

Ariel Kyrrou

# Les imaginaires de l'intelligence artificielle



- > #Numéro 2
- > Ce que le calcul fait à nos pratiques
- > Working papers
- > Humanités numériques plurielles

## Références de citation

Kyrrou, Ariel. "Les imaginaires de l'intelligence artificielle.", 28 novembre 2018, mäj 0000, *Cahiers COSTECH*<http://www.costech.utc.fr/CahiersCOSTECH/spip.php?article68>

## Résumé

Selon le rapport de Cédric Villani, « Donner un sens à l'intelligence artificielle », qui a fait date au printemps 2018, les recherches en IA sont indissociables des craintes et fantasmes que porte la discipline, en particulier *via* la littérature et les grands films de science-fiction. Ce constat s'applique parfaitement aux géants du numérique californiens. Mais les œuvres d'anticipation n'en restent pas moins des antidotes à leurs délires les plus démiurgiques.

## Auteur(s)



Ariel Kyrrou est un journaliste, écrivain, essayiste et animateur radio, spécialisé dans les nouvelles technologies, les musiques électroniques, la science-fiction et les grandes avant-gardes artistiques du siècle dernier. Il est chercheur associé au laboratoire Costech (UTC).

Et si les recherches en intelligence artificielle étaient inséparables de l'imaginaire de nos machines « intelligentes » et de nos créatures « fabriquées » ? Et si le devenir de cette discipline et de cette technologie, dont les applications se développent comme jamais au cœur du quotidien de chacun d'entre nous, ne pouvait se concevoir sans référence aux rêves et aux cauchemars qu'elles suscitent ? Et s'il fallait convoquer Frankenstein, Terminator, HAL 9000, R2D2, la belle androïde de *Ex Machina* et leurs multiples incarnations dans le cinéma, la littérature et le jeu vidéo pour donner du sens au futur de l'IA ? C'est ce que sous-entend le rapport rendu fin mars 2018 par le mathématicien et député de l'Essonne Cédric Villani : « Donner un sens à l'intelligence artificielle », pourtant sous-titré de façon plus opérationnelle « Pour une stratégie nationale et européenne ». En effet, pour un gouvernement revendiquant la « tech for good », il s'agit de rendre à la fois crédible et désirable l'Eldorado du numérique. Or si la crédibilité suppose une rationalité, voire une prise de distance vis-à-vis de certaines lubies des géants du numérique états-uniens, susciter le désir du grand public impose une mise en récit accessible et séduisante de l'IA. Autrement dit : un *storytelling* nourri de ces imaginaires fictionnels.

Dès la première page de l'introduction du rapport, page 9 pour être précis, le robot dopé à l'IA, le scientifique, le citoyen, l'écrivain et le metteur en scène de science-fiction sont conviés sur la même planète bien au-delà du numérique : « *Du fait de ses ambitions, qui en font un des programmes scientifiques les plus fascinants de notre époque, la discipline de l'IA s'est toujours développée de concert avec les imaginaires les plus délirants, les plus angoissants et les plus fantastiques, qui ont façonné les rapports qu'entretient le grand public avec l'IA mais également ceux des chercheurs eux-mêmes avec leur propre discipline. La (science) fiction, les fantasmes et les projections collectives ont accompagné l'essor de l'intelligence artificielle et guident parfois ses objectifs de long terme : en témoignent les productions fictionnelles abondantes sur le sujet, de 2001 l'Odysée de l'espace, à Her en passant par Blade Runner et une grande partie de la littérature de science-fiction. Finalement, c'est probablement cette alliance entre des projections fictionnelles et la recherche scientifique qui constitue l'essence de ce qu'on appelle l'IA.* »

**De HAL 9000 de 2001 *L'Odysée de l'espace* à l'OS1 de *Her*, un**

## imaginaire qui prend corps dans notre XXI<sup>e</sup> siècle

La notion d'imaginaire, selon Wikipédia, se définirait comme « *le fruit de l'imagination d'un individu, d'un groupe ou d'une société, produisant des images, des représentations, des récits ou des mythes plus ou moins détachés de ce qu'il est d'usage de définir comme la réalité.* » Sous le double regard de ce résumé et du parti-pris de Cédric Villani, l'imaginaire de l'intelligence artificielle serait de l'ordre des « fantasmes et projections collectives ». Il prendrait la forme de mythes actualisés par le présent, d'images et de représentations partagées avec d'autres personnes de nos groupes, de nos sociétés, de nos environnements proches ou lointains, *via* des fictions qui, au-delà de leur fonction divertissante, agiraient profondément sur notre vision du monde. Sur notre appréhension du progrès et de ses pièges. Sur nos espoirs et nos désespoirs vis-à-vis de la chose technologique.

Les trois films cités par ce rapport – dont l'écho remarquable n'a d'égal que le risque qu'il finisse dans un tiroir de l'administration de l'État – méritent un arrêt images.

Commençons par le premier long métrage mentionné. Sorti en 1968, détail tout sauf neutre, *2001 L'Odyssée de l'Espace* concrétise la vision d'un ordinateur *vraiment* intelligent. HAL 9000 n'a pas de corps humanoïde. Il s'incarne dans le vaisseau et se manifeste par son œil qui semble tout voir et savoir de son horizon proche. Il est le pilote autonome au service de l'humanité dans sa conquête des galaxies, symbole métaphorique autant qu'auxiliaire que l'on penserait infaillible de notre quête de connaissances. Grâce à son scénariste majeur, l'écrivain de science-fiction et prospectiviste Arthur C. Clarke, ainsi qu'une ribambelle de conseillers scientifiques, allant de l'astronautique à la paléontologie en passant par les bases lunaires, la recherche informatique ou l'hypothèse d'une intelligence extraterrestre, le film de Stanley Kubrick rend plausible le syndrome Frankenstein d'un bug absolu de la machine intelligente. De fait, il crédibilise, et notre rapport démiurgique aux nouvelles technologies de l'intelligence et de la vie, et le danger d'un dépassement de l'humain, ou du moins d'une erreur imprévue de la mécanique que l'on croyait parfaite. Or la perspective d'une rébellion a priori « rationnelle » de machines potentiellement « plus intelligentes » que nous vis-à-vis de leurs créateurs humains, qu'elle soit acceptée ou rejetée, est aujourd'hui au cœur des imaginaires de l'IA.

Les répliquants de *Blade Runner* (1982) ajoutent d'autres dimensions au glacial intellect de l'ordinateur de *2001 L'Odyssée de l'Espace* : la ressemblance physique entre le créateur humain et le robot doté d'une IA remarquable par la grâce de la Tyrell Corporation ; et puis les rôles cruciaux, d'une part de la mémoire, qu'elle soit fabriquée *a priori* ou construite par des interactions, d'autre part de l'empathie, c'est-à-dire de la capacité de la mécanique à se glisser dans notre peau et notre esprit, pour mieux nous comprendre, nous servir et pourquoi pas déjouer ses propres programmes. Nécessité ou non d'une proximité esthétique entre la machine et l'humain ? Mémoire ? Empathie ? Autant de questions déterminantes de la robotique de 2018 – que l'on retrouve d'ailleurs dans un autre film, le bien nommé *AI* de Steven Spielberg (2001), dont on peut s'étonner qu'il ait été oublié par Villani.

Reste une troisième référence essentielle : *Her* de Spike Jonze (2013). Cette fois – et il s'agit d'un changement déterminant – l'IA perd la grandiloquence à la Big Brother de HAL. Elle ne semble plus du tout motivée par la quête d'humanité des androïdes Rachel ou Roy Batty de *Blade Runner*. Mieux : elle assume son identité virtuelle, avec cette capacité inhumaine de se connecter au même moment, en un même élan invisible, à une myriade d'êtres humains voire d'IA comme elle. Elle n'est qu'un système d'exploitation, OS1, auquel le héros donne la voix d'une femme séduisante. Et ce système d'exploitation, pour satisfaire cet homme qui a personnalisé sa programmation finale, apprend par lui-même, évolue au fur et à mesure des interactions avec lui. Ce simple OS agit ainsi à la façon de nos IA de 2018 et plus encore de demain, au cœur de relations banales, au quotidien. Sauf qu'au final, l'écrivain public joué par Joaquin Phoenix, désabusé de ses amours avec l'autre sexe, tombe amoureux de cette mécanique sophistiquée, pourtant désincarnée. Quoi qu'il en soit, dans *Her*, l'IA redescend sur Terre, et en acquiert une présence familière, immédiate et permanente, comme ses sœurs de notre présent qui vivent « à l'insu de notre plein gré » dans nos connexions et discussions en ligne avec nos banques, nos assurances, nos services juridiques, nos livreurs, nos centres d'appels, nos antennes médicales ou nos *coachs* numériques en tous genres.

Cette intelligence artificielle-là, si banale, serviable mais insidieuse, Cédric Villani et ses collaborateurs la font vivre au travers de nouvelles de science-fiction qui émaillent leur rapport, sur fond bleu pour rythmer le passage d'un chapitre à l'autre. Comme pour assumer ses affirmations sur l'importance des imaginaires en matière d'intelligence artificielle, un

texte signé Anne-Caroline Paucot met en scène un couple heureux qui en vient à se déchirer à cause d'une petite IA de l'ère de la prévoyance intégrale. Disposant en effet de leurs datas à tous deux, « *Madame Pizza3D fait dégringoler la cote de bonheur* » de Kim. Une pizza mexicaine pour la femme de Timéo ? « *Non, ce n'est pas possible. Madame Pizza3D dit qu'avec tes hémorroïdes, le piment n'est pas conseillé.* » Pire, comme le dit le jeune homme, dépité, à sa compagne : « *Kim, si tu maintiens ta Mexicaine, tu auras un malus d'assurance de 30%.* » Vous mangez trop gras ? Vous n'accomplissez pas vos 10 000 pas mesurés chaque jour grâce aux miracles de votre montre connectée ? La vendeuse robotique dopée à l'IA vous apostrophe, car tout se sait à l'ère du tout numérique et des Big data globalisées. Et si vous voulez protester contre ces inqualifiables ingérences dans votre vie privée ? « *Ils sont équipés d'un système anti-insultes avec dépôt de plainte immédiat pour outrage à commerçant dans l'exercice de sa prévention*<sup>2</sup>. » La nouvelle d'après cette « Perte de commande » s'appelle « Impulsion naturelle » et déroule son scénario sur cinq pages entre l'utopie et la dystopie, sous la plume d'Olivier Paquet. Dès les premières lignes apparaît la figure à poil et ronronnante d'un chat artificiel capable d'auto-apprentissage, « *coffre-fort portable* » des données personnelles de l'héroïne, à même de se connecter à ses frères et sœurs félins mécaniques, mais aussi de lancer « *ses requêtes en utilisant les implants réseau* » installés sur les arbres de l'île d'Oléron. La nouvelle mentionne des applications, à l'instar des essaims de drones « *fonctionnant sur une base coopérative avec des IA basiques* », notamment utiles « *pour la pollinisation* ». Elle dessine aussi les contours d'un « *SuperBrain Project* », travaillant à « *l'élaboration d'une IA qu'on laissait se développer sans l'orienter, comme un enfant que ses parents surveilleraient de loin.* » Sauf qu'ici, les IA connectées s'avèrent diablement malines et adaptées à des fins écologiques, prenant en compte « *l'écosystème, depuis la forêt jusqu'à la mer* », en « *osmose avec les éléments naturels* », là où les personnages de la nouvelle décrivent le SuperBrain Project comme une initiative « *tout à fait inutile, un jouet pour des chercheurs*<sup>3</sup>. » Sous-entendu : c'est à nous, en Europe, de construire et de valoriser un imaginaire de l'intelligence artificielle toute de vies connectées en phase avec nos ambitions environnementales, afin de mieux contrer le rêve – absurde ? – d'une IA super intelligente risquant de nous « dépasser », telle qu'en fantasment pas mal des apôtres de l'idéologie transhumaniste, si puissante dans la Silicon Valley<sup>4</sup>...

## Le cas exemplaire de l'imaginaire de science-fiction de Google

Le rapport Villani n'entre jamais dans le détail de l'activation des imaginaires de l'IA, ce qui est bien dommage. Qu'en est-il vraiment dans la Silicon Valley ? Les images partagées, les représentations collectives y sont-elles assumées, comme de réels moteurs de l'innovation ? Et si oui, avec quels effets indirects sur notre quotidien ? Le détour par Google, dont la maison-mère est devenue l'entreprise Alphabet en octobre 2015, s'impose.

Andy Rubin a fondé la startup *Android*, rachetée le 17 août 2005 par Google. Puis à partir de fin 2007, c'est sous son imprimatur que le système d'exploitation en *open source* Android en est arrivé à équiper plus de 80% des smartphones LG, HTC, Samsung et autres fabricants sur la planète, loin devant l'iOS d'Apple. Rubin a ensuite pris la direction des huit sociétés de robotique acquises par l'ogre californien fin 2013, histoire de cultiver sa passion pour les robots. Mais cette aventure a tourné court. Et depuis qu'il a lâché Google pour monter sa société de capital-risque, Playground, puis sa marque d'élégants smartphones, Essential, il se lâche sans crainte d'un retour de bâton des héritiers de l'auteur de science-fiction Philip K. Dick. Le 7 janvier 2016 à 19 heures en Californie, soit 5 heures du matin le 8 janvier en France, Andy Rubin a en effet envoyé un tweet pour fêter la « *naissance / activation* » de Roy Batty, l'androïde Nexus 6 joué par Rutger Hauer dans *Blade Runner* de Ridley Scott, long métrage tiré du roman de Philip K. Dick *Les androïdes rêvent-ils de moutons électriques ?*. Sur son message figure une photo visiblement tirée du film, la tête du répliquant positionnée à deux centimètres de sa fiche d'identité sur un écran : « *Replicant (M), Des : BATTY (Roy), Nexus 6, N6MAA10816, Incept Date : 8 JAN., 2016, Func : Combat, Colonization Defense Prog. Phys : LEV. A. Ment : LEV. A* ». Or la gamme de smartphones et tablettes tactiles sous la marque autonome du géant de Mountain View, créée en son temps par Andy Rubin pour mieux montrer à tous les potentialités de l'OS Android, s'appelle Nexus. C'est comme si les Nexus de Google, repoussant toujours « *plus loin les limites de ce qu'il est possible de faire avec un mobile* », permettant de tout commander de sa vie en *live*, au seul son de la voix *via* les merveilles de son IA, représentaient les premiers pas vers la fabrication des androïdes Nexus 6 de la Tyrell Corporation du film, tels la sublime Rachel ou le terrifiant Roy Batty<sup>5</sup>.

Cette anecdote dévoile une réalité d'habitude cachée, car propre à la

culture, voire à l'histoire intime des acteurs concernés : ici, un imaginaire de science-fiction habite la subjectivité d'une entreprise à l'ambition démiurgique. Mais cet univers partagé se situe très en amont de l'idéologie transhumaniste, comme l'une de ses multiples sources peut-être, à la fois profonde et terriblement ambiguë. Que *Blade Runner* soit aussi et surtout une critique du manque d'humanité de nos sociétés, le chasseur Deckard se révélant moins « humain » que certains Nexus 6, ne change rien à la donne. Au risque de glisser vers l'idéologie, dérive hélas trop commune, ce type d'imaginaire est un formidable moteur de création. S'appuyant sur des technologies se voulant transparentes d'évidence ainsi que sur leurs usages surmultipliés, Google s'est porté vers les cimes grâce à cette puissance-là. Il se veut le net à lui seul. Relais de toute l'information de l'univers, il agit pour le bien – à l'instar du motto « *Don't be evil* » et du « *Do the right thing* » qui l'a remplacé avec Alphabet. Il est la vie. Il est l'information. Il est HAL 9000, imaginé en 1968 par Stanley Kubrick avec le gourou de l'intelligence artificielle Marvin Minsky, lui-même fondu de science-fiction, ainsi comme déjà dit d'Arthur C. Clarke, écrivain, prospectiviste et principal scénariste de *2001 L'Odyssée de l'Espace*.

La perspective que l'on pensait impossible d'une IA s'avérant « naturelle », voire plus humaine que l'être humain, hante les cultures, occidentale comme nipponne – même si de façon différente. Des *Simpsons* à *Star Wars* en passant par Achille Talon, la « *liste des références culturelles à 2001 L'Odyssée de l'espace* » sur Wikipédia est dantesque. Ne nous trompons pas, c'est la silhouette de HAL 9000 – CARL dans sa version française – qui se profile derrière le programme AlphaGo de la société DeepMind, intégrée depuis son rachat en janvier 2014 au labo Google X. Car après avoir battu le numéro 3 mondial du jeu de go en mars 2016, Lee Sedol, elle semble désormais capable de jouer de façon « *surhumaine* » autant au go qu'aux échecs... Et ses concepteurs ont promis au printemps 2018 qu'elle battrait bientôt les caïds du jeu vidéo *Starcraft6*. L'horizon affiché de DeepMind ? Qu'AlphaGo devienne une IA « généraliste », exerçant ses réseaux de neurones sur une multitude de domaines avec un minimum de connaissances *a priori*. Bref, que cette IA devienne capable de résoudre tous types de problèmes tarabiscotés – et pas simplement de vaincre les champions d'un jeu unique, aussi complexe soit-il à l'instar du go qui offre plus de possibilités de parties qu'il y a d'atomes dans l'univers.

Autre perspective fantasmagique, qu'incarment les grandes IA de la

science-fiction : celle de l'autonomie, passant forcément par la capacité de la machine à apprendre « par elle-même ». L'introduction d'un article sur l'intelligence artificielle conçue par Google Brain (autre équipe du géant de Mountain View) illustre parfaitement ce vertige : « Cette IA qui peut générer d'autres IA a effectivement créé un véritable "enfant", qui a surpassé tous ses homologues conçus par l'Homme<sup>6</sup>. » Vision superlative, qui fait écho au SuperBrain Project de la deuxième nouvelle du rapport Villani, dont le modèle démiurgique semble avoir été l'intelligence artificielle rêvée par Google Brain. Mais à en croire les mots bien pesés de Jean-Gabriel Ganascia dans son livre sur *Le mythe de la Singularité*, cette métaphore de la naissance tromperait son monde. L'autonomie des IA de DeepMind ou Google Brain, magnifiées, est obtenue le plus souvent grâce à un apprentissage par renforcement (avec récompenses). Au sens technique, elle « signifie qu'il existe une chaîne de causalités matérielles allant de la prise d'information par des capteurs, à la décision, puis à l'action, qui ne fait pas intervenir d'agent extérieur, en particulier d'agent humain. Mais d'un autre côté, au plan philosophique, l'autonomie tient à la capacité à se donner sa propre loi, à savoir les règles et les finalités de son comportement. » Or, écrit Jean-Gabriel Ganascia, l'humain « qui configure l'algorithme d'apprentissage par renforcement choisit lui-même le critère à optimiser, sans que la machine soit en mesure de le changer.<sup>7</sup> » L'avancée bien réelle vers de plus en plus d'autonomie, que le chercheur français ne nie pas, n'en suscite pas moins une incroyable excitation des imaginaires de la Silicon Valley...

« Le moteur de recherche ultime est quelque chose d'aussi intelligent que les êtres humains, voire davantage », avançait déjà Larry Page dans une conférence du début des années 2000. Ce qui se traduit, dans une interview de Sergey Brin à *Newsweek* en mars 2004, par une variante de la même démente cybernétique : « Il est certain que si vous aviez toute l'information du monde directement connectée à votre cerveau, ou un cerveau artificiel plus intelligent que votre propre cerveau, vous vous en porteriez d'autant mieux.<sup>8</sup> » D'où les *Google Glass* ou les projets de puces et autres implants de connexion pour toujours, « version allégée de Google » selon Brin, à glisser sous notre boîte crânienne tels les « papies » et les « mamies », périphériques de connaissance ou de personnalité imaginés en 1987 par l'écrivain cyberpunk George Alec Effinger dans son roman *Gravité à la manque*<sup>9</sup>. De fait, avant de pratiquer un discours bien plus prudent, les deux fondateurs de Google ont souvent parlé de leur moteur de recherche comme s'il s'agissait d'un bébé HAL 9000. Est-ce la raison pour laquelle ils ont embauché Ray Kurzweil en 2012 ? Celui-ci, en quelque sorte pape d'une idéologie transhumaniste

aux multiples facettes, semble encore très loin d'avoir inventé puis élevé l'IA de leurs rêves. Sauf que l'impossible d'hier semble devenir envisageable, ne serait-ce qu'à très long terme. Et c'est bien pourquoi l'apport de la science-fiction, pour imaginer l'inimaginable, semble plus crucial que jamais. Le sérieux, mais aussi le plaisir voudraient que cette « intelligence » des machines ne puisse être du même ordre que la nôtre. Sera-t-elle une « intelligence *alien* », une « *seconde nature* » algorithmique, explorant l'incalculable *via* l'intégralement calculé, comme la dépeint la chercheuse Luciana Parisi dans un curieux article de la revue *Multitudes*, titré « *La raison instrumentale, le capitalisme algorithmique et l'incomputable*<sup>10</sup> » ? À moins que l'activation des imaginaires de l'IA par les autres cadors de la Silicon Valley n'accouche d'une tout autre souris cybernétique ?

## L'iminaire IA de la Silicon Valley

C'est en juin 2016, lors de la « Code conférence » des magazines en ligne californiens *The Verge* et *Recode*<sup>11</sup>, qu'Elon Musk, fantasque PDG de la société d'aéronautique et lanceurs spatiaux SpaceX, ainsi que du fabricant de voitures électriques haut de gamme, *high tech* et sportives Tesla Motors, a révélé au monde son projet Neuralink. L'objectif de cette *startup* est d'augmenter nos capacités cérébrales grâce à un « *lacet neuronal* ». Ce terme, à la fois imagé et nébuleux, induit l'idée d'un entrelacement de ce qui est décrit comme une « *couche d'intelligence digitale* » avec nos presque cent milliards de neurones, à l'intérieur même de notre boîte crânienne. Malgré la perspective annoncée d'une commercialisation de premiers « produits » en 2021, Elon Musk n'explique jamais le comment d'une telle opération. Mais il justifie la création de Neuralink par la conviction que « l'intelligence » des machines est appelée à surpasser la nôtre d'ici une trentaine d'années : « *Nous serons à ce point surclassés, au niveau de l'intelligence, que nous nous retrouverons face à l'IA comme nos animaux de compagnie face à nous*<sup>12</sup> ». Autrement dit : serviles comme un chien fidèle, ou un chouïa rebelles à la façon d'un chat, nous vivons sous la domination de nos IA super intelligentes. La perspective imaginaire est ici celle d'un monde à la *Matrix*. Musk, sous ce regard, serait une sorte de Néo d'avant le personnage de Keanu Reeves. Comme s'il alertait le monde avant que n'advienne l'ère de la Matrice. Avec une solution simpliste, qui n'est autre que celle du transhumanisme : pour éviter que nous ne devenions demain inutiles, que nous tombions sous le contrôle intégral de nos machines devenues autonomes, opérons la fusion ou du moins le rapprochement le

plus intime entre l'intelligence des machines et la nôtre – sans besoin d'ailleurs de préciser la définition de l'intelligence, réduite à sa version la plus opérationnelle et compétitive, de l'ordre du QI...

Car, argumente le patron de SpaceX et de Tesla Motors, « *nous sommes déjà des cyborgs. Vous avez déjà, en ligne, une version digitale de vous-même, du moins partielle (...). Grâce à votre ordinateur, votre smartphone, vos applications, etc., vous avez déjà des super pouvoirs, bien plus que le Président des États-Unis d'il y a une vingtaine d'années !* ». Nous voilà dans l'imaginaire, assez primaire, de *Robocop* et des *Quatre Fantastiques* des Marvel Comics. L'objectif, sous ce regard si plein d'amalgames, est de pousser un cran plus loin la route vers toujours plus de digitalisation de nos êtres, afin d'augmenter sans cesse ces soi-disant « super pouvoirs » technologiques. Or, continue-t-il en identifiant l'humain à une boîte noire cybernétique, nos limites par rapport aux machines tiennent à la faible performance des « entrées et sorties » (« input » et « output ») de notre cerveau. « *Surtout les sorties, extrêmement lentes* », à cause du temps infini que nous passons à transmettre nos messages, par exemple « *avec nos doigts sur un clavier* », alors même qu'à l'inverse, côté entrées (input donc), « *nos yeux prennent en compte un nombre beaucoup plus important de données en un instant.* » Que l'ouïe, l'odorat, le goût ou le toucher soient des modes de perception du monde tout aussi importants que la vision ? Que ce que nous voyons ou ressentons de notre environnement ne soit pas réductible à des données ? Bref, qu'il y ait du connu et de l'inconnu, du descriptible et de l'indescriptible, du rationnel et de l'irrationnel, du conscient et de l'inconscient, du réel et du rêve dans ce que nous percevons de notre quotidien ? Musk n'en a cure, et reste cantonné au « solutionnisme technologique » de la Silicon Valley<sup>13</sup>. Ainsi se positionne-t-il en grand super héros des petits super héros que nous sommes, agitant le chiffon rouge d'une intelligence artificielle qui prendrait demain le contrôle de la civilisation et la détruirait dans le plus pur style Terminator : « *Je n'arrête pas de tirer la sonnette d'alarme, mais tant que les gens ne verront pas des robots descendre dans la rue pour tuer tout le monde, ils ne sauront pas comment réagir.* »

« *Imaginer un scénario apocalyptique, c'est irresponsable* », lui répond à l'été 2017 Mark Zuckerberg. Ce à quoi l'ami Musk rétorque, sur Twitter : « *J'en ai parlé avec Mark. Sa compréhension du sujet est limitée*<sup>14</sup>. » Vraie ou fausse polémique ? L'imaginaire du patron de Facebook est-il tout autre ? Pas sûr. Car trois mois avant cette passe d'arme, le PDG du premier réseau social de la planète avait annoncé une initiative proche du

« lacet neuronal » dont rêve son *alter ego* de Tesla et de SpaceX. C'est en effet le 18 avril 2017 que Mark Zuckerberg a présenté son projet d'interface directe du cerveau avec Facebook, histoire de pouvoir écrire et commander tout demain depuis son smartphone *via* sa pensée plutôt qu'en utilisant ses doigts d'homme de Cro-Magnon<sup>15</sup>. Les deux projets d'interface de Zuckerberg et de Musk convergent, entre eux, mais aussi avec les visions d'un Ray Kurzweil – qui imagine un cerveau augmenté *via* ses connexions directes avec les IA de serveurs de quelque *cloud*, pour ne citer que l'une de ses idées les moins farfelues. Le Monde, dans son dossier dédié à l'intelligence artificielle du passage de 2017 à 2018<sup>16</sup>, voit ces personnages comme les porteurs archétypaux de trois des « *cinq façons d'envisager la pensée des machines* » : Mark Zuckerberg pour les « *techno-optimistes* », qui « *font gagner beaucoup d'argent à leur entreprise grâce à l'intelligence artificielle* » ; Elon Musk pour les « *alarmistes militants* » qui « *dénoncent les dangers* » de l'IA « *tout en popularisant les thèses les plus radicales* » ; et Ray Kurzweil pour les « *apôtres de la "superintelligence"* », qui « *croient à la thèse controversée de la "singularité", selon laquelle l'IA va surpasser l'intelligence humaine* » et « *y voient une chance* ». Si leurs postures semblent diverger, c'est qu'elles reflètent le positionnement à court terme vis-à-vis de l'intelligence artificielle, et de leurs sociétés, et d'eux-mêmes en tant que princes médiatiques. À l'instar d'un Ben Goertzel, directeur scientifique de Hanson Robotics en quête d'une superintelligence, ou d'un Bryan Johnson, dont la *startup*, Kermel, conçoit des interfaces cerveau-machine dont il espère qu'elles nous rendront presque télépathes, les figures de la Silicon Valley ne questionnent guère leur imaginaire de science-fiction en mode Marvel et Walt Disney, dans lequel nage selon eux leur public. Zuckerberg s'identifie à son réseau social, au centre de l'écosystème qu'il construit dans le monde, et à la puissance qu'il lui donne pour « faire le bien ». De SpaceX à Hyperloop (trains à ultra grande vitesse), Musk multiplie au contraire les entités, dans le transport futuriste plutôt que le service client dopé par l'IA. Sur un autre registre, il a été, dans les trois films *Iron Man*, le modèle assumé du milliardaire et playboy Tony Stark – archétype de l'homme augmenté. Et il n'aime rien tant que s'exprimer, dans la vraie vie, à la façon de son double cinématographique, en tant que sauveur de l'humanité, qu'il souhaite d'ailleurs transformer en « espèce multiplanétaire » grâce à la colonisation de Mars. Mais au-delà de leurs différences de discours ou de caractère, Musk et Zuckerberg partagent une même vision performative de l'intelligence et un *credo* dans les « super pouvoirs » qu'offrent les nouvelles technologies, à même de résoudre à elles seules tous les maux de l'humanité. Et dès lors que

les apôtres comme les fidèles s'y retrouvent d'une façon enfantine, assumée à la Musk ou induite à la Zuckerberg, leurs apparentes différences de *storytelling* se complètent plus qu'elles ne se heurtent.

Amazon et Apple investissent eux aussi fortement dans l'intelligence artificielle, entre autres pour leurs « agents conversationnels » respectifs, Alexa et Siri. Une analyse des annonces et déclarations de leurs mentors ne laisse pourtant guère transparaître cet imaginaire démiurgique de l'IA. Jeff Bezos, Tim Cook et leurs sbires seraient-ils plus prudents ? Ou plus, proche, *in fine*, de la vision imaginaire d'un Cédric Villani, telle qu'elle se concrétise dans les nouvelles sur fond bleu de son rapport de fin mars 2018 ? Car les Ray Kurzweil, les Elon Musk, les Ben Goertzel ou les Bryan Johnson adoptent une version superficielle des imaginaires de l'IA, moins inspirée des subtilités de la littérature que des clichés de films dont ils ne semblent percevoir que l'écume. Mais que signifie cette façon de prendre son fantasme pour de l'imagination, ou pire ses fictions pour de la science ? Serait-ce de la naïveté ou du calcul ? De la bêtise pour cultiver son ego ou de la stratégie médiatique pour faire mieux faire passer la pilule IA dans son business le plus quotidien ?

## Les contre-fictions de la littérature de science-fiction

La science-fiction, lorsqu'elle ne se contente pas de divertir, fabrique des souvenirs du futur. Et ces souvenirs, à l'instar de dispositifs d'art contemporain, deviennent autant de tableaux dessinant des hypothèses de l'à venir sur le registre d'une humaine imperfection. Autant de points de repère, de bouées intellectuelles et spirituelles pour ne pas succomber aux idéologies ou ne pas nous noyer sous les océans de promesses des vendeurs de bonheur en kit scientifique ou technologique de notre aujourd'hui. Et pour anticiper ici et maintenant ce qui était inconcevable hier ou avant-hier, avec ou sans intelligence artificielle.

Dans son décryptage du mythe de la Singularité, analysant avec minutie le *storytelling* autour de l'avènement prochain d'une hyperintelligence à même de nous dépasser, Jean-Gabriel Ganascia écrit : « *Outre cette argumentation soumise plus à l'impératif narratif qu'à la logique et à la rigueur démonstrative, tout se passe aussi comme si science et science-fiction permutaient leurs rapports logiques et chronologiques : des scientifiques et des ingénieurs tirent désormais les justifications de leur recherche de la science-fiction alors qu'originellement l'inverse prévalait, en cela que les résultats scientifiques servaient d'aliment à l'imaginaire*

d'écrivains et de cinéastes<sup>17</sup>. » Mais ce rapport entre la réalité de la recherche et les rêves de la science-fiction est-il toujours à ce point univoque ? Les idéologies comme les imaginaires ne vivent-ils pas en amont autant qu'en aval de toute démarche scientifique et technologique ? Autrement dit : la confusion palpable que dénotent les discours de bien des technologues de la Silicon Valley ne vient-elle pas d'un glissement de sens et d'un mélange des genres bien plus que d'une méconnaissance de la flèche du temps, ne pouvant partir que de la science pour entrer ensuite dans l'imaginaire – et pas le contraire – comme l'affirme un peu vite Ganascia ?

Quand Vernor Vinge, auteur de science-fiction, met en scène la « Singularité » comme l'une des clés contextuelles de son roman *La captive du temps perdu*, publié en 1986, il n'assène aucune vérité. Chacun de ses personnages, selon ses lubies, avance une hypothèse différente sur ce qui a pu susciter cette « Singularité », autrement nommé « L'Extinction ». Apocalypse écologique ? Intervention de quelque force extraterrestre ? Disparition de l'humanité par la grâce de machines superintelligentes ? L'un de ses personnages affirme certes : « Nous savions dès 2200 accroître notre intelligence. Et l'intelligence est la base de tout progrès. (...) Il n'y a pas eu d'Extinction, Wil. L'humanité a atteint un état supérieur<sup>18</sup>. » Mais une autre figure du livre dit à peu près l'inverse : « La science a donné à chaque animal humain l'audace d'agir comme un petit dieu. La Terre ne pouvait tout simplement plus le supporter. Merde, je suis certaine qu'il n'y a même pas eu de guerre. Que toute la structure s'est effondrée sous son propre poids, abandonnant les violeurs à la merci de leur victime : la nature<sup>19</sup>. » Or il se trouve que Vernor Vinge est aussi mathématicien à l'université de San Diego, et qu'il est communément crédité par le courant transhumaniste pour avoir été le premier à affirmer, en 1993, le concept dans un court essai se voulant scientifique, justement titré « La Singularité ». Sauf que le texte regorge de clichés à l'emporte-pièce sur « les mystères du développement humain » pour la plupart « élucidés » ou surtout « l'émergence d'un intellect superhumain<sup>20</sup> » pouvant aboutir à notre extinction. Là où le romancier ouvrait les portes de la réflexion en multipliant les éventualités d'un futur tragique, le chercheur peu rigoureux avance des hypothèses de fiction plus ou moins plausibles comme s'il s'agissait de vérités scientifiques. L'idée même de cette rupture qu'on appelle « Singularité » a été évoquée de façon moins précise dans les années 1960 ou 1970 par des scientifiques tel le mathématicien Stanislaw Ulam, mais aussi dans une nouvelle d'Isaac Asimov, comme le dit justement Ganascia. En revanche,

l'enjeu n'est pas, comme il semble le sous-entendre, de positionner systématiquement la science avant le travail de l'imaginaire, mais de faire en sorte que l'un et l'autre territoire s'alimentent en permanence sans se confondre l'un l'autre.

Au-delà des imaginaires de l'IA, Elon Musk raconte volontiers avoir été marqué par le cycle de *Fondation*, d'Isaac Asimov, et ses récits dans le temps long d'une lutte de l'humanité – devenue « espèce multiplanétaire » – contre son inéluctable déclin. Aussi est-ce savoureux de constater que le critique le plus vif de son plan de colonisation de Mars, présenté fin 2016, a été Kim Stanley Robinson, auteur d'une monumentale trilogie martienne : le plan de Musk, dit-il non sans humour, « ressemble à un gros cliché de science-fiction des années 1920, imaginé par un garçon qui tente de construire une fusée pour atteindre la lune depuis son jardin à l'arrière cours<sup>21</sup>. »

Quand Ian M. Banks invente « *la Culture* », utopie libertaire et civilisation intergalactique d'un très lointain futur dont les « *Mentaux* », drones et autres Intelligences artificielles dessinent une « intelligence *alien* », il ne valide aucune thèse *a priori*. Il pratique juste un exercice d'anthropologie expérimentale, il est vrai moins fondamentalement techno-critique que celle de Philip K. Dick lorsqu'il crée dans une nouvelle le *Swibble*, « *métazoaire télépathique artificiel* » qui réduit à un simple problème de plomberie technologique la question de l'obéissance de tous les citoyens à l'idéologie dominante. La science-fiction crée d'un même élan l'utopie et la contre-utopie, la fascination et la répulsion vis-à-vis de nos futurs numériques, computationnels ou biotechnologiques. À l'instar de toute technologie, elle cultive bien des caractères du *pharmakon* de l'antiquité, à la fois poison pour qui se laisse dériver aux illusions dominantes et remède pour celui qui saurait en fabriquer une thérapie.

Ses souvenirs du futur peuvent nous aider à décrypter et à rire des pires affabulations transhumanistes sur l'intelligence artificielle, sans pour autant refuser l'immensité de ses potentiels. Leur beauté tient à l'infinité des subjectivités qu'ils placent en orbites philosophiques, sociologiques, culturelles, spirituelles ou tout simplement divertissantes. Nos subjectivités ont d'abord été *littérairement* numérisées par les imaginaires de la SF, bien avant d'être effectivement computées par les algorithmes des Big data et des intelligences artificielles. Et les machines littéraires, notamment quand elles sont reliées par les jeux vidéo, les séries comme *Black Mirror* et des films comme *Matrix*, s'avèrent en fin de

compte aussi puissantes que les ordinateurs les plus performants. Dès lors, nourrir nos subjectivités des multiples hypothèses et de la pluralité des univers de la science-fiction peut nous servir à contrer la pauvreté intellectuelle et spirituelle des fantasmes médiocres et imaginaires de dieux *liliputiens* des apôtres de cette IA censée nous dépasser... Cette démarche ouvre en effet nos esprits aux errements et explorations mentales qui s'avèrent indispensables pour nos réponses les plus techno-critiques aux fabriques du conformisme de l'hypercapitalisme de la Silicon Valley.

Par la grâce de la subjectivité technoscientifique, ce qui était de l'ordre du divin devient de l'ordre de l'humain. Et c'est la science-fiction qui, depuis un siècle, actualise l'imaginaire de cette nouvelle toute-puissance en réinventant les mythes éternels de l'humanité : elle met en scène la prétention ahurissante ou la vérité inconsciente du « devenir dieu » de l'homme. Et ce faisant, elle ouvre le champ à une pacifique guerre des subjectivités. Sous ce regard, n'est-il pas urgent de répondre aux fictions en actes des rêves d'IA *made by Silicon Valley*, avouées ou non, par des *contre-fictions* décapantes nous plongeant dans la pluralité des IA possibles ?

---

1 Rapport de Cédric Villani : « Donner un sens à l'intelligence artificielle (IA), Pour une stratégie nationale et européenne », 28 mars 2018, Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, page 9.  
<http://www.enseignementsup-recherche.gouv.fr/cid128577/rapport-de-cedric-villani-donner-un-sens-a-l-intelligence-artificielle-ia.html>

2 Rapport de Cédric Villani, *op. cit.*, « Perte de commande », nouvelle d'Anne-Caroline Paucot, pages 68 à 70.

3 Rapport de Cédric Villani, *op. cit.*, « Impulsion naturelle », nouvelle d'Olivier Paquet, pages 94 et 95.

4 L'ensemble des développements qui suivent utilisent notamment un article publié dans le n°205 (avril 2018) de la revue *Raison présente*, « L'humain augmenté, pour quoi faire ? » : Ariel Kyrou, « Les imaginaires transhumanistes de la Silicon Valley », pages 49 à 61.

5 Voir pour l'ensemble de cette partie « *Nos subjectivités baignent dans un imaginaire de science-fiction* », par Ariel Kyrou, *Multitudes* n°62, printemps 2016, p. 126-132.

6 « *L'intelligence artificielle de Google a élaboré sa propre IA, qui surpasse tout ce qui a été fait jusqu'à présent par les humains* », par Stéphanie Schmidt, *Trust My Science*, 4 décembre 2017.

<http://trustmyscience.com/une-ia-genere-une-nouvelle-ia-plus-performante-que-toutes-les-autres/>

7 Jean-Gabriel Ganascia, *Le mythe de la Singularité. Faut-il craindre l'intelligence artificielle ?*, Science ouverte / Seuil, février 2017, p. 51 et 52.

8 Ariel Kyrou, *Google God, Big Brother n'existe pas, il est partout*, Inculte (2010), p. 205-216.

9 George Alec Effinger, *Gravité à la manque*, Denoël / Présence du futur (1987, 1989 pour la traduction française), p. 12.

10 « *La raison instrumentale, le capitalisme algorithmique et l'incomputable* », par Luciana Parisi, Multitudes n°62, printemps 2016, p. 98-109.

11 « *Elon Musk | Full interview | Code Conference 2016* » (1h24'14").  
<https://www.youtube.com/watch?v=wsixsRI-Sz4>

12 « *We are already cyborgs | Elon Musk | Code Conference 2016* » (5'11").  
<https://www.youtube.com/watch?v=ZrGPuUQsDjo&t=119s>

13 Evgeny Morozov, *Pour tout résoudre, cliquez ici ! L'aberration du solutionnisme technologique*, Fyp Éditions, septembre 2014. Voir aussi : « *Evgeny Morozov, contre l'Internet centrisme* », par Ariel Kyrou, Culture Mobile, 7 novembre 2014.  
<http://www.culturemobile.net/visions/evgeny-morozov-contre-internet-centrisme>

14 « *Clash entre Mark Zuckerberg et Elon Musk sur l'intelligence artificielle* », par Gregory Rozieres, Huffpost, 25 juillet 2017.  
[http://www.huffingtonpost.fr/2017/07/25/clash-entre-mark-zuckerberg-et-elon-musk-sur-lintelligence-arti\\_a\\_23046661/](http://www.huffingtonpost.fr/2017/07/25/clash-entre-mark-zuckerberg-et-elon-musk-sur-lintelligence-arti_a_23046661/)

15 « *Facebook travaille sur une interface cérébrale qui vous permettra de "communiquer uniquement avec votre esprit"* », Transhumanisme et intelligence artificielle, 22 avril 2017.  
<https://iatranshumanisme.com/2017/04/22/facebook-travaille-sur-une-interface-cerebrale-qui-vous-permettra-de-communiquer-uniquement-avec-votre-esprit/>

16 « *Intelligence artificielle, Promesses et périls* », Le Monde, 31 décembre 2017, 1<sup>er</sup> et 2 janvier 2018, dossier de huit pages, avec des articles signés notamment Morgane Tual, David Larousserie, Sandrine Cassini, Alexandre Piquard et Yves Eudes.

17 Jean-Gabriel Ganascia, *Le mythe de la Singularité. Faut-il craindre l'intelligence artificielle ?*, Science ouverte / Seuil, février 2017, p. 72.

18 Vernor Vinge, *La captive du temps perdu*, Le Livre de Poche / L'Atalante, 1986 pour la version originale, 1996 pour la traduction française, p. 168.

19 *op.cit.*, p. 114-115.

20 Vernor Vinge, *La Singularité*, 1993, traduit de l'anglais par Emmanuel Mie.

<http://www.lesconfins.com/SINGULARITE..pdf>

21 « *Why Elon Musk's Mars Vision Needs "Some Real Imagination"* », par Éric Roston, Bloomberg, 17 octobre 2016.