

DE FACTO

GASPARD KOENIG

La fin de l'individu

Voyage d'un philosophe au pays
de l'intelligence artificielle

La Fin de l'individu

Du même auteur

ESSAIS :

Voyages d'un philosophe aux pays des libertés, Éditions de l'Observatoire, 2018.

Time to Philo, Larousse, 2017.

Les Aventuriers de la liberté, Plon, 2016.

Le Révolutionnaire, l'expert et le geek. Combat pour l'autonomie, Plon, 2015 (prix Turgot, prix Zerilli-Marimò de l'Académie des sciences morales et politiques).

Leçons sur la philosophie de Gilles Deleuze, Ellipses, 2013.

Leçons de conduite, Grasset, 2011.

Les Discrètes Vertus de la corruption, Grasset, 2009.

ROMANS :

Kidnapping, Grasset, 2016.

La Nuit de la faillite, Grasset, 2013.

Un baiser à la russe, Grasset, 2006 (prix Publicis).

Octave avait vingt ans, Grasset, 2004 (prix Jean-Freustié).

Dans la même collection

Gaspard Kœnig, *Voyage d'un philosophe aux pays des libertés*, 2018.

Peter Grete, *Voyage d'un reporter au pays de la censure*, 2018.

Nicolas Gardères, *Voyages d'un avocat au pays des infréquentables*, 2019.

Gaspard Kœnig

La Fin de l'individu

Voyage d'un philosophe
au pays de l'intelligence
artificielle

Collection « De Facto »

L'Éditions de
Observatoire

Cet ouvrage est publié dans le cadre d'un partenariat
entre *Le Point* et les Éditions de l'Observatoire.

ISBN : 979-10-329-0721-4

ISSN « De Facto » : 2648-7888

Dépôt légal : 2019, septembre

© Éditions de l'Observatoire/Humensis 2019
170 *bis*, boulevard du Montparnasse, 75014 Paris

Collection « De Facto »

« Alexis de Tocqueville déplorait “l’effrayant spectacle” des philosophes français, enfermés dans des spéculations abstraites : “même attrait pour les théories générales, les systèmes complets de législation et l’exacte symétrie dans les lois ; même mépris des faits existants ; même confiance dans la théorie.”

C’est pour remédier à ce travers national que “De Facto” accueille, autour du thème de la liberté, des textes à la première personne mêlant théorie et pratique, idées et expériences, réflexion et récit. Nos auteurs peuvent aussi bien être des théoriciens (de toutes disciplines) exposés à la pratique, que des praticiens forgeant leur théorie. Ils alternent rigueur argumentative et sincérité personnelle.

Plutôt que d’être simplement confronté à un système de pensée, le lecteur est ainsi embarqué dans une aventure intellectuelle, avec ses doutes, ses révélations, ses combats. Un voyage, au sens métaphorique comme souvent au sens propre.

De quoi redonner à la pensée la vigueur de l’expérience et la sève de la vie ! »

Gaspard Kœnig
directeur de collection

Homo deus

« *Sorry, I can't help you.* » Telle fut la réponse laconique d'Eliezer Yudkowsky, l'un des chercheurs les plus en pointe de la Silicon Valley sur l'intelligence artificielle (IA), à ma demande d'entretien. On ne peut mieux dépeindre le gouffre qui s'est creusé entre les maîtres de la technologie et le public, entre ceux qui créent les algorithmes et ceux qui vivent sous leur règne, entre ceux qui écrivent les lignes de code et ceux qui essayent de les comprendre.

Sorry, I can't help you. Autrement dit : je le voudrais bien, mais vraiment je ne peux pas. Comment un barbare qui n'a jamais codé, un profane qui peine encore à utiliser PowerPoint pourrait-il comprendre les subtilités du *deep learning* ? Ici, on est occupés à *make the world better*. On ne se pose pas de questions, on résout des problèmes. Les tribulations philosophiques, luxe d'un esprit oisif, ne sont pas à l'ordre du jour. C'est la réponse de Gorgias à Socrate : il faut arrêter les enfantillages.

Sorry, I can't help you. Je ne serais pas surpris d'apprendre que ce message avait été généré automatiquement par Gmail. M. Yudkowsky doit recevoir de nombreuses demandes similaires. L'algorithme a dû apprendre à les reconnaître, ainsi qu'à suggérer sa

réponse favorite. Quoi de plus naturel qu'une IA gère les questions sur l'IA ?

Et pourtant, nous avons plus que jamais besoin de nous parler, de nous entraider, pour comprendre le bouleversement majeur qui inquiète nos sociétés, rebat les cartes de nos économies, ébranle nos systèmes politiques et envahit nos existences, nous laissant partagés entre l'espoir du progrès et l'inquiétude de l'avenir. Lors des précédents épisodes de rupture technologique, penseurs, inventeurs, scientifiques, investisseurs et politiques se trouvaient rassemblés dans des lieux uniques, centres névralgiques des « économies-mondes » décrites par Fernand Braudel. Ces villes-mondes représentaient non seulement des carrefours économiques mais aussi de hauts lieux de culture. « La splendeur, la richesse, le bonheur de vivre, écrit Braudel, se rassemblent au centre de l'économie-monde, en son cœur. C'est là que le soleil de l'histoire fait briller les plus vives couleurs¹. » C'est ainsi que Spinoza sema les graines d'une philosophie de l'immanence à Amsterdam, capitale du Siècle d'or, qu'Adam Smith put théoriser le capitalisme depuis Édimbourg, au cœur de la révolution industrielle, ou que Karl Marx conçut la lutte des classes dans le Londres victorien. Les villes servaient de pôle d'attraction et de creuset intellectuel. Les esprits bouillonnants y opéraient des fusions étranges, hasardeuses et parfois miraculeuses. Force est de constater qu'aujourd'hui les cerveaux sont davantage éparpillés. Personne n'est capable de désigner le centre mondial de l'IA. Même la Silicon Valley avec ses banlieues-dortoirs ne saurait passer pour le parangon du « bonheur de

1. Fernand Braudel, *La Dynamique du capitalisme* [1977], Arthaud, 1985.

vivre » ; San Francisco, devenue l'une des villes les plus chères du monde, chasse les jeunes innovateurs ; la contre-culture californienne des années 1970 s'efface dans une indifférence patente pour tout ce qui concerne les sciences humaines, comme si la passion du changement évacuait toute réflexion anthropologique, comme si l'être humain était une pâte à modeler sans biologie ni histoire. *The Economist* proposa d'ailleurs il y a quelques années aux leaders de la « tech » de prendre des cours de philosophie, sans succès¹. N'y a-t-il plus de ville-monde ?

C'est donc pour reconstituer une sorte de ville-monde virtuelle, pour tâcher de jeter un pont entre les fulgurances de la tech et les permanences de la métaphysique, que j'ai entrepris un long voyage autour de l'IA. Au cours de plusieurs mois, je me suis entretenu avec 125 spécialistes, plus chaleureux qu'Eliezer Yudkowsky, ou simplement lassés de mes relances incessantes : chercheurs, entrepreneurs, investisseurs, professeurs, régulateurs, artistes... J'ai voulu les rencontrer là où ils vivent et travaillent, dans leur habitat naturel fait d'ordinateurs et d'embouteillages, ce qui m'a entraîné dans un tour du monde par l'ouest : Cambridge, Oxford, Boston, New York, Washington, San Francisco, Los Angeles, Shanghai, Pékin, Tel-Aviv, Copenhague et enfin Paris. Au passage, à force de chaparder des heures à la course du soleil, j'ai sincèrement cru que mon existence serait rallongée d'une journée, avant de m'apercevoir en survolant le détroit de Béring que la ligne de changement de date, qui court le long du 180° méridien, me reprenait tout. J'avais fait la même erreur que Phileas

1. <https://www.economist.com/business/2014/10/04/philosopher-kings>

Fogg, mais en sens inverse : c'est vous dire si mon esprit n'est pas scientifique.

La route ne fut pas de tout repos. Elle commença au laboratoire IA de Facebook en Europe, dédié à la recherche fondamentale. En voyant une jeune femme absorbée par des milliers de lignes de code sur une bonne demi-douzaine d'écrans, j'ai pris conscience que je m'aventurais déraisonnablement dans le domaine du sacré, et qu'on ne franchit l'iconostase qu'à ses risques et périls. En l'occurrence, cette chercheuse avait pour ambition de prédire automatiquement le déplacement des objets dans une rue à partir d'une simple image : telle voiture va-t-elle démarrer, tel piéton va-t-il traverser, tel enfant va-t-il laisser tomber son ballon... ? Je me reculai instinctivement, de peur peut-être de voir apparaître une ligne qui décrive mon propre comportement, comme si ces millions de millions de 1 et de 0 pouvaient contenir l'ensemble de la réalité passée et future.

Je n'étais pas au bout de mes frayeurs. Quelques jours plus tard, Aurélie Jean, une jeune informaticienne passée par le Massachusetts Institute of Technology (MIT), profita témérairement d'un voyage en TGV pour m'initier au Python, un des plus célèbres langages informatiques. Ce qui acheva de me traumatiser fut de voir sur l'écran d'ordinateur d'Aurélie non pas des dossiers et des fichiers, mais une simple fenêtre noire remplie de signes cabalistiques. Car Aurélie, comme nombre de ses semblables, ne s'abaisserait pas à cliquer banalement avec une souris sur les icônes trop commodes d'une interface utilisateur. Elle travaille sous le capot de la machine, au plus proche de ses fonctions premières. Elle lui donne ses instructions sous forme

de code. Au lieu d'ouvrir des dossiers pour accéder à un document contenant du texte, par exemple, elle ordonne à l'ordinateur, dans un langage qu'il comprend, d'aller le chercher. Elle a ainsi le sentiment de communiquer avec l'outil informatique de manière plus naturelle. Nous autres profanes sommes comme des enfants qui, pour effectuer des opérations mathématiques primaires, doivent ajouter et soustraire des parts de gâteau : nous avons besoin d'une représentation (c'est d'ailleurs ce qui a fait le succès de Microsoft et d'Apple dans les années 1980). Aurélie, elle, manipule directement les chiffres. Elle se dispense d'une strate symbolique supplémentaire. « Et c'est plus rapide », osa-t-elle me dire en mitraillant son clavier.

L'IA multiplie les embûches pour décourager les explorateurs. Elle semble se développer de préférence dans les villes les plus congestionnées du monde : je comprends mieux après des dizaines d'heures de surplace la passion des geeks pour la ville intelligente et les voitures autonomes. Surtout, la technologie la plus révolutionnaire de ces dernières décennies se niche dans une science mystérieuse. Il existe peu d'ouvrages de synthèse destinés aux néophytes. Je ne prétendrai pas avoir dépassé l'introduction de la bible des étudiants en informatique, *Artificial Intelligence* par Stuart Russell et Peter Norvig : une fois compris une poignée de définitions et de rappels historiques, la matière devient exponentiellement technique, me rappelant sans plaisir toutes les raisons qui m'avaient fait abandonner la terminale scientifique pour une terminale littéraire après quelques semaines de calvaire. Néanmoins, à force de lectures et de conversations autour de l'IA, je pense en avoir acquis une certaine « teinture », selon le mot de Montaigne pour désigner nos connaissances toujours

imparfaites. Une teinture nécessaire, quoique jamais suffisante, pour prétendre philosopher. Une teinture empreinte de hasards, de découvertes, d'obsessions : il faut accepter dans le reportage une part de chance et de malchance, de révélations et d'ignorance. Mon agenda était souvent à moitié vide en arrivant dans une ville et se remplissait au fur à mesure des rencontres. J'ai embrassé dans mon enquête cette sérendipité que l'IA voudrait justement abolir.

Durant mes quatre semaines sur la côte Ouest, il m'est arrivé une seule fois d'entrer dans un bureau où les étagères croulaient sous les classiques : à la fondation de Peter Thiel à Los Angeles, je retrouvai soudain mes marques, entre les volumes de Saint-Simon en Pléiade et les œuvres de René Girard. J'eus même la surprise de découvrir, dans le salon d'attente, un exemplaire des *Discours sur la condition des grands* de Pascal. Il ne serait sans doute pas inutile, pour un entrepreneur de la tech en pleine levée de fonds, de se voir rappeler de temps à autre la distinction entre « grandeur d'établissement » et « grandeur naturelle ». La première, liée au statut social, implique une déférence parfaitement légitime envers les puissants mais qui ne saurait préjuger des qualités humaines réelles liées à la seconde. Pascal n'est pas un révolutionnaire : il ne propose pas de renverser les grands de ce monde mais nous invite à entretenir une « double pensée » qui sache distinguer les conventions sociales des vertus morales. C'est une recommandation fort à propos pour nos entrepreneurs qui, en se drapant des faux-semblants de l'authenticité, à coups d'emojis et de selfies, feignent d'ignorer les rapports de pouvoir et de capital qui gouvernent leur relation à autrui. Confondraient-ils eux aussi leur grandeur d'établissement avec une grandeur naturelle ?

Car l'étiquette de la Valley n'a rien à envier à celle des cours de jadis. Le cool a généré ses propres codes, dont le respect est vital pour prospérer ou simplement exister dans cet écosystème de compétition permanente. J'ai très vite constaté qu'un email solitaire, envoyé avec fougue et bonne volonté (« *Hi Mark ! I am a French philosopher* »), était voué au silence éternel. Comme toujours, il ne faut pas croire la pub : rien là-bas n'est horizontal, fluide ni transparent. « *Think different* », mais pas trop. Sauf exception, un rendez-vous ne s'obtient qu'après un long parcours de mise en relation ; solliciter une introduction est un préliminaire indispensable qui requiert lui-même de curieuses contorsions rhétoriques. Oublier le point d'exclamation ou le smiley trahit un manque d'enthousiasme impardonnable. Il faut à la fois être sérieux, paraître badin, démontrer son ardeur et suggérer son dévouement, le tout en trois bullet-points : que Saint-Simon semble simpliste en comparaison ! Dans le monde du Google Calendar, il existe encore des armées de secrétaires pour filtrer les quémandeurs. À l'ère du post, rien ne fonctionne mieux que le bouche-à-oreille. Un Français expatrié de longue date à San Francisco a d'ailleurs écrit un article juste et drôle décrivant les règles strictes de « Silicon-Versailles », de la gestion du temps aux formules de politesse¹.

Il suffit de se rendre au restaurant Madera sur Rosewood Sand Hill, l'épicentre du capital-risque, le « temple du deal » selon l'expression de l'investisseur qui m'y invitait, pour constater combien le monde de la Valley reste codifié et hiérarchisé. Tout d'abord, le

1. <https://medium.com/@romainserman/silicon-valley-etiquette-6934cf6f8f73>

lieu est introuvable, déjouant le GPS. « C'est à dessein, me dit mon hôte en mordant dans le hamburger le plus cher du monde. C'est un monde d'*insiders*. Ce n'est pas démocratique ici. » À travers les baies vitrées, on aperçoit le panorama paisible de la réserve de Jasper Ridge ; sous le ciel bleu de l'été californien, les massifs forestiers moutonnent en vagues successives. On devine les séquoias au tronc rougeoyant, muscles à vif palpitant dans le vent. Au premier plan scintillent sous le soleil des oliviers parfaitement alignés. Quelques dizaines de maisonnettes d'allure modeste s'égrènent le long d'allées ombrées. « Ce sont les plus gros fonds de *venture capital* (capital-risque) du monde : Sequoia Capital, Menlo Ventures, Schlumberger, Makena Capital, Andreessen Horowitz, Coaetue Management, Silver Lake Partners, Kleiner Perkins... Le cash est là, autour de nous. Plusieurs dizaines de milliards prêts à être déboursés à tout instant. » L'argent s'est mis au vert, trompant ceux qui le cherchent encore au sommet de gratte-ciels en verre. Les entrepreneurs font leur pèlerinage à pied, toquant de porte en porte comme des enfants pour Halloween. On les reconnaît aisément : ils ont fait l'effort de mettre une veste, parlent vite et sourient trop. Les autres, les investisseurs, ceux qui détiennent la clé des ambitions, les reçoivent nonchalamment, si possible chaussés de baskets fluo. C'est le capitalisme en jean, énième variation sur les drapés noirs des Fugger ou les costumes impeccables des Rothschild. Contrairement à ce que nous répètent les génies de la disruption à longueur de TED Talks, le monde ne change pas tant.

Comme le conclut Pascal, « il n'est pas nécessaire, parce que vous êtes duc, que je vous estime, mais il est nécessaire que je vous salue ». De même dans la

Valley, il n'est pas nécessaire, parce que vous êtes *venture capitalist*, que je vous aime, mais il est nécessaire que je vous *like*.

La leçon ultime de mon hôte, avant de bondir vers un *meeting* naturellement urgent, c'est que ni les entrepreneurs ni les investisseurs n'ont la moindre idée de l'impact social et politique des technologies qu'ils créent. À ses yeux, l'intelligence artificielle s'accompagne d'une forte dose d'« intelligence superficielle ». Afin de retendre le lien primordial entre innovation technologique et réflexion philosophique, il faut davantage de Pascal sur Rosewood Sand Hill : lui-même n'était-il pas entrepreneur, créateur des premiers autobus urbains, les « carrosses à cinq sols » ?

Il faut néanmoins résister à la tentation facile de la technophobie. Rétrospectivement, les prophètes d'apocalypse nous paraissent bien ridicules : ainsi Paul Valéry qui, il y a près d'un siècle, dénonçait « l'intoxication insidieuse » du progrès technique en se lamentant (déjà !) de la disparition du temps libre, de la tendance à parcourir les livres plutôt qu'à les lire, de la dictature de l'émotion... « Le courrier ni le téléphone ne harcelaient Platon¹ », regrette Valéry. Pauvre poète, interrompu par le facteur ! Qu'aurait-il pensé des notifications et des tweets ! Il faut se faire violence, et faire violence au lecteur, pour tâcher de comprendre son époque sans la maudire.

D'autant que l'intelligence artificielle devrait être le rêve de tout philosophe. Ne serait-il pas commode de fabriquer une machine pensante qui nous dispense des

1. Paul Valéry, *Le Bilan de l'intelligence* [1935], Allia, 2016.

fautes de logique, des préjugés individuels et des errements conceptuels, un algorithme qui calcule la vérité et nous donne enfin, après des millénaires de controverses répétitives, la réponse à nos questions les plus existentielles ? La pensée conceptuelle n'est au fond qu'une approximation ; tandis qu'un système complet de symboles, gouverné par des lois scientifiques, permettrait de se tenir au plus près de la vérité. Le premier à en rêver fut Leibniz, génie des mathématiques et de la métaphysique, qui chercha la formule d'une sorte de machine à calculer la pensée, une « Caractéristique universelle » capable de raisonner juste. Une combinatoire géante, que Leibniz baptise « *calculus ratiocinator* », évacuerait automatiquement les chimères de l'esprit. Dans ce monde parfaitement rationnel, « il ne sera plus besoin entre deux philosophes de discussions plus longues qu'entre deux mathématiciens, puisqu'il suffira qu'ils saisissent leur plume, qu'ils s'asseyent à leur table de calcul et qu'ils se disent l'un à l'autre : Calculons¹ ! »

On retrouve cet idéal chez tous les grands précurseurs de l'intelligence artificielle : Hilbert, Frege, et bien sûr Alan Turing, qui entretenirent tous des relations étroites avec la logique et la philosophie analytique². Si l'exactitude logique est ce qui peut permettre à l'esprit humain d'appréhender au mieux le réel, alors la conversion du réel en une combinaison chiffrée pourrait avantageusement remplacer l'esprit humain. D'où la réplique hilarante de Deep Thought, le superordinateur de la série *Hitchhiker's Guide to*

1. G. W. Leibniz, *Nova methodus pro maximis et minimis*, 1668.

2. Pour une histoire intellectuelle de l'IA, on pourra lire Martin Davis, *The Universal Computer*, W. W. Norton, 2000.

the Galaxy, quand on lui demande de résoudre « la question ultime de la vie, de l'univers et du reste » : 42. Un nombre devenu mythique, qui continue à animer les conversations des geeks. Ne serait-ce pas apaisant si toutes les significations qui travaillent nos existences pouvaient se résoudre en un chiffre totale-ment a-signifiant ?

Cette affinité entre science informatique et logique perdure aujourd'hui et poursuit le visiteur dans l'immeuble biscornu du département de *Computer Science* rattaché au MIT : quand je pénétrais dans le bureau de Leslie Kaelbling, une vétérane de la recherche en IA, je la trouvais entourée de livres de philosophie analytique, et passionnée par l'œuvre de Willard Quine. Cette approche est assez éloignée de la tradition intellectuelle de l'Europe continentale. À la fin du siècle dernier, Gilles Deleuze a poussé l'opposition à son terme en définissant la philosophie comme « production de concepts », une activité relevant d'un processus créatif davantage que de la rigueur scientifique. Un concept n'est pas résoluble en une succession de 1 et de 0 : il projette sur notre monde une signification nouvelle, une perspective supplémentaire. Il y a souvent un moment d'épiphanie quand, en reprenant pour la dixième fois un aride traité métaphysique, on en « comprend » soudain le sens. Tous les concepts s'emboîtent alors naturellement et l'œil glisse sur les pages avec une facilité inattendue. Comme si les arguments développés au cours des chapitres étaient moins des étapes sur un chemin logique que des lentilles de vue modifiant progressivement notre perception des choses. On décortique en s'arrachant les cheveux les propositions de l'*Éthique* de Spinoza, on s'escrime à en analyser l'enchaînement, et voilà qu'à un moment,

sans crier gare, surgit le « *Deus sive Natura* » dans toute sa profondeur conceptuelle. On peut alors relire les propositions qui nous avaient tant mystifié : elles paraîtront d'une évidence étonnante. C'est un phénomène qui ne mobilise pas seulement l'esprit mais, d'une manière confuse, l'ensemble de notre corps. « Une modification de ses rapports de mouvement et de repos », dirait Deleuze.

À l'inverse, je me rappelle mon effarement quand, en suivant des cours de philosophie à Columbia University, on m'a fait décomposer des phrases en équations, en cherchant la « vérité » d'un énoncé dans le jeu des prémisses et des déductions : en quoi ces exercices d'écolier studieux avaient-ils le moindre rapport avec la pensée ? De leur côté, les Américains considèrent avec scepticisme ce qu'ils ont baptisé la *French Theory*, où ils ne voient bien souvent qu'une fanfaronnade poétique. Il n'est pas impossible que le rejet de la philosophie analytique contribue à expliquer, chez les Européens, une défiance instinctive à l'égard de l'IA. En tout état de cause, il faut se rappeler que l'IA, avant de devenir une technique industrielle, représente un projet philosophique de compréhension du monde.

J'avais également une raison plus personnelle d'entreprendre ce long périple. Libéral, je défends l'idée d'un individu autonome, libre de ses choix et responsable de ses actions, donc censé faire usage d'une forme de libre arbitre. Une idée qui se trouve peu ou prou au fondement de nos sociétés occidentales depuis les Lumières, et qui justifie à la fois les droits individuels, les mécanismes de marché, le droit de vote et la justice pénale. La Cour suprême des États-Unis ne s'y est d'ailleurs pas trompée, en consacrant le libre