

**CATHY O'NEIL**  
**ALGORITHMES**  
**LA BOMBE À RETARDEMENT**

**PRÉFACE DE CÉDRIC VILLANI**

**« LE MANUEL DU CITOYEN  
DU 21<sup>E</sup> SIÈCLE »**

**FINANCIAL TIMES**

# ALGORITHMES

## LA BOMBE À RETARDEMENT

Qui choisit votre université? Qui vous accorde un crédit, une assurance, et sélectionne vos professeurs? Qui influence votre vote aux élections? Ce sont des formules mathématiques.

Ancienne analyste à Wall Street devenue une figure majeure de la lutte contre les dérives des algorithmes, Cathy O'Neil dévoile ces « armes de destruction mathématiques » qui se développent grâce à l'ultra-connexion et leur puissance de calcul exponentielle. Brillante mathématicienne, elle explique avec une simplicité percutante comment les algorithmes font le jeu du profit.

Cet ouvrage fait le tour du monde depuis sa parution. Il explore des domaines aussi variés que l'emploi, l'éducation, la politique, nos habitudes de consommation. Nous ne pouvons plus ignorer les dérives croissantes d'une industrie des données qui favorise les inégalités et continue d'échapper à tout contrôle. Voulons-nous que ces formules mathématiques décident à notre place? C'est un débat essentiel, au cœur de la démocratie.

« UN LIVRE FASCINANT ET  
PROFONDÉMENT PERTURBANT »

YUVAL NOAH HARARI, AUTEUR DE *SAPIENS*



# ALGORITHMES

## LA BOMBE À RETARDEMENT



Titre original : *Weapons of Math Destruction.*  
*How Big Data Increases Inequality and Threatens Democracy.*  
Copyright © 2016 by Cathy O’Neil

© Les Arènes, Paris, 2018,  
pour la traduction française.

Les Arènes  
27 rue Jacob, 75006 Paris  
Tél. : 01 42 17 47 80  
[arenas@arenas.fr](mailto:arenas@arenas.fr)  
[www.arenas.fr](http://www.arenas.fr)

CATHY O'NEIL

# ALGORITHMES

## LA BOMBE À RETARDEMENT

PRÉFACE DE CÉDRIC VILLANI

TRADUIT DE L'ANGLAIS (ÉTATS-UNIS) PAR SÉBASTIEN MARTY

LES ARÈNES





*À tous ceux que l'on donne perdants*

# PRÉFACE

Impossible de ne pas remarquer Cathy O’Neil ! Son style, son regard malicieux et son charisme font effet dès la première rencontre, et ne cessent d’impressionner. Ton joyeux, propos graves, elle incarne la rigueur et l’engagement tout à la fois.

Cathy est aussi parfaitement légitime. Elle connaît bien les algorithmes pour les avoir pratiqués, conçus, vérifiés des années durant. Formée en géométrie algébrique et arithmétique, elle maîtrise tout aussi tout autant le monde des concepts et des équations. Revenue des *hedge funds*, elle a une conscience aigüe du monde de la finance algorithmique internationale. Engagée dans le mouvement Occupy Wall Street, elle est très au fait de la politique. Enfin, comme ancienne directrice d’un programme de formation de journalistes, le monde des médias lui est familier. Cette combinaison rare de talents rend sa contribution aux débats précieuse quand il s’agit de mêler politique, algorithmique et information des citoyens.

En 2016, Cathy lâcha une véritable bombe : *Weapons of Math Destruction*. Sous sa couverture jaune flamboyant, avec son logo morbide, c’était un concentré explosif qui devint immédiatement emblématique. Derrière le jeu de mots, auquel aucune traduction ne pourra rendre pleinement justice, le concept d’« armes de destruction mathématiques » annonçait l’angle nouveau qu’adoptait l’ouvrage : un courant de contestation de la mathématisation du monde, basé non sur l’idéologie ou sur les émotions, mais sur la dénonciation étayée des ravages massifs que ces outils peuvent produire.

Pour moi comme pour tant d’autres, la lecture de ce brûlot fut un coup de tonnerre. Chapitre après chapitre, Cathy O’Neil analysait des situations dans lesquelles l’intelligence artificielle avait un effet néfaste, parce qu’utilisée à mauvais escient, sans réflexion éthique, sans garde-fous, ou sans contrôle suffisamment rigoureux. L’ampleur des dégâts était annoncée

sans détour par le sous-titre : « Comment le traitement des données accroît les inégalités et menace la démocratie ».

Pour Cathy O'Neil, les coupables ne sont pas tant des acteurs que des systèmes pervers, dans des secteurs aussi divers que l'éducation, la justice, l'enseignement supérieur, les banques, le crédit, la planification des horaires de travail. Sous sa plume défilent des schémas opaques, non régulés, non contestables, opérant à grande échelle, justifiant la formule « destruction massive ».

