ? De même, elle nous intéresse non pas au titre de la protection de tel ou tel individu, mais au titre de ce qu'elle induit en termes de construction d'un collectif harmonieux, apaisé, solidaire. Je conclus ici le premier point de mon exposé, sans réponse mais avec des questions qui, j'espère, sont à présent plus parlantes : quand on dit sagesse collective, qui est le collectif qui décide ? Et quel est le périmètre du collectif dont on veut sauvegarder la dignité, l'humanité, etc. ?

Le deuxième point de mon exposé se place davantage sur le terrain pratique. Admettons que nous disposions de cette sagesse collective. Nos actions, nos décisions n'en sont pas moins individuelles. Donc comment cette sagesse collective se propage-t-elle jusqu'à moi, hic et nunc, au moment de décider ou d'agir ? L'idéal des éthiciens, c'est que la sagesse arrive jusqu'à moi par l'effet de la morale, des valeurs qui déterminent mon action : imprégnée des notions du bien et du mal, je vais faire ou ne pas faire ceci ou cela. Ce serait l'idéal mais compte tenu de ce qu'on a vu précédemment, la variabilité des valeurs, la difficulté à s'abstenir lorsque le progrès semble proposer des solutions à nos besoins, on peut douter que cela suffise.

C'est pourquoi il existe une deuxième courroie de transmission de la sagesse collective à nos actes individuels : la loi. Cette fois, il n'est plus laissé à mon appréciation personnelle de savoir si je peux faire ceci ou cela, mais la loi, censée résumer la sagesse collective en autorisations et interdictions, va cadrer ma marge de manoeuvre. La difficulté ici est de nature différente : c'est celle de la gouvernance, de la légitimité des institutions, de leur capacité à décider. Sans même parler d'une introuvable gouvernance internationale, même l'échelon national est en proie à des fragilités à cet égard. Les lois, particulièrement quand elles touchent à des enjeux éthiques, ont du mal à s'imposer : leur élaboration est controversée, leur mise en application est très inégale. A tel point que l'on voit aujourd'hui se développer de multiples réflexions et expérimentations sur d'autres moyens que le vote majoritaire pour construire la norme collective : les conférences citoyennes, le vote quadratique, les mécanismes de décision par consentement plutôt que par consensus, etc.

Si la loi elle-même ne suffit pas à déployer la sagesse collective dans nos actions individuelles, reste alors, pour le meilleur ou pour le pire, les inducteurs de comportement, ce qu'on appelle le nudge. Ce sont finalement les machines elles-mêmes qui, de plus en plus, induisent nos comporte-

ments. En voici quelques exemples. Vous avez en Suède une néo-banque, appelée Doconomy, proposant une carte de paiement qui se veut gardienne de l'environnement. Au fur et à mesure que vous faites des achats, votre carte bancaire calcule l'empreinte carbone de vos dépenses ; quand vous avez atteint un certain plafond, elle se désactive, vous ne pouvez plus rien acheter. En somme, votre carte de crédit devient votre bonne conscience, comme le Jiminy Cricket de Pinocchio, pour vous inciter à préserver la planète.

Autre exemple, à Taiwan, pour combattre la propagation des fake news sur les réseaux sociaux, vous pouvez, si vous le souhaitez, ajouter en ami sur votre compte de messagerie sociale, un personnage virtuel, qui est en fait une intelligence artificielle nommée Tante Meiyu pour la rendre sympathique. Tante Meiyu veillera sur vos communications et, quand vous vous apprêtez à propager une nouvelle dont la véracité est douteuse, elle vous signalera que cette information est sujette à caution. Elle le sait parce qu'elle surveille tout le reste du web. Tante Meiyu ne va pas vous interdire de diffuser cette nouvelle, elle va juste vous alerter : « Es-tu vraiment sûr que tu veux propager cette nouvelle ? Es-tu bien conscient qu'elle n'est pas absolument digne de confiance ? » Ainsi, finalement, si d'une part notre force morale est insuffisante pour nous empêcher de faire le mal -disons-le comme ça-, et si d'autre part la loi est en retard sur les événements, inadaptée, ou tout simplement inappliquée, alors ce qui va se passer, ce qui est en train de se passer, c'est que ce sont les fournisseurs, les fabricants, les industriels qui conçoivent les objets et services que nous utilisons, qui vont combler le vide. Par des règles de fonctionnement des gadgets qu'ils nous fournissent, par des sécurités qu'ils y intégreront, par des scénarios d'usage qu'ils auront validés ou pas.

Un dernier exemple pour montrer comment la boucle se boucle. C'est une histoire documentée par *China AI*, une revue chinoise sur l'intelligence artificielle, dont certains articles sont traduits et diffusés en Occident par des chercheurs de l'université d'Oxford. L'histoire débute par un drame familial : une maman chinoise avait une fille unique, âgée de 14 ans, qui est morte d'un cancer. Sous le coup de la tristesse, cette dame interpelle le laboratoire d'intelligence artificielle d'Alibaba via Weibo, un équivalent chinois de Twitter, par un message demandant s'il serait possible de créer un avatar numérique de la défunte. Remarque au passage : pour ceux d'entre vous qui regardent les séries de science-fiction, c'est l'intrigue d'un épisode de *Black Mirror*. La réalité rejoint donc l'imagination : quelqu'un dans la vraie vie demande à Alibaba de faire ce que la fiction a décrit. Les chercheurs d'Alibaba, ainsi interpellés, répondent : « Pas de problème, montrez-nous les données que vous avez ». Les données en question, ce sont les traces numériques de la fille décédée : ses emails, ses billets sur les réseaux sociaux, etc... Par ce biais, on peut savoir quels sujets l'intéressaient, comment elle s'exprimait. Une intelligence artificielle nourrie de ces données pourrait imiter la défunte : et donc l'avatar de votre fille décédée pourrait vous envoyer des emails ou des messages sur une conversation en ligne. C'est ce que la maman demandait. Si par ailleurs, vous aviez des enregistrements audio, elle pourrait aussi vous laisser des messages vocaux. Et puis si vous aviez des images vidéo, elle pourrait ... Après analyse des données transmises par la mère, les chercheurs ont déduit qu'ils avaient assez de traces écrites mais pas assez de traces audio et vidéo pour reconstituer l'avatar complet. Ils ont donc proposé à la maman : « L'avatar numérique de votre fille décédée pourrait vous envoyer, de temps en temps, des messages de réconfort ». Puis quand même, dans un moment de lucidité, ils ont eu l'idée de demander au département de psychologie d'une université à Pékin si c'était vraiment judicieux. Comment faire votre deuil si votre fille décédée continue à vous envoyer des messages ?

Le problème auquel nous sommes confrontés, c'est que si la fiction l'a imaginé, si les ingénieurs savent le faire, alors inévitablement l'un ou l'autre d'entre nous va demander le bénéfice de telle ou telle technologie. D'où la nécessité d'un peu de discernement individuel et de la sagesse collective que Tesaco nous invite à définir.

Débat : Les enjeux sociétaux des technologies émergentes

Intervenants

THIERRY MÉNISSIER

Professeur de philosophie à Grenoble et titulaire d'une nouvelle chaire "Éthique et intelligence artificielle" dans l'institut 3IA à Grenoble.

BERNARD CHARPENTIER

Vice-président de l'Académie de médecine et doyen honoraire de la Faculté de médecine Paris Sud.

JEAN LORENCEAU

Directeur de recherche au CNRS, neurosciences et sciences cognitives.

JACQUES DUBUCS

Directeur scientifique du secteur SHS au Ministère de l'enseignement supérieur, de la recherche et de l'innovation.

SONIA DESMOULIN-CANSELIER

Chargée de recherche au CNRS, Laboratoire Droit et Changement social, chercheuse associée à l'Institut des sciences juridique et philosophique de la Sorbonne et au Centre d'études des normes juridiques CENJ/Yann Thomas, EHESS

Jean-Paul Laumond

Je voudrais intervenir en rapport à votre dernière intervention Madame Grosdhomme pour souligner dans un premier temps qu'il faut être très prudent quant à ces études pseudo-scientifiques qui sont issues de laboratoires prestigieux du MIT. Je commence par dire qu'il n'y a pas une communauté de pensée dans les communautés scientifiques (et je pense que ce sera un problème à aborder dans le sens où il n'y a pas d'un côté le scientifique qui considère les faits et de l'autre la société avec ses valeurs et que nous devons mettre les faits en rapport avec les valeurs). La communauté scientifique n'est pas homogène. L'étude à laquelle vous faisiez référence sur la voiture autonome est un exercice de pensée qui est extrêmement du type anglo-saxon et qui repose sur une méconnaissance complète de l'état de la technologie. Cette étude a été relayée en France. C'est une étude internationale qui a été publiée dans la revue prestigieuse *Nature*, dont on peut considérer qu'elle publie parfois des inepties. Il faut être extrêmement prudent à ce niveau-là. Pourquoi ? Parce que, juste deux mots sur la voiture

autonome. Il y a derrière, la voiture autonome un problème de compréhension de la technologie. Qu'est-ce qu'une voiture autonome ? Tout le monde, et c'est normal, considère que c'est une voiture qui va circuler dans le quartier du Marais tel qu'il est aujourd'hui. Mais ça ne se passera pas comme ça. Si on veut des voitures autonomes qui circulent dans le Marais, on va totalement transformer le Marais. J'aime bien prendre cet exemple, à un moment donné on a voulu faire du 130 km/h avec nos voitures. Ça a été un désir d'aller plus vite avec nos voitures. Que s'est-il passé ? On a créé des infrastructures qui s'appellent des autoroutes. Et sur ces autoroutes on peut considérer qu'il y a très peu de dames âgées qui traversent et très peu de jeunes garçons qui jouent au ballon ; qui vont nous poser le problème du choix entre écraser le petit garçon ou la vieille dame. Eh bien le problème de la voiture autonome va être exactement celui-là : une machine automatique, qui va être automatique dans un milieu parfaitement contrôlé comme on a l'habitude de le faire en aéronautique. Nos avions décollent et atterrissent de manière complètement automatique et le problème de la technologie va être de certifier les technologies qui sont mises là. Donc là où je veux insister c'est sur le rapport du scientifique à la société à travers les seules possibilités qu'il a de communiquer avec elle grâce à un langage de tous les jours, compréhensible, qui ne soit pas des équations. L'idée est donc de créer une culture de la technologie qui soit partageable et participer au débat de cette manière là. Je pense qu'il est nécessaire de travailler en interdisciplinarité.

Réponse d'Élisabeth Grosdhomme

J'ai pris l'exemple de la voiture autonome car c'est un exemple qui parle à tous. En France la question posée est non-pertinente pour une raison simple : notre constitution stipule l'égalité de tous les individus devant la loi, qu'ils soient jeunes, vieux, hommes ou femmes. Au-delà du véhicule autonome, même si l'étude du MIT est critiquable, je pense que sa conclusion reste juste : il sera difficile de trouver des références culturelles communes dans toutes les géographies. Ce qui ne pose pas de problème pour les technologies dont on peut localiser l'usage, mais pour toutes celles où il y a des effets de fuite, ça va être compliqué.

Thierry Ménissier

Mon ancrage en philosophie est à la fois épistémologique, anthropologique et politique. Je voulais dire que l'article de *Nature* est épouvantable également du point de vue éthique, du fait de la méconnaissance des auteurs de la théorie morale classique. Le problème est d'avoir enfermé l'analyse éthique du problème et des usages techniques ce qui est encore très différent dans ce qu'on appelle le conséquentialisme (qui est la manière dominante de poser les problèmes aujourd'hui en éthique de l'IA quand les choses sont très appauvries). À vrai dire, tout est à faire. Je ne vais pas continuer à karchériser ce mauvais article, et cette expérience très orientée, très buzz.

Ce qui m'a frappé dans toutes les interventions c'est la mise en relation de la sagesse (qui n'est pas l'éthique) et ce qui est en crise aujourd'hui : la question de l'institution. En effet, aujourd'hui il n'y a pas de sagesse sans institution. Alors qu'est ce qu'une institution ? Ce n'est pas seulement l'État ni la loi, ce n'est pas seulement la science. C'est une structure forcément collective à laquelle on adhère et

qui donne du sens car elle dure au-delà de nos petites activités. Ainsi les religions et leurs superstructures comme l'Église font institutions. Et ces institutions durent sur de grandes durées. La loi ne peut pas être une institution car elle est le produit d'une institution particulière qui est l'État. Alors peut-être que le savoir quel qu'il soit fait institution. Toujours est-il qu'aujourd'hui ce qui risque de faire institution si on ne réfléchit pas et si on ne se mobilise pas c'est quand la technique elle-même est laissée à des opérateurs qui sont d'abord économiques et financiers (des entreprises qui ont leur raison d'être). C'est un problème majeur, je crois, de l'évolution du capitalisme que de penser des compagnies qui font institutions d'une certaine manière ou que les technologies et projets technologiques font aujourd'hui institutions. Alors que cela n'a pas été pensé comme tel. Le défi est très grand de repenser ce qui pour une civilisation globalisée ou en voie de l'être, peut faire institution.

Réponse d'Élisabeth Grosdhomme

J'observe que les entreprises, la nature ayant horreur du vide, occupent le vide pour l'instant. Ce n'est pas un idéal, y compris du point de vue des entreprises elles-mêmes. Voyez que Google, par la voix de son dirigeant Sundar Pichai, est récemment allé demander à Bruxelles que l'Union européenne veuille bien réglementer l'usage de la reconnaissance faciale. Parce qu'ils savent bien qu'à défaut d'une discipline des usages, il va y avoir des phénomènes de rejet. Donc de fait si l'État et nous les citoyens ne faisons rien, les entreprises vont occuper le vide. Mais ce n'est une situation optimale ni pour nous ni pour elles.

Réponse d'Alain Bravo

Je voudrais confirmer ce que vous venez de dire [Thierry Ménissier]. Je crois que sur ces problèmes qui sont des phénomènes très innovants, la loi n'est pas adaptée. Je constate assez fréquemment, mais je n'en connais pas l'efficacité réelle, qu'il y a l'émergence de chartes entre les "stakeholders" (parties prenantes) dans lesquels on inclut tous les acteurs concernés qui ont le droit de s'exprimer. Par exemple, quand j'ai évoqué la régulation telle que je l'ai vu émerger au niveau européen en matière de télécommunications. Au début, j'étais complètement inexpérimenté dans les auditions publiques organisées par l'Europe sur un sujet de l'époque, la 3G.

Je dis seulement des remarques, je ne donne pas de solution. À l'Académie on commence à travailler sur les changements climatiques, immédiatement j'ai cherché s'il y avait des chartes pour essayer de savoir ce que certains collectifs d'ores et déjà conviennent. Mais il faut trouver des manières nouvelles, surtout si on passe à l'échelle globale.

Réponse de Yannick Blanc

Je partage l'idée que s'interroger sur ce qui fait institution est une interrogation absolument inséparable de la problématique de cette enquête. Il y a des phénomènes de résonance ou de non résonance entre des structures institutionnelles installées ou émergentes et des technologies. Et on peut

explorer l'hypothèse que l'une des raisons de la faiblesse de l'Europe dans la compétition pour la puissance technologique c'est l'inadaptation de ses institutions à ce contexte. C'est pourquoi j'ai cité le cas de la Chine. Les technologies numériques viennent de la Californie, donc le territoire libéral, libertaire par excellence. Et on est en train d'assister à leur appropriation et leur transformation en réelle puissance par la Chine. Il faut donc analyser ces phénomènes.

Puisqu'on a évoqué la loi et les chartes, je pense que la substitution progressive de l'exercice "charte" à l'énoncé du droit par la loi est une mutation majeure de notre façon de faire institution. La charte est un phénomène universel, proliférant. C'est-à-dire que dans la vie quotidienne là où nous sommes dans nos institutions, nos entreprises et nos associations on ne peut plus rien faire sans adopter une charte, y compris au sein de l'administration publique. Le droit ne sait pas ce qu'est une charte. Et aujourd'hui il n'y a pratiquement plus d'action collective possible sans qu'à un moment ou l'autre on éprouve le besoin d'énoncer une charte. Ce phénomène à lui seul est un phénomène d'analyse de sciences politiques, de philosophie politique mais aussi de prospective particulièrement intéressant car la définition que je donne de l'institution (sous l'influence de Wittgenstein), c'est que c'est un système d'énoncés, c'est un jeu de langage. Et à partir de là, dans la façon dont l'environnement technologique modifie notre façon d'émettre et de percevoir des énoncés, il y a une fermentation dans la fabrication des institutions qui est en train de se produire mais dont nous ne savons pas vers quoi elle se dirige.

C'est là que la prospective dit : travaillons sur ces phénomènes émergents, ces signaux faibles, ces tendances et construisons des esquisses de scénarios pour voir si nous avons des choix possibles ou si nous sommes déjà pris dans un mouvement qui nous a dépassés. Les deux phénomènes existent en prospective. Il y a des phénomènes dont on ne s'est pas aperçu qu'ils allaient transformer la société (les phénomènes démographiques par exemple qui ont profondément transformé notre société y compris notre vie intime), et puis il y a des phénomènes sur lesquels on a prise. L'exercice c'est de discerner les deux catégories de phénomènes.

Florian Forestier

Juste une petite remarque : on peut dire que le juridique est un phénomène institutionnel très singulier d'occidental, c'est l'objet de travaux d'Alain Supiot, Gilles Legendre. Il y a vraiment cet horizon-là derrière, de repenser notre rapport au juridique comme institution.

Bernard Charpentier

Ma remarque concerne le moment où il y a eu le choix à la Commission européenne des différents commissariats. Quand il s'est agi de choisir le commissariat "Innovation, recherche, culture, éducation et jeunesse", il est arrivé en 23ème position alors qu'il en va de l'avenir de l'Europe. Donc j'ai quand même été choqué par cela. Dans d'autres pays ce commissariat aurait été choisi en première ou deuxième position, plutôt que la régulation du marché intérieur ou du secteur bancaire. Ce choix m'apostrophe au moins dans mon pays l'Europe et vous savez que l'Europe est la fille d'Agénor, et Agénor c'est le dieu tout puissant c'est le vainqueur.

Réponse d'Alain Bravo

Ne me demandez pas de considérer que le fonctionnement de l'Europe est satisfaisant. Ça résulte d'un jeu d'équilibre de la distribution des postes de commissaires. C'est la cour de récréation.

Réponse de Bernard Charpentier

Sauf que la commission de l'Europe c'est 100 milliards d'euros.

Réponse d'Alain Bravo

Je vous disais et j'en suis convaincu, Thierry Breton c'est quelqu'un de vertébré. Donc à mon avis quand il y a aura des arbitrages il s'en mêlera. C'est un jeu où on essaye de donner quelque chose à chacun des 27.

Réponse de Yannick Blanc

J'ai une réaction totalement différente de votre interprétation. Les grandes technologies dont on parlait tout à l'heure se sont répandues grâce à des mécanismes d'investissement, à des comportements d'investisseurs, à des outils (des outils numériques) et à toute une culture de l'investissement et du risque. Je fais une comparaison avec ce que dans le social on appelle "le contrat à impact social", le social impact bond. Les Anglais l'ont inventé, ça a fasciné quelques experts français qui se sont dit qu'il allaient l'adapter chez nous. Quelle est la différence ? Chez les Anglais, le social impact bond c'est un investisseur, un parieur qui fait de l'investissement social sur une dizaine de projets. Il a une stratégie de portefeuille. Il sait que sur ces 10 projets, il y en a 5 qui sont nuls, 3 ou 4 qui sont à peu près équilibrés et un qui va vraiment lui permettre de récupérer sa mise et à l'arrivée, il n'aura pas gagné de l'argent (car on ne gagne pas d'argent dans le social), mais il se sera imposé comme un investisseur social et aura répondu à un certain nombre d'exigences. On a essayé de faire la même chose en France, mais la bureaucratie s'est emparée de l'outil et avec l'aversion la prise de risques, elle l'a verrouillé à tous les étages. Résultat, on n'arrive pas à faire essaimer le dispositif. Une des grandes faiblesses de l'Europe aujourd'hui, une des grandes faiblesses de l'Allemagne qui en est la puissance dominante, est son incapacité à investir. Dans la mentalité collective, dans la vision du monde et dans le mécanisme de décision collective de ce qu'est aujourd'hui l'Europe il n'y a aucune capacité à prendre des risques, à se projeter dans l'avenir et à imaginer quels seront les enjeux de demain. C'est la grande différence qu'on a avec la Californie et la Chine.

Là, il y a un sujet proprement politique. Au demeurant, dans l'enjeu de la transition climatique, dans la réaction et la réponse à l'enjeu climatique, la controverse la plus intéressante est celle de l'investissement public. On voit aujourd'hui les meilleures autorités de l'économie financière mondiale dire au gouvernement européen qu'il est temps de faire du déficit public car il faut faire de l'invest-

tissement massif pour attaquer la transition écologique et les puissances européennes dire qu'il n'en est pas question. Cette capacité à investir, à faire des paris, de prendre des risques touche aussi bien le domaine de l'innovation technologique, de la recherche et l'enseignement que le domaine de la réponse climatique et écologique.

Jean Lorenceau

Je vais revenir à quelque chose de différent en proposant une question et un cadre théorique. Ne pensez-vous pas au fond qu'au fil de ces vingt dernières années nous avons glissé du cerveau 2 au cerveau 1 ? Je fais référence au livre de Kahneman¹, le cerveau 1 c'est le cerveau de l'immédiateté, celui qui traite immédiatement l'information et qui suscite émotion, décision, jugement. Le cerveau 2 au contraire est celui qui revient sur cette information primaire et ascendante et qui remet en boucle la question. Donc, quand on pense à l'immédiateté et à la compression du temps, on se dit : oui, on est sous l'effet du cerveau 1, le cerveau de l'émotion où la rationalité n'est plus présente. Alors on pourrait généraliser et se demander si on n'est pas dans un aller-retour entre le cerveau 1 qui est le cerveau de la population des gens dominés par des émotions liées à l'immédiateté de l'information, l'immédiateté des sentiments suscitée par les médias et le cerveau 2 qui serait peut-être le cerveau des institutions, celui qui doit revenir sur ce cerveau 1 pour le réguler et lui donner un contexte institutionnel acceptable. Au fond, je reprends à partir de l'exemple de la centrale nucléaire, qui est l'exemple que vous avez donné Monsieur Blanc, en disant qu'il y a eu là une formidable concentration étatique. On a mis des scientifiques, de l'armée, de l'État, on a fait quelque chose. C'était l'approche centralisée. Les technologies qui sont aujourd'hui consommées sont essentiellement décentralisées. Et la question est : y aura-t-il une recentralisation de ces données ? On a l'impression que oui, à la fois à travers les modèles économiques de la diffusion d'information, à la fois avec tout ce que vous dites et la Chine est une bonne reprise en main recentralisée. C'est peut-être au fond une alternance entre du cerveau 1 et du cerveau 2 dans une dynamique longue cette fois-ci mais qui est un dialogue avec une population qui consomme et qui est avide d'émotions rapides et de transformations rapides et qui souhaite de l'émotion et une réflexion rationnelle, raisonnée qui va venir soit des institutions soit des citoyens. Alors il y a des éléments intéressants qui sont totalement orthogonaux au développement technologique mais qui sont des changements rapides. Moi je ne suis pas sûr, mais j'ai eu l'impression, par exemple sur le changement climatique, qu'il y a eu un glissement au niveau des élections et des citoyens d'une prise en compte du climat mais plutôt à partir des cataclysmes qui ont été diffusés à la télévision (les incendies en Australie, les inondations dans le sud-est de la France, les cyclones aux États-Unis) plutôt que les recommandations du GIEC qui ont peut-être moins ébranlé l'opinion publique. Ces premiers éléments sont à nouveau dominés par le cerveau 1, celui de l'émotion, de la compassion, de l'empathie.

Un mot sur la compréhension, et donc on en revient sur la question de l'éducation. Aujourd'hui la difficulté n'est pas de savoir comment marche notre smartphone ; je pense que les gens qui utilisent le téléphone ou l'ampoule électrique au moment où Edison les avait proposés n'avaient pas une

¹ Daniel Kahneman (2011), *Thinking Fast and Slow*; trad. fr. *Système 1 / Système 2 : Les deux vitesses de pensée* (2012).

meilleure compréhension des phénomènes physiques sous-jacents. Je crois que ce qui est difficile aujourd'hui c'est de connaître les clauses du contrat qu'on signe. Au fond, quand on signe on ne va pas sur les options de la RGPD noter que tel site va vous proposer de tout accepter, tel site vous propose de ne rien accepter. Quelques mots sur l'alimentation et la santé. Les technologies sont là certes, mais le mouvement vegan et le mouvement végétarien ne sont pas liés aux technologies. Ce sont pourtant des mouvements essentiels d'opinion qui vont changer le rapport à la santé, à l'alimentation et au climat. Là nous n'avons pas affaire à des technologies mais à de l'émotion, du cerveau qui collectivement appréhende des phénomènes qui provoquent émotion, compassion, empathie ou non.

Réponse de Yannick Blanc

Moi, je fais l'analyse diamétralement opposée. Ce que j'observe aujourd'hui c'est une emprise du cerveau 1 sur les institutions. On travaille sur l'émotion, l'immédiateté, l'image, l'événement, pour le meilleur et pour le pire. Le pire c'est la démagogie et le meilleur c'est la façon dont nos dirigeants politiques ont réagi face aux attentats de 2015. Faire de la politique à ce moment-là c'était prendre des décisions opérationnelles, maîtriser le registre de l'émotion qui a été politiquement fondamental à ce moment-là. Inversement lorsque j'observe la façon dont travaillent, s'organisent, agissent collectivement les hommes et les femmes sur le terrain dans les communes, les associations, les petites entreprises, je trouve qu'il y a beaucoup plus d'intelligence collective locale dans la conception des projets, dans la capacité à imaginer des stratégies coopératives et dans le rapport au réel. Je trouve qu'il y a plus d'intelligence que d'émotion dans la population de manière répartie que dans le sommet des institutions. J'en reste sur ce paradoxe pour l'instant.

Réponse de Nicolas Chaillet

Cela rejoint tout à fait mon point sur la contraction du temps. Les institutions ont tous les défauts de la terre mais essayons de nous projeter dans un monde sans institution. Donc il est important de pouvoir les sauvegarder. Et si je reprends vos propos de maintenir ou de développer un peu de cerveau 2 dans les institutions alors il faut gérer ce temps court, cette immédiateté qui peut percuter la maturation de la réponse à un sujet complexe où la proactivité pour éviter de n'être qu'en réactivité. Pour cela je suis totalement d'accord avec vous.

Sur la maîtrise et la non-compréhension des technologies, j'entends votre sujet. Il me semble néanmoins que démystifier une technologie, en particulier dans le domaine du numérique, cela rejoint l'exemple du tumeur que je donnais tout à l'heure, c'est aussi éclairer un peu ces fameuses chartes qui sont si absconses pour bon nombre d'entre nous.

Réponse d'Élisabeth Grosdhomme

Je réagis sur cette idée de ménager un espace institutionnel qui soit un peu moins dans l'immédiateté et dans l'émotion : on l'a déjà essayé par le passé, ou on l'essaye dans différents domaines, par

des gouvernances particulières. Pensez aux banques centrales, dont l'indépendance vise à extraire la politique monétaire des emballements du débat politique. Cela fonctionne tant que, dans un système démocratique, la population n'a pas le sentiment d'être dépossédée de son pouvoir politique. Mais ne sommes-nous pas en passe de franchir ce point critique ? La défiance vis-à-vis des institutions a atteint une cote d'alerte. J'en donne un signe parmi mille autres : le succès du groupe d'activistes Extinction Rebellion . Ils sont un peu près 10 000 participants en France.

Ils prônent la désobéissance civile et organisent des camps d'entraînement à cette fin pour leurs adhérents : ils leur apprennent par exemple ce qu'il faut faire s'ils sont arrêtés. Plus prosaïquement, pour ceux d'entre vous qui habitent Paris, vous avez peut-être vu sévir dans votre rue le « Gang des dégonflés » : ce sont des activistes qui considèrent qu'il est inadmissible de conduire un SUV dans Paris car ces modèles de voiture polluent trop. Pour se faire justice, ils dégonflent les pneus de ces véhicules et laissent un petit papillon explicatif sur le pare-brise. Autrement dit : même si la loi n'interdit pas les SUV, ils considèrent, eux, que c'est inacceptable moralement. Cette résurgence des phénomènes de désobéissance civile n'est pas propre à la France. Aux États-Unis, Waymo, la filiale de mobilité autonome de Google, est confrontée à des actes de vandalisme : des gens attaquent les voitures autonomes au feu rouge à coups de battes de baseball. Tous ces exemples témoignent de ce qu'une part au moins de la population considère que les institutions ne la défendent pas et ne la représentent pas.

J'aime l'idée d'avoir des institutions sages pour tempérer une population qui serait toujours en risque d'être dominée par l'émotion, mais cela nécessite un degré de confiance institutionnelle qu'on n'a plus aujourd'hui. J'espère qu'on le retrouvera un jour.

Réponse d'Alain Bravo

C'est vrai qu'à l'heure actuelle on vit beaucoup dans l'immédiateté mais je me demande et c'est peut-être l'étude qui devrait le mettre en évidence, je crois qu'il y a une prise de conscience d'un certain nombre de ce que Daniel Andler a appelé les "hyper objets complexes". Avec la mondialisation on sait qu'il y a un grand nombre de problèmes qui nécessitent une solidarité large et la construction de solutions qui soient partagées. Alors, pour le changement climatique, il n'y a pas la même sensibilité selon les générations mais par rapport à l'émergence de ces problèmes hypercomplexes c'est là qu'on retrouve d'un seul coup la nécessité de l'action longue et qu'on revient vers la nécessité d'investissement. Mais je suis assez d'accord, les médias nous donnent une image du temps en permanence contracté. Je me demande si on est pas à l'aube de la prise de conscience d'un certain nombre de ces "hyper objets complexes".

Patrick Couvreur

En entendant la discussion je me pose la question du sens de la sagesse collective, ce qu'on va mettre derrière ce terme "collectif". On voit qu'il y a une série de minorités aux avis diamétralement opposés. Comment va-t-on pouvoir réaliser collectivement l'orientation que devrait prendre telle et telle technologie ? Je pense par exemple, et au départ ça partait d'une bonne intention, il y a quelques

années Valérie Pécresse avait organisé des débats citoyens donc collectifs concernant les nanotechnologies. Ça ne s'est pas très bien terminé. On n'a tiré aucun enseignement de ces histoires-là. Sauf que sur certains débats il y a eu des déprédations de locaux etc. Je pense qu'il faudrait qu'on réfléchisse quand même, étant donné que c'est le titre de nos discussions, sur ce qu'on doit mettre derrière "collectif" et comment on peut faire du collectif sur les technologies.

Réponse d'Alain Bravo

En particulier l'organisation des discussions de débats publics. C'est très intéressant sur le Plateau de Saclay, on avait eu un événement violent à propos des nanotechnologies. Il y a des sujets de débats publics où l'institution d'organisation de débats publics fonctionne très bien. Quand il s'agit par exemple d'infrastructures qui sont partagées par un certain nombre de collectivités. Dans le cas où on est sur un débat plus difficile à percevoir comme les nanotechnologies on peut dire qu'il y a eu une erreur de méthode à l'origine. Quand cela a été ouvert, au même moment, il y a eu un plan gouvernemental sur les nanotechnologies. Cela donnait l'idée que les débats publics ne servaient à rien. Il y a donc un art d'organiser le débat public et un savoir faire peut-être à regarder. Dans un certain nombre de domaines sur le plan pratique cela fonctionne.

Patrick Couvreur

Je peux répondre aussi. Je suis assez d'accord pour des infrastructures collectives, parce que les gens peuvent se mettre d'accord, parce qu'ils sont tous des utilisateurs et ils comprennent très vite pourquoi l'utilisation du métro par exemple peut être utile ou pas. Quand on parle de sciences et de technologies c'est beaucoup plus complexe. Derrière tout cela il faut une connaissance scientifique et technologique minimale à avoir pour comprendre les implications de ces technologies, sur ce qu'elles peuvent amener et ne pas amener. On fantasme aussi sur ce que peuvent apporter certaines technologies. Ma remarque était de dire que cela serait intéressant d'avoir l'avis de la commission nationale du débat publique sur les problèmes de cette nature.

Réponse de Nicolas Chaillet

Juste un mot du côté de la recherche et du ministère chargé de la recherche. On voit des évolutions et on essaye de lancer ces évolutions dans l'exercice de la recherche. Il ne s'agit pas de passer d'un exercice en chambre dans un laboratoire à tout le monde s'improvise chercheur du jour au lendemain. On voit des évolutions qu'on essaye de porter et que les organismes de recherche veulent aussi porter notamment dans la recherche participative. Elle consiste à impliquer d'autres catégories de population que les chercheurs eux-mêmes dans la conduite d'une recherche. Cela peut paraître assez banal mais c'est assez nouveau. Il a aussi cette fameuse interdisciplinarité ou pluridisciplinarité (parfois on a tendance à confondre ces deux concepts). Ce n'est pas récent. Je crois qu'on voit apparaître dans la recherche une émergence depuis quelques années d'objets de recherche qui ne sont pas l'avancement du front de la connaissance dans un domaine disciplinaire particulier mais la prise en charge d'un

enjeu qui dépasse un domaine particulier. Cela peut générer de l'interdisciplinarité mais cela génère surtout d'attaquer cet enjeu qui peut être un enjeu sociétal sous différents angles à travers différentes compétences disciplinaires.

Jacques Dubucs

J'avais deux remarques relatives à l'exposé d'Élisabeth Grosdhomme. La première est plutôt destinée, encore qu'elle n'est visiblement pas besoin du tout de ma défense pour illustrer son point, à la question de la relativité des normes. Cela peut concerner le périmètre de l'action de Daniel Andler. L'impact des technologies de quelque type qu'elles soient sur la société est une question exagérément abstraite. Quand on parle d'impact, l'impact est autant dû à la nature de la chose impactée que la nature de la chose impactante. Il y a un paramètre culturel déterminant qui devrait être pris en compte. Juste pour dire un mot du genre de recherche qui pourrait soutenir cette réflexion : on s'efforce y compris par des infrastructures sérieuses d'arriver à objectiver le subjectif. En même temps il est important de prendre en compte cette variante culturelle y compris pour ne pas encourir le risque de catastrophes financières ou industrielles quand les produits inappropriés restent sur étagère.

Le second point est un point de désaccord par rapport à ce qu'elle a dit. Je vais donc être plus explicite. Au fond ce qui a été dit c'est que dans les voies de résolution des grands problèmes pour dire les choses à grandes pelletées (avec la sévérité qu'impose l'heure qui passe) on a l'espace suivant (je mentionne que le recours à la dictature est ce qui est de plus efficace pour changer les comportements même si elle a évidemment un certain nombre d'inconvénients) : on a le recours à la loi en effet, c'est-à-dire la tentative de rendre délictuels les comportements non souhaitables. Il y a une défiance vis-à-vis des institutions qui est telle en effet que ceci n'est plus possible. Un autre recours qu'on a par exemple, c'est comme la confiance dans la connaissance, c'est-à-dire l'optimisme épistémique, l'idée que le stock d'informations détenues et le niveau scientifique (à défaut du niveau moral) des citoyens allait augmenter et que les citoyens allaient comprendre que les comportements appropriés était tel et tel. Ce qui, me semble t-il, est en train d'être réalisé (je ne vais pas broder sur le rôle des réseaux dans la prédominance du crétin ou dans la décélération de la croissance de la connaissance). Ce qui est en train d'être réalisé c'est une défiance vis-à-vis des institutions et des discours tenus par les scientifiques. Cela tient à plusieurs choses, notamment au fait que l'autorité scientifique a perdu le privilège de la matérialité de l'imprimé.

Dans un certain nombre de domaines critiques, comme la sécurité des citoyens ou le changement climatique, nous n'avons plus le loisir d'attendre que le niveau général de connaissances augmente et que le cerveau lent domine. En d'autres termes, on arrive là à quelque chose qui est adressé au cerveau rapide, ce qu'on appelle le nudge. J'ai un léger désaccord avec Élisabeth Grosdhomme qui a semblé dire que le recours au nudge était lié aux automates, aux machines et à ceux qui les concevaient. Je ne crois pas du tout que tel soit le cas. Pour prendre un exemple relatif à la transition climatique, au problème énergétique ou au problème hydrique, on sait qu'on a réussi à faire baisser la consommation d'eau en Californie pendant les grandes sécheresses simplement en prenant comme mesure dans les districts de rendre publique la consommation de chacun dans le voisinage. Cela a eu une grande efficacité qui ne repose pas forcément sur la connaissance des nappes phréatiques aux États-Unis. Donc la question du recours à des incitations peut se présenter dans le futur du programme qui est inauguré aujourd'hui.

Réponse d'Élisabeth Grosdhomme

Je me contente, en guise de réponse, de broder sur ce que vous venez de dire. Si je reprends vos différentes pistes : la dictature, je ne l'ai pas mentionnée. Mais gardons en tête que des études de la Fondation pour l'innovation politique relatives à l'adhésion à l'idéal démocratique montrent une chute vertigineuse. Une question a été posée dans une vingtaine de pays, pas seulement en France. La question était : est-ce que cela vous paraît important de vivre dans une démocratie ? Les résultats sont décomposés selon l'âge du répondant. Les personnes de 60 ans et plus, ceux qui ont connu la Seconde Guerre mondiale ou du moins l'après-guerre, le rideau de fer, etc, répondent à 80% : oui c'est très important. Mais pour le groupe d'âge des moins de vingt-cinq ans, il n'y en a plus que 20% qui disent que oui c'est important ; les autres pensent qu'après tout un gouvernement militaire ou un gouvernement par les algorithmes serait tout aussi bien. On n'en est pas encore, Dieu merci, au point où la population appelle de ses vœux un gouvernement dictatorial mais les anticorps contre un régime moins démocratique que celui dont nous avons pu bénéficier semblent s'affaiblir.

Sur la régulation des comportements par la loi, ce n'est pas parce que la loi a perdu de sa légitimité et de sa force d'application qu'il faut baisser les bras. En revanche, ce qui est frappant aujourd'hui c'est l'apparition de nouvelles déviances, de nouvelles dissidences. Je prends un exemple, qui est la manière dont on peut faire obstacle à la reconnaissance faciale. Vous avez toutes sortes de créateurs, essentiellement des créateurs de maquillage, de mode ou de joaillerie, qui vous proposent des astuces pour y échapper : par exemple des maquillages qui dissimulent l'écartement de vos pommettes, ou bien un T-shirt sur lequel est imprimé un paysage, si bien que l'intelligence artificielle qui fait de la reconnaissance de formes est perturbée car elle voit une tête mais pas le corps en dessous. Je vous donne un autre exemple : quand vous remplissez des formulaires sur internet pour vous inscrire à ceci ou cela, on vous demande quel est votre pays de résidence. Une société, dont je tairai le nom, m'a confié avoir une proportion incroyable de gens qui déclarent habiter en Afghanistan : ils n'habitent évidemment pas en Afghanistan mais quand ils déroulent le formulaire, ils cliquent sur le premier pays de la liste qui est l'Afghanistan. Ils ne veulent pas transmettre leurs données et répondent n'importe quoi. Ce sont des perturbations mineures du système ; reste cependant qu'avec la surveillance technologique on voit très nettement apparaître des phénomènes de nouvelles déviances.

Dans la gouvernance du futur, il y a un thème qui s'appelle la « dataxation ». Jusqu'à présent, vous avez différents moyens de matérialiser votre appartenance à la communauté citoyenne, dont payer vos impôts. Demain un autre mode participation sera de transmettre vos données, d'où cette notion de « dataxation ». Eh bien, il y aura demain de l'évasion de données comme aujourd'hui de l'évasion fiscale.

Pour revenir sur votre point sur la loi, nous sommes d'accord sur le constat que la loi est un outil moins puissant qu'il n'a été. Mais je recommande de ne pas considérer ce constat comme le dernier mot de l'histoire, et de se battre pour réhabiliter la loi. Quant au nudge, il est souvent considéré comme le pis-aller. On aimerait avoir des citoyens qui font preuve d'une éthique parfaite, ce qui nous dispenserait d'avoir recours à d'autres moyens pour les policer. Mais en attendant, le nudge est mieux que l'inaction. Les entreprises ne sont pas les seules à le développer ; les administrations le font aussi – avec notamment des travaux très intéressants à ce sujet d'un centre d'innovation publique britannique nommé Nesta.

Sonia Desmoulin-Canselier

Je voulais aller dans votre sens et dire que ce n'est pas parce que l'on constate actuellement un déclin, quoique relatif (car il y a aussi une véritable appétence pour la loi et une demande de nouvelles lois: c'est donc à la fois un surinvestissement et un désinvestissement un peu paradoxal), du droit que celui-ci n'est plus une ressource pertinente. Pour l'instant dans un contexte d'institutions étatiques et démocratiques, la loi (dans son sens générique) apparaît tout de même comme le moins mauvais système normatif. C'est un système procéduralement organisé, permettant du débat public et qui a l'avantage par rapport aux différentes formes de chartes et de guides qui ont été évoqués plus tôt, de présenter un certain nombre de garanties en terme de représentation des différents intérêts en présence et de représentativité des personnes qui prennent la décision.

Bien sûr tous ces éléments sont contestés, mais si on passait lesdites chartes au crible de critères équivalents à ceux qui ont été élaborés pour les normes juridiques, on verrait que les chartes résistent beaucoup moins bien qu'un certain nombre de lois de ce point de vue là. De plus, il ne faut pas penser que la normativité "souple" des acteurs économiques est une alternative totale à l'ordre juridique (j'ajoute que la loi, ce n'est pas juste la loi pénale). Même la normalisation technique évoquée plus tôt s'adosse à tout un système réglementaire. Le pluralisme normatif existe déjà. Le problème est plutôt : qu'attend-t-on de la loi ? Comment fait-on pour qu'elle soit appliquée ? Mais je vous rejoins totalement sur l'idée qu'il y a un important travail à faire sur l'après de la décision politique et législative, et plus généralement de la décision normative. D'ailleurs, les chartes et autres normes volontaires reviennent dans le système judiciaire, par exemple au moment où l'on s'aperçoit qu'un comportement est problématique et qu'il a causé un dommage. Pour apprécier ce comportement, le juge saisi peut consulter lesdites chartes. Mais le problème principal de ces documents est qu'ils sont élaborés seulement par les acteurs réunis sur la base du volontariat, mais aussi de leurs réseaux et des moyens matériels et humains dont ils disposent. La question de savoir comment ces gens ont été désignés pour participer est toujours d'actualité. C'est souvent l'angle mort de la procédure. On a un effet normatif et éventuellement judiciaire de textes dont le mode d'élaboration est très souvent sujet à critique.

Réponse de Yannick Blanc

Juste un mot pour dire que dans mon propos je ne portais aucun jugement de valeur sur la validité, la supériorité ou l'infériorité de tel type de normes sur une autre. Je ne faisais qu'observer la prolifération du phénomène des chartes. J'invitais aussi à l'interrogation collective de l'observation de ce phénomène parce que cela signifie quelque chose dans le rapport que nous avons à la règle, dans le rapport que nous avons à l'élaboration et à l'application de la règle. Il y a beaucoup de choses qui sont en train de changer. Donc observer ces phénomènes de près nous permet de ne pas camper sur des positions de principe mais de mieux comprendre les questions que posent par exemple dans le champ du droit public l'opposition entre la prolifération de la production de normes et l'affaiblissement des outils de l'application de ces normes. C'est un paradoxe qui mérite d'être analysé de près.

Conclusions

PIERRE DELVOLVÉ

Pierre Delvolvé, agrégé des facultés de droit (1966), a été successivement professeur aux universités de Beyrouth, de Toulouse, de Paris II Panthéon-Assas. Il est professeur émérite depuis le 1er septembre 2010. Il est membre de l'Institut, ayant été élu à l'Académie des sciences morales et politiques en 2009 ; il en est le président pour l'année 2020.

Le programme *Technologies émergentes et sagesse collective* a été lancé par Daniel Andler dans le cadre de l'Académie des sciences morales et politiques. Le colloque le présentant a été ouvert par M. Gabriel de Broglie, membre de cette Académie, chancelier honoraire de l'Institut et président de la Fondation Del Duca qui finance ce programme. Pour clore les deux journées du colloque, il a été demandé au président de l'Académie de présenter des conclusions. C'est bien audacieux pour une personne qui n'assume qu'une présidence « filante » et qui n'a aucune compétence dans le domaine des technologies.

Peut au moins être soulignée l'exceptionnelle richesse des travaux de ces deux journées. Ils ont d'abord mis en valeur de manière générale « les enjeux sociétaux des technologies émergentes ». Puis, de celles-ci, ont été exposés les principaux aspects : le numérique, la robotique, les relations homme-machine, la génétique, les biotechnologies, l'intelligence artificielle, les nanotechnologies. Dans tous les cas, il ne s'est pas agi seulement de présenter l'état actuel des solutions et les perspectives à venir ; ont été posées les questions essentielles de ces développements, notamment en termes d'éthique, de responsabilité, de protection, de démocratie.

D'emblée apparaît l'ampleur du programme, non seulement par les objets sur lesquels il porte, mais tout autant par les problèmes fondamentaux qu'il soulève et les réponses qu'il

appelle à leur donner.

Il est dû à l'initiative de Daniel Andler, qui manifeste par là son entière « implication » dans l'activité de l'Académie et la profondeur de sa réflexion sur des sujets essentiels de la société actuelle. Il doit être remercié pour ce qu'il entreprend ainsi, dans le prolongement d'une oeuvre très riche.

Si l'on essaie de dégager des travaux de ces journées les réflexions qu'elles suscitent pour un non-spécialiste, on peut les rassembler autour de deux considérations : la puissance, la résistance.

Puissance

Le mot est revenu à plusieurs reprises dans les interventions, soit comme substantif soit comme adjectif. Il vient en résonance avec le programme de l'Académie pour cette année, qui porte sur le pouvoir. C'est bien la puissance qui ressort des travaux de ces journées, dans trois domaines.

Il s'agit d'abord de la *puissance scientifique*. Par les recherches entreprises, la science contemporaine est arrivée à des découvertes et à des résultats qui dépassent tout ce qui avait pu être trouvé auparavant. Les études antérieures ont été le support nécessaire de toutes les nouvelles avancées. Mais celles-ci sont arrivées à un degré d'invention extraordinaire et qui continuera à se développer. On n'en voit pas les limites. La recherche est assurée d'une liberté que les chercheurs veulent complète. Elle est considérée comme une valeur en soi, qu'aucune contrainte ne saurait entraver. Les moyens que les industries mettent à sa disposition en soutiennent la réalisation.

Apparaît ainsi en deuxième lieu la *puissance économique*. Celle-ci se trouve dans les entreprises qui soit poussent à des découvertes nouvelles, soit exploitent celles qui ont été ré-

alisées, soit à la fois suscitent des recherches pour bénéficier des réalisations qu'elles permettent. Le cas des GAFAM vient immédiatement comme exemple. Ils ne sont pas les seuls : on pourrait en donner beaucoup d'autres, dans l'industrie pharmaceutique ou encore dans les industries de l'informatique et de la robotique. Cette puissance peut prendre la forme de monopoles ou d'oligopoles, donnant à leurs détenteurs le pouvoir de maîtriser tout un secteur d'activité. Au cours des travaux a été donné l'exemple d'une entreprise qui est arrivée à avoir l'exclusivité de l'attribution des noms de domaine.

Elle peut arriver par là à exercer une *puissance politique*. C'est la troisième qui ressort des analyses de ces deux journées. Les nouvelles technologies sont un instrument du pouvoir. Elles peuvent être bénéfiques lorsqu'elles associent un large public à l'élaboration de nouvelles normes. L'exemple a été donné en France de la loi pour une république numérique, dont l'adoption à l'unanimité par le Parlement a été précédée d'une consultation par internet qui a donné vingt-trois mille réponses et a permis une amélioration des dispositions initialement projetées. En sens inverse, les nouvelles technologies peuvent être pour le pouvoir un instrument de contrôle des individus, dans leurs échanges, dans leurs déplacements et, plus grave encore, dans tous les aspects de leur personnalité : on en arrive à une police de l'intime.

C'est pourquoi il faut résister.

Résistance

C'est le mot qui vient à l'esprit devant le caractère terrifiant (l'expression a été employée) des inventions nouvelles. Il est évidemment fort puisqu'il a désigné les mouvements qui ont lutté contre la tyrannie hitlérienne et qu'on n'est pas ici en présence de dominations aussi dramatiques. Un des intervenants a employé une formule plus modérée lorsqu'il a dit qu'il

fallait garder la main. Mais l'idée est la même : contre les puissances, il faut s'élever.

Les pouvoirs publics eux-mêmes peuvent avoir à le faire lorsque c'est, non pas seulement l'exercice de leur pouvoir qui est compromis, mais leur légitimité même qui est remise en cause. Que peuvent faire les petits Etats contre les puissances politiques et économiques dont non seulement le budget peut être supérieur aux leurs, mais qui surtout, par les techniques, peuvent contrôler leur activité, contraindre et commander leurs actions. Que signifie la souveraineté de l'Etat lorsqu'elle se heurte, non plus seulement à la puissance d'argent, mais maintenant à la puissance de la technologie ?

Ce sont évidemment, plus encore, *les particuliers* qui doivent, sinon se rebeller, du moins être protégés. La liberté personnelle est menacée.

Elle l'est par exemple lorsque tout individu peut être identifié partout où il est et à tout moment. On peut considérer comme un progrès le message reçu par un voyageur sur ses appareils téléphoniques ou autres lorsqu'il débarque d'avion à l'étranger (Rome, Taipei, pour prendre des exemples personnels récents), lui souhaitant bonne arrivée et lui indiquant les modalités et le tarif d'un appel : c'est un service qui lui est rendu, c'est un lien qui lui est assuré. C'est tout autant d'un lien qu'il est serré. Il est certes venu à l'étranger en bénéficiant de la liberté d'aller et venir. Mais il n'est pas libre : les données de son déplacement sont enregistrées ; on sait où il est ; il n'a pas la liberté qu'on ne le sache pas. Il serait libre seulement si c'est lui-même qui demandait qu'on l'informe des conditions d'appel du pays où il arrive. Sa propre démarche provoquerait le message ; il obtiendrait par elle ce qu'il veut. Il exercerait sa liberté.

On peut encore citer le système du prélèvement à la source, tel qu'il est en vigueur en France depuis le 1er janvier 2019. C'est une

admirable réussite : le changement s'est réalisé sans accroc ; l'administration fiscale et les informaticiens ont réussi une prouesse alors qu'il fallait traiter des millions de comptes et de contribuables. Tout cela s'est fait en douceur et sans résistance. Mais quelle inversion par rapport au principe de la *contribution* fiscale ! Lorsque le contribuable recevait un avis d'imposition lui commandant de payer une certaine somme, il était certes obligé de le faire ; mais il lui restait à accomplir une démarche, celle de payer, constituant sa participation aux dépenses publiques. Désormais rien de tel : le prélèvement à la source supprime tout acte de volonté du redevable.

Tout cela peut paraître secondaire et faciliter la situation de l'usager dans le premier cas, du contribuable (et plus encore du fisc) dans le second cas. Il n'y en a pas moins une inversion de principes fondamentaux. Ces petits pas dans la restriction des libertés, sans aucune réaction, vont en permettre de plus grands.

L'un d'eux a été exposé au cours des travaux : celui de l'enregistrement des données personnelles, qui donne une prise à toute intervention. Ce fut le cas aux Pays-Bas : l'enregistrement de la religion des citoyens pour déterminer le montant de la subvention accordée aux différentes religions a grandement facilité ensuite l'identification des Juifs et leur déportation. Aujourd'hui en Estonie sont enregistrées toutes les données des citoyens permettant de leur assurer un système de soins perfectionné ; ce dispositif a été démocratiquement approuvé par référendum à une majorité de 75%. Le pouvoir démocratique peut être un pouvoir tyrannique, comme l'ont montré plusieurs auteurs (Tocqueville, Jouvenel notamment). Il faut pouvoir lui résister autant qu'à un autre pouvoir.

Que dire alors des propositions de « fabrication » d'embryons selon des critères précis (sexe, couleur des yeux, des cheveux par ex.) pour satisfaire à une demande ? Le cas a été ex-

posé. On peut soutenir qu'il n'y a là que la mise en œuvre de la liberté de ceux qui demandent et de ceux qui répondent. Ces derniers cherchent d'ailleurs à s'abriter derrière une clause d'irresponsabilité si le résultat n'est pas atteint – ce qui est d'ailleurs juridiquement contestable. Devant cette dénaturation de la procréation, il y a encore matière à résistance.

Les pouvoirs publics en sont conscients, qui, dans l'ordre international comme dans l'ordre interne, cherchent à formuler des normes (par exemple Convention d'Oviedo du 4 avril 1997 sur les droits de l'homme et la biomédecine, Règlement général du 27 avril 2016 de l'Union européenne sur la protection des données ; lois françaises sur l'informatique et les libertés, et sur la bioéthique), et mettent en place des organes de contrôle (par exemple en France la Commission nationale de l'informatique et des libertés).

Mais ces normes et ces organes ne sont jamais suffisants : il faut les adapter à l'évolution des technologies. On a même dit qu'il faut les adapter à l'évolution de l'éthique, celle-ci étant contingente et dépendant des options politiques. Cette affirmation est un renversement de l'ordre des notions : s'il y a une éthique, qui définit les normes fondamentales des comportements humains, c'est elle qui doit commander les dispositifs politiques, et non l'inverse.

La recherche des normes éthiques peut être difficile ; la détermination de leur contenu peut ne pas être unanime.

Mais il existe au moins deux principes qui les commandent : celui de l'intégrité de la science, celui de la dignité de la personne humaine¹.

Ils sont au cœur de la sagesse collective qui doit encadrer les technologies émergentes. C'est une tâche qui relève des sciences morales et politiques.

¹ Pierre Delvolvé, « L'encadrement normatif de la science », *Revue française de droit administratif*, 2018, p. 487 - 493

Les Cahiers de TESaCo n°1

Ce premier Cahier rassemble l'essentiel des communications du colloque de lancement d'une enquête « Technologies émergentes et sagesse collective » (TESaCo), qui a eu lieu à Paris, à l'Institut de France le 30 et 31 janvier 2020 sous le titre : « Technologies émergentes : comprendre, faire comprendre, maîtriser. Les conditions d'une sagesse collective ».

L'objet de l'enquête est constitué par l'ensemble des technologies dites « émergentes ». Il s'agit des technologies de l'information et de la communication (TIC), de l'intelligence artificielle et de la robotique ; des biotechnologies ; des technologies dérivées des sciences cognitives et des neurosciences ; des nanosciences. Ces domaines en plein essor appellent une adaptation en continu de la part tant des usagers que des experts et des décideurs, qu'il s'agisse de l'adoption de nouveaux outils, de l'évaluation de leur fiabilité, de la prévision de leurs effets sur le long terme, des conditions économiques, industrielles et politiques de leur développement. Cet état de fait appelle un effort pour mieux comprendre les technologies et leurs effets, informer le public et les responsables politiques, et proposer des dispositifs pouvant contribuer à maîtriser l'évolution en cours.