Quand le gouvernement m'a chargé d'une mission sur l'intelligence artificielle, j'ai relu avec attention ce livre explosif, pour en incorporer les enseignements dans mon rapport. Et quand la mission est arrivée à son terme, j'ai invité Cathy à participer à la restitution, à discuter directement avec le Président Emmanuel Macron, et à partager son expérience et ses conseils avec un public aussi large que possible.

Car une chose est certaine : pour mettre fin aux mauvais usages de l'intelligence artificielle, il ne faut pas les cacher, ou se bercer de l'illusion qu'ils seraient réservés aux autres : il faut, au contraire, en tirer parti et en parler. Profiter de ce que les États-Unis, moins méfiants que les pays européens, ont déjà essuyé les plâtres avec des expérimentations instructives. En parler aussi pour convaincre nos concitoyens que tous ces problèmes ont bien été identifiés et que les pouvoirs publics ne seront pas naïfs, mais vigilants. Et si les pouvoirs publics baissent un jour la garde, ce sera à nous de les réveiller !

Certains chapitres m'ont été particulièrement précieux : le catastrophique algorithme de notation automatique des enseignants, qui aggrave les inégalités qu'il prétend corriger ; les effets pervers de la police prédictive, qui tend à renforcer la frustration et la défiance précisément là où c'est déjà le plus difficile ; les procédures automatiques d'attribution de crédit, incrustant pernicieusement les biais déjà pratiqués par les humains ; la faculté de manipuler les masses via les réseaux sociaux. Il y a dans tout cela de quoi faire réfléchir les lecteurs loin au-delà des considérations algorithmiques : car l'intelligence artificielle, dans sa mise en œuvre, reflète nos valeurs, notre société, notre âme parfois.

Malgré tout cela, Cathy et moi-même gardons l'espoir que demain, bien encadrée, l'intelligence artificielle pourra nous apporter aide et progrès inouïs !

La chercheuse française Laurence Devillers, grande pourfendeuse des dérives de l'IA, se plaît à dénoncer les méfaits de ce qu'elle appelle la « fée Clochette », cet algorithme qui simule une compagne attentionnée pour célibataires en mal de tendresse. Cathy O'Neil, quant à elle, incarne une fée Clochette bien humaine, la voix de la conscience qui nous incite à ne pas céder aux facilités, à la fascination de la technologie, et de n'utiliser celle-ci que dans des conditions de rigueur et d'éthique incontestables. Elle est bien décidée à ne pas baisser les bras, et à faire porter sa voix autant qu'il le faut pour que nous puissions conserver notre humanité.

Écoutez-la attentivement.

Cédric Villani

# INTRODUCTION

Quand j'étais enfant, j'avais pour habitude d'observer la circulation par la vitre, dans la voiture, et d'étudier les chiffres sur les plaques d'immatriculation. Je réduisais chaque plaque à des éléments fondamentaux, en me servant des nombres premiers qui la composaient.  $45 = 3 \times 3 \times 5$ . C'est ce qu'on appelle la factorisation, et c'était mon passe-temps favori. Les nombres premiers intriguaient tout particulièrement la mathématicienne en herbe que j'étais.

Cet amour des mathématiques a fini par devenir une passion. À 14 ans, je suis partie en stage de mathématiques pour l'été et suis rentrée en serrant un Rubik's Cube sur ma poitrine. Les mathématiques m'offraient un refuge bien ordonné, à l'abri de la confusion du monde réel. Elles nous propulsaient vers l'avant, élargissant le champ des connaissances démonstration après démonstration. Et je pouvais y apporter ma contribution. Je me suis inscrite en mathématiques à l'université, et j'ai poursuivi jusqu'au doctorat. Ma thèse portait sur la théorie algébrique des nombres, un domaine de recherche qui trouvait son origine dans toutes ces factorisations que j'avais effectuées enfant. J'ai finalement été nommée professeur adjoint au Barnard College, qui partageait son département de mathématiques avec celui de l'université Columbia.

Puis j'ai complètement bifurqué. J'ai quitté mon poste et je suis partie travailler comme analyste quantitatif chez D. E. Shaw, l'un des principaux fonds spéculatifs du pays. En abandonnant l'université pour la finance, je suis passée de la théorie abstraite à la pratique. Nos opérations sur les nombres se traduisaient en milliers de milliards de dollars, qui circulaient d'un compte à l'autre. Au départ, j'étais à la fois enthousiaste et épatée de travailler dans ce nouveau laboratoire – l'économie mondiale. Mais à l'automne 2008, un peu plus d'un an après mon arrivée, tout s'est brusquement effondré.

Le krach financier a révélé avec la plus grande clarté que les mathématiques, mon refuge de toujours, étaient non seulement intimement mêlées aux problèmes du monde, mais alimentaient également nombre d'entre eux. En récitant leurs formules magiques, les mathématiciens s'étaient rendus complices de la crise de l'immobilier, de la chute d'institutions financières majeures, de la montée du chômage. Grâce aux extraordinaires pouvoirs que je vénérerais tant, les mathématiques s'étaient associées à la technologie pour décupler le chaos et le malheur, conférant une ampleur et une efficacité redoutables à des systèmes que je savais désormais défectueux.

Si nous avions fait preuve de lucidité à ce stade, nous aurions tous pris un peu de recul pour comprendre le mauvais usage qui en avait été fait, et comment empêcher à l'avenir une catastrophe similaire. Au lieu de cela, dans le sillage de la crise, les nouvelles techniques mathématiques suscitèrent plus d'engouement que jamais, gagnant davantage de secteurs. Vingt-quatre heures sur vingt-quatre et sept jours sur sept, elles brassaient des pétaoctets d'informations, extraites pour l'essentiel des médias sociaux ou des sites de commerce en ligne. Et elles cessèrent de se focaliser uniquement sur les mouvements des marchés financiers mondiaux, pour s'intéresser de plus en plus aux individus. Les mathématiciens et les statisticiens étudiaient nos désirs, nos déplacements et notre pouvoir d'achat. Ils prédisaient le niveau de confiance que l'on pouvait nous accorder et calculaient notre potentiel en tant qu'étudiants, travailleurs, amants ou criminels.

C'était l'économie du « Big Data<sup>1</sup> », et elle promettait des gains spectaculaires. Un programme informatique pouvait parcourir en une ou deux secondes des milliers de curriculum vitæ ou de demandes de prêt, et les trier bien soigneusement, avec en tête de liste les meilleurs candidats. Non seulement une telle méthode faisait gagner du temps, mais elle était aussi présentée et vendue comme équitable et objective. Il n'était plus question d'individus pétris de préjugés fouillant parmi des monceaux de documents, mais de simples machines qui manipulaient des chiffres impartiaux. Autour de 2010, les mathématiques commencèrent à s'immiscer comme jamais auparavant dans les affaires humaines, recevant un accueil très favorable du public.

Mais je sentais venir les ennuis. Les applications mathématiques qui alimentaient l'économie des données reposaient sur des choix, eux-mêmes

faits par des personnes faillibles. Certains de ces choix procédaient sans nul doute des meilleures intentions possible. Pour autant, nombre des modèles utilisés encodait justement les préjugés, les malentendus et les partis pris humains au cœur des systèmes informatiques qui régissaient de plus en plus nos vies. Telles des divinités mystérieuses, ces modèles mathématiques étaient opaques, leurs rouages, invisibles de tous, sauf des plus grands prêtres en la matière : les mathématiciens et les informaticiens. Leurs verdicts, fussent-ils nuisibles ou erronés, étaient sans appel et ne souffraient aucune discussion. Et ils avaient tendance à punir les plus défavorisés et les opprimés, tout en rendant les riches encore plus riches.

J'ai trouvé un nom pour désigner ces modèles nocifs : ce sont des Armes de Destruction Mathématique, ou ADM en abrégé. En voici un exemple, dont je soulignerai tout du long les caractéristiques destructrices.

Comme c'est souvent le cas, tout était parti d'un objectif louable. En 2007, le nouveau maire de Washington D.C., Adrian Fenty, avait résolu d'améliorer les écoles les moins performantes de sa ville. Il avait du pain sur la planche : à cette époque, à peine un lycéen sur deux parvenait jusqu'au bac, et 8 % seulement des élèves de troisième obtenaient la moyenne en mathématiques. Fenty embaucha Michelle Rhee, championne de la réforme éducative, et lui confia le nouveau et puissant poste de recteur des écoles de Washington.

De l'avis général, les élèves n'apprenaient pas correctement parce que les professeurs faisaient mal leur travail. En 2009, Rhee lança donc un plan visant à écarter les enseignants les moins performants. C'est la tendance qui prévaut actuellement à travers tous les États-Unis dans les districts scolaires en difficulté, et du point de vue d'un concepteur de systèmes, le raisonnement s'avère tout à fait sensé : évaluer les enseignants, se débarrasser des pires et placer les meilleurs là où ils peuvent être le plus utiles. Dans le jargon des experts en données, cela a pour effet d'« optimiser » le système scolaire et, selon toute vraisemblance, d'assurer aux enfants de meilleurs résultats. À l'exception des « mauvais » professeurs, qui pourrait s'y opposer ? Rhee élaborait un outil d'évaluation baptisé IMPACT et, dès la fin de l'année scolaire 2009-2010, le district scolaire de Washington renvoya tous les enseignants dont le score se situait parmi les 2 % les plus bas. À la fin de l'année suivante, 5 % supplémentaires des effectifs, soit 206 enseignants, furent mis à la porte.

Sarah Wysocki, institutrice de l'équivalent du CM2, ne semblait avoir aucune raison de s'inquiéter. Elle n'enseignait à la MacFarland Middle School que depuis deux ans, mais recevait déjà d'excellents avis de la part de son directeur et des parents d'élèves. L'une de ces appréciations louait sa prévenance à l'égard des enfants ; une autre saluait en elle « l'une des meilleures enseignantes que j'aie rencontrées ».

Pourtant, au terme de l'année scolaire 2010-2011, Wysocki obtint un score déplorable lors de son évaluation IMPACT. Le problème venait d'un nouveau système de notation, connu sous le nom de « modèle de la valeur ajoutée<sup>2</sup> », qui prétendait mesurer son efficacité dans l'enseignement des compétences mathématiques et linguistiques. Ce score, généré par un algorithme, comptait pour moitié de son évaluation globale et l'emportait sur les appréciations positives de l'administration et de la collectivité. Les administrateurs du district scolaire n'eurent d'autre choix que de la renvoyer, et avec elle 205 autres enseignants qui avaient reçu un score IMPACT inférieur au seuil minimal.

Cela ne ressemblait ni à une chasse aux sorcières, ni à un règlement de comptes. La démarche du district de Washington relevait même d'une certaine logique. Après tout, les administrateurs d'une école pouvaient très bien apprécier des enseignants lamentables, admirant davantage leur façon d'être ou leur prétendu dévouement. De mauvais enseignants pouvaient *sembler* performants. À l'instar de nombreux autres districts, celui de Washington allait donc réduire au minimum ces critères d'appréciation humains – et donc subjectifs – et accorder plus d'attention à des scores fondés sur les résultats concrets des élèves : leur réussite en maths et en lecture. Les chiffres parleraient d'eux-mêmes, promettaient les responsables de l'académie. Ils seraient plus justes.

Bien entendu, Wysocki les trouva au contraire horriblement injustes, et voulut savoir d'où ils provenaient. « Je pense que personne n'y comprenait rien », m'a-t-elle avoué plus tard. Comment un bon enseignant pouvait-il obtenir des scores aussi déplorables ? Que mesurait vraiment ce modèle de valeur ajoutée ?

Elle finit par apprendre que c'était assez compliqué. Le district de Washington avait engagé le cabinet de conseil Mathematica Policy Research, de Princeton, pour concevoir le système d'évaluation. Le défi pour Mathematica consistait à mesurer les progrès des élèves de l'académie en question, puis à calculer quelle part de leur amélioration ou de leur



régression était imputable aux enseignants. Naturellement, c'était tout sauf facile. Les chercheurs savaient bien que de nombreuses variables, depuis le milieu socioéconomique jusqu'aux conséquences des troubles d'apprentissage, pouvaient affecter les résultats des élèves. Les algorithmes devaient prendre en compte ce type de paramètres différentiels, d'où leur grande complexité.

Réduire le comportement, les performances et le potentiel des êtres humains à des algorithmes n'est certes pas une mince affaire. Pour comprendre les difficultés auxquelles se heurtait Mathematica, imaginez une fillette de 10 ans habitant un quartier pauvre du sud-est de Washington D.C. Une fois l'année scolaire terminée, elle passe le test standard de fin de primaire. Et puis la vie suit son cours. Elle rencontre peut-être des problèmes dans sa famille ou des soucis d'argent. Elle est peut-être amenée à changer sans arrêt de logement, ou s'inquiète pour un grand frère aux prises avec la justice. Peut-être aussi est-elle complexée par son poids ou terrorisée par un ou une camarade de classe qui la brutalise. Quoi qu'il en soit, elle passe l'année suivante un autre test standard, conçu cette fois-ci pour les élèves de sixième.

Si l'on compare les résultats des deux tests d'un élève, les scores devraient normalement rester stables ou, avec un peu de chance, s'améliorer. Si à l'inverse les résultats de la fillette dégringolent, l'écart entre sa performance et celle des élèves qui ont réussi est d'autant plus flagrant.

Mais dans quelle mesure cet écart est-il dû à l'enseignant ? Difficile de le savoir, et les modèles utilisés par Mathematica n'ont que peu de chiffres pour affiner la comparaison. Dans les entreprises du Big Data comme Google, en revanche, les chercheurs réalisent en permanence quantité de tests et suivent des milliers de variables. Ils peuvent faire passer du bleu au rouge la police de caractères d'une publicité, soumettre chaque version à 10 millions de personnes et observer celle qui suscite le plus grand nombre de clics. Cette remontée d'information leur permet d'affiner leurs algorithmes et d'en ajuster le fonctionnement. Même si j'ai beaucoup de reproches à adresser à Google – nous y reviendrons par la suite –, ce type de test fait un usage efficace des statistiques.

Vouloir calculer l'impact qu'une personne peut exercer sur une autre au cours d'une année scolaire se révèle bien plus complexe. « Il y a tellement de facteurs en jeu dans l'apprentissage et l'enseignement qu'il serait très

difficile de tous les mesurer », explique Wysocki. D'un point de vue statistique, chercher à noter l'efficacité d'un enseignant en analysant les résultats aux tests de vingt-cinq à trente élèves à peine apparaît en outre assez douteux, voire ridicule. Au vu de la quantité de choses qui sont susceptibles de mal se passer au cours de l'année, l'échantillon s'avère beaucoup trop restreint. Si l'on devait analyser les enseignants avec la rigueur statistique d'un moteur de recherche, il faudrait en réalité les tester sur des milliers ou même des millions d'élèves sélectionnés de façon aléatoire. Les statisticiens comptent sur le nombre pour compenser les exceptions et les anomalies (et les ADM, comme nous le verrons, ont précisément tendance à punir les individus qui font exception).

Autre facteur crucial de réussite, les systèmes statistiques ont besoin d'un « feed-back », d'une remontée d'information qui puisse les alerter lorsqu'ils s'égarent : en somme, les statisticiens se servent des erreurs pour éprouver leurs modèles et les rendre plus performants. Si Amazon, du fait d'une corrélation erronée, se mettait à conseiller aux adolescentes des ouvrages sur l'entretien des pelouses, les clics s'effondreraient et on modifierait alors l'algorithme jusqu'à ce qu'il tombe juste. Faute de feed-back, un moteur statistique peut au contraire continuer de produire des analyses erronées et préjudiciables, sans jamais apprendre de ses erreurs<sup>3</sup>.

Nombre des ADM dont je parlerai dans ce livre, y compris le modèle de valeur ajoutée en vigueur dans le district de Washington, se comportent de la sorte. Elles définissent leur propre réalité et l'utilisent ensuite pour justifier leurs résultats. Ce type de modèle s'autojustifie et s'autoperpétue ; il est extrêmement destructeur, mais surtout très courant.

Quand le système d'évaluation de Mathematica désigne Sarah Wysocki et 205 autres enseignants comme incompetents, ceux-ci sont congédiés. Mais comment les administrateurs du district de Washington peuvent-ils savoir s'ils ont bien fait ? En réalité, c'est impossible. Le système lui-même a déterminé que ces personnes étaient incompetentes, c'est donc comme cela qu'elles sont considérées. 206 « mauvais » enseignants écartés. Ce simple chiffre semble à lui seul démontrer l'efficacité du modèle mesurant la valeur ajoutée. Il purge le district scolaire de ses enseignants médiocres. Au lieu de rechercher la vérité, le score en vient à l'incarner, et à l'imposer.

Nous avons là l'un des exemples de cercle vicieux d'une ADM. Nous verrons beaucoup d'autres de ces effets boomerang tout au long de ce livre.

Les employeurs utilisent par exemple de plus en plus les scores de crédit<sup>4</sup> pour évaluer de potentielles recrues – le raisonnement étant que les candidats qui règlent leurs factures sans délai seront davantage enclins à arriver à l’heure au travail et à respecter les règles. Or, il existe en réalité quantité de gens responsables et d’excellents travailleurs qui subissent des coups durs et voient leur score de crédit chuter. Mais la corrélation présumée entre une mauvaise note de solvabilité et de mauvaises performances professionnelles laisse moins de chances à ceux qui ont un score faible de trouver du travail. Le chômage les pousse vers la pauvreté, ce qui dégrade encore leur score et rend d’autant plus difficile l’obtention d’un emploi. C’est une spirale infernale. Et les employeurs ne sauront jamais combien d’employés de qualité ils ont laissé échapper en se focalisant ainsi sur les scores de crédit. À cause des ADM, beaucoup d’hypothèses dommageables sont camouflées par les mathématiques et ne sont ni vérifiées en pratique, ni remises en cause.

Voilà qui met en évidence une autre caractéristique des ADM : leur tendance à punir les pauvres. Cela est dû notamment au fait qu’elles sont conçues pour évaluer de vastes quantités de personnes. Ces modèles se spécialisent dans la gestion des masses et ne coûtent pas cher : c’est ce qui fait une partie de leur attrait. Les riches, à l’inverse, bénéficient souvent d’un traitement de faveur, puisqu’il est personnalisé. Un cabinet d’avocats prestigieux ou un collège privé élitiste auront bien plus recours à des recommandations personnelles et à des entretiens en tête à tête qu’une chaîne de restauration rapide ou qu’un district scolaire en milieu urbain à court d’argent. Les privilégiés, nous le constaterons maintes fois, sont plus volontiers pris en charge par des êtres humains, et le bas peuple, par des machines.

L’impossibilité pour Wysocki de trouver quelqu’un capable d’expliquer son score exécration est également révélatrice. Les verdicts des ADM tombent comme des diktats énoncés par les démiurges algorithmiques. Le modèle lui-même est une boîte noire, il renferme un secret industriel farouchement gardé, et des consultants comme Mathematica en profitent pour facturer plus cher leurs services. Cela sert aussi un autre but : si on laisse les personnes évaluées dans l’ignorance, l’idée est qu’elles chercheront moins à déjouer le système. Elles devront simplement travailler dur, suivre les règles, et prier pour que le modèle enregistre et tienne

compte de leurs efforts. Quand les informations sont cachées, il est aussi plus difficile de mettre en doute un score, voire de le contester.

Pendant des années, les enseignants de Washington se plaignirent de l'aspect arbitraire des scores et réclamèrent des éclaircissements sur les paramètres de leur calcul. « C'est un algorithme, c'est très complexe » fut la seule réponse qu'ils obtinrent. Beaucoup n'eurent pas le courage d'insister. Les mathématiques intimident, hélas, un grand nombre de gens. Une professeure de mathématiques du nom de Sarah Bax continua cependant d'assaillir le recteur du district, Jason Kamras, un de ses anciens collègues, pour obtenir les renseignements demandés. Après plusieurs mois de tergiversations, Kamras la pria d'attendre la publication imminente d'un rapport technique. « Comment justifiez-vous d'évaluer les gens selon un barème que vous êtes incapable d'expliquer ? », lui rétorqua Bax. Mais telle est la nature des ADM. On sous-traite l'analyse à des codeurs et à des statisticiens. Et la règle en vigueur est de laisser parler les machines.

Sarah Wysocki avait bien conscience que les résultats de ses élèves lors des tests devaient peser lourdement dans la balance de l'algorithme en question. C'est cela qui finit par lui mettre la puce à l'oreille. Avant d'entamer ce qui devait être sa dernière rentrée à la MacFarland Middle School, elle s'aperçut avec plaisir que ses futurs CM2 avaient étonnamment bien réussi leur test de fin d'année de CM1. Au sein de la Barnard Elementary School, dont sortaient beaucoup d'entre eux, 29 % des enfants avaient été classés « niveau supérieur » en lecture. Soit cinq fois plus que la moyenne de ce district scolaire.

Quand les cours débutèrent, Sarah constata cependant qu'un grand nombre de ses élèves avaient du mal à lire, même lorsqu'il s'agissait de phrases simples. Bien plus tard, des enquêtes menées par le *Washington Post* et *USA Today* révélèrent que les tests standard de quarante et une écoles du district, y compris Barnard, comportaient une importante proportion de réponses raturées. Un taux élevé de rectifications indique souvent une plus grande probabilité de tricherie. Dans certaines écoles, jusqu'à 70 % des classes étaient suspectées.

Quel rapport avec les ADM ? Plusieurs éléments. D'une part, les algorithmes d'évaluation des enseignants constituent un puissant outil de modification comportementale. Ils ont été conçus pour cela et font office à la fois de carotte et de bâton dans les écoles de Washington. Si leurs élèves trébuchaient lors du test, les enseignants savaient que leur emploi était

menacé. D'où une forte motivation pour s'assurer de leur réussite, surtout au moment où la crise économique frappait si durement le marché du travail. En revanche, si leurs élèves obtenaient de meilleurs résultats que les autres, les enseignants et membres de l'administration pouvaient recevoir une prime allant jusqu'à 8 000 dollars. Ces puissantes incitations s'ajoutant aux preuves réunies – le nombre important de réponses raturées et les scores anormalement élevés –, on était tenté de soupçonner que les enseignants de CM1 – soit par peur, soit par cupidité – avaient bel et bien falsifié les copies de leurs élèves.

Il est donc tout à fait concevable que ceux de Sarah Wysocki aient débuté l'année scolaire avec des scores artificiellement gonflés. Dans ce cas, leurs résultats l'année suivante montreraient une chute de la progression en CM2 – et donc que leur institutrice n'était pas assez performante. Voilà ce qui lui était arrivé, Sarah en était persuadée. Cette explication cadrait avec les observations des parents, de ses collègues et de son directeur, qui témoignaient de ses réelles qualités d'enseignante. Elle permettait de démêler l'imbroglio. Sarah Wysocki avait de solides arguments à faire valoir.

Mais impossible de faire appel contre un modèle mathématique du type ADM. C'est aussi cela, leur terrifiant pouvoir. Ils n'écoutent pas. Ils ne plient pas. Ils restent non seulement insensibles au charme, aux menaces et aux flatteries mais aussi à toute logique – même lorsqu'on a de bonnes raisons de mettre en doute les données qui alimentent leurs conclusions. Certes, s'il devient évident que des systèmes automatisés se trompent de manière embarrassante et systématique, les programmeurs se remettront à l'ouvrage et ajusteront les algorithmes. Mais la plupart du temps, les programmes livrent des verdicts inébranlables et les êtres humains qui les utilisent ne peuvent que hausser les épaules, en se disant justement : « Que peut-on y faire ? »

Et c'est précisément la réponse que Sarah Wysocki finit par obtenir de la part de son district scolaire. Jason Kamras expliqua plus tard au *Washington Post* que les ratures « suggéraient » l'existence d'un problème et qu'il se pouvait que les chiffres au sein de sa classe de CM2 aient été erronés. Mais les preuves n'étaient pas concluantes. Il affirma que Wysocki avait été traitée de façon équitable.

Voyez-vous le paradoxe ? Un algorithme analyse une multitude de statistiques et en tire la probabilité que telle ou telle personne pourrait être

une mauvaise recrue, un emprunteur à risque, un terroriste ou un enseignant déplorable. De cette probabilité, on extrait un score qui peut bouleverser l'existence d'un individu. Or, quand l'individu en question cherche à se défendre, les indices qu'il présente « suggérant » un problème ne suffisent pas. Ses arguments doivent être imparables. Alors que l'on n'exige pas de justifications de la part des ADM lorsque leur verdict tombe, les individus qui en sont victimes doivent en revanche fournir des preuves infaillibles de leur dysfonctionnement : le niveau d'exigence est donc très déséquilibré.

Passé le choc de son licenciement, Sarah Wysocki ne resta pas plus de quelques jours sans emploi. Quantité de gens, à commencer par son directeur, pouvaient la recommander, et elle décrocha très vite un poste dans une académie aisée du nord de la Virginie. Du fait d'un modèle extrêmement discutable, une école pauvre avait ainsi perdu une bonne enseignante, au profit d'une école riche qui ne renvoyait pas les siens en fonction des scores de leurs élèves.

\*\*\*

À la suite de la crise des *subprimes*, et de l'effondrement de l'immobilier qui en découla, je pris conscience de la prolifération des ADM dans le milieu bancaire et du danger qu'elles faisaient peser sur notre économie. Début 2011, je quittai mon poste au sein du fonds spéculatif qui m'employait. Plus tard, après m'être reconvertie comme experte en données, je rejoignis une jeune pousse du commerce en ligne. Je bénéficiai là d'un point de vue privilégié et constatai que des légions d'autres ADM étaient à l'œuvre dans tous les secteurs d'activité imaginables, et que nombre d'entre elles exacerbaient les inégalités et punissaient les individus. Elles étaient au cœur d'une économie des données totalement débridée.

Pour informer le public, je lançai un blog baptisé Math-Babe. J'avais comme objectif de mobiliser mes collègues mathématiciens contre l'utilisation de statistiques approximatives et de modèles biaisés, qui généraient leurs propres boucles de rétroaction aux effets délétères. Les experts en données furent tout particulièrement attirés par mon blog, et m'avertirent de la propagation des ADM à de nouveaux domaines. Mais au milieu de l'année 2011, avec la naissance du mouvement Occupy Wall Street dans le sud de Manhattan, je compris que nous avions un travail à mener auprès du grand public. Des milliers d'individus s'étaient réunis pour

réclamer justice et prise de responsabilité sur le plan économique. Or, lorsqu'ils étaient interrogés, les « occupants » semblaient souvent tout ignorer de certaines problématiques fondamentales liées au monde de la finance. Ils n'avaient de toute évidence jamais lu mon blog (même si je me dois d'ajouter qu'il n'est pas nécessaire de comprendre tous les détails d'un système pour savoir qu'il a échoué).

Il me fallait faire un choix : soit les critiquer, soit me joindre à eux. Je décidai de les rejoindre, et me retrouvai très vite à animer les réunions hebdomadaires de l'Alternative Banking Group à l'université Columbia (New York), où nous débattions de réformes financières. Ce faisant, j'en vins à réaliser que mes deux expériences hors du milieu universitaire, l'une dans la finance, l'autre dans la science des données, m'avaient offert une fabuleuse occasion d'accéder aux technologies et à la culture qui soutiennent les ADM.

Des modèles mathématiques mal conçus contrôlent aujourd'hui les moindres aspects de l'économie, depuis la publicité jusqu'à la gestion des établissements pénitentiaires. Ces ADM partagent de nombreuses caractéristiques avec le modèle de valeur ajoutée qui a fait dérailler la carrière de Sarah Wysocki dans les écoles publiques de Washington. Opaques, jamais remises en cause et comptables de rien, elles opèrent sur une échelle qui leur permet de trier, de cibler ou d'« optimiser » des millions de personnes. En confondant leurs conclusions avec la réalité du terrain, la plupart génèrent des boucles de rétroaction néfastes.

Il existe néanmoins une distinction importante entre le modèle de valeur ajoutée utilisé par un district scolaire et une ADM qui cible des candidats potentiels pour un prêt sur salaire à taux exorbitant. L'un et l'autre génèrent un gain différent. Pour le district scolaire en question, c'est une sorte de monnaie d'échange politique, l'impression de résoudre des problèmes. Pour les entreprises, c'est un gain beaucoup plus ordinaire : l'argent. Beaucoup de compagnies qui emploient ces algorithmes véreux voient dans l'afflux de trésorerie la preuve que leurs modèles fonctionnent. Il suffit de se mettre à leur place pour le comprendre. Quand elles bâtissent des systèmes statistiques pour trouver des clients ou manipuler des emprunteurs désespérés, l'augmentation de leurs bénéfices semble indiquer qu'elles sont sur la bonne voie, et que le logiciel remplit donc sa tâche. Le problème vient de ce que les profits concrets finissent par suppléer ou générer leur

propre vérité. Nous verrons maintes fois resurgir cette dangereuse confusion.

Cela se produit particulièrement lorsque les experts en données finissent par perdre de vue ceux qui se trouvent à l'autre extrémité de la transaction. Ils admettent qu'un programme traitant de grandes quantités d'informations est condamné à se méprendre sur un certain pourcentage d'individus, à les classer dans les mauvais groupes et à leur refuser un emploi ou la chance d'acquérir la maison de leurs rêves. Mais en règle générale, les personnes qui manient les ADM ne s'attardent pas sur ces erreurs. Leur « feedback » à elles, c'est le gain financier, qui est aussi leur motivation. Leurs systèmes sont conçus pour englober toujours plus de données, affiner leurs outils d'analyse et rapporter ainsi toujours plus d'argent. Les investisseurs se repaissent bien entendu de ces profits et financent très généreusement les entreprises qui conçoivent les ADM.

*Quid des victimes ?* Un expert travaillant dans l'une de ces entreprises pourrait vous répondre qu'aucun système statistique ne saurait être parfait. Ce sont des dommages collatéraux. Et souvent, à l'image de Sarah Wysocki, on estime qu'elles sont sans valeur et peuvent être sacrifiées. Oubliez-les un instant, vous dirait-il peut-être, et concentrez-vous plutôt sur tous ces gens qui obtiennent des suggestions utiles grâce aux systèmes de recommandation sur Internet, ou qui trouvent la musique qu'ils aiment sur Deezer, l'emploi idéal sur LinkedIn, voire l'amour de leur vie sur Meetic. Songez donc à quelle échelle stupéfiante on se situe, et ne tenez pas compte des imperfections.

Le Big Data possède beaucoup d'apôtres, mais je n'en fais pas partie. Ce livre s'orientera résolument dans la direction opposée, et se focalisera sur les dégâts infligés par les ADM et les injustices qu'elles perpétuent. Nous étudierons des exemples funestes touchant à des moments critiques de l'existence : entrer à l'université, emprunter de l'argent, purger une peine de prison, ou trouver et conserver un emploi. Tous ces aspects de la vie sont de plus en plus contrôlés par des modèles tenus secrets, qui assènent des punitions arbitraires.

Bienvenue dans le côté obscur du Big Data.

---

1. « Big Data » signifie littéralement « mégadonnées ». Né en 1997, le terme désigne la quantité gigantesque de données générées chaque jour par des utilisateurs du monde entier, et qui a suscité une problématique nouvelle ces dernières années : le stockage, la gestion et la protection des informations privées (*toutes les notes sont de l'éditeur, sauf indication contraire*).



2. Le *value-added model* est appliqué à l'enseignement aux États-Unis. En comparant les résultats scolaires d'un groupe d'élèves d'un même niveau une année sur l'autre, on évalue les progrès d'un professeur dans l'enseignement d'une matière donnée.
3. Phénomène que l'auteur appelle « *feedback loop* », ou « boucle de rétroaction ».
4. Aux États-Unis, les scores de crédit sont attribués aux sociétés ou individus en fonction de leur historique de crédit : rapidité des remboursements, dépassement des plafonds, nombre de crédits demandés, etc.

# LES MÉCANISMES

## Qu'est-ce qu'un modèle ?

# D'UNE BOMBE

1001Ebooks

C'était en 1946, par un chaud après-midi d'août. Lou Boudreau, joueur-entraîneur de l'équipe de base-ball des Indians de Cleveland, passait une journée épouvantable. Dans le premier des deux matchs à l'affiche, Ted Williams, peut-être à l'époque le plus grand batteur, avait presque à lui tout seul anéanti son équipe. Il avait réussi trois *home runs*<sup>1</sup> et rapporté huit points. Les Indians s'étaient finalement inclinés onze à dix.

Boudreau devait réagir. Lors du second match, quand Williams se présenta au poste de batteur pour la première fois, les joueurs des Indians se positionnèrent. Boudreau, habituellement au poste d'arrêt-court, alla prendre la place traditionnellement occupée par le deuxième base, et le deuxième base se positionna en retrait dans le champ intérieur droit. Le troisième base se décala sur sa gauche, dans la zone de l'arrêt-court. Boudreau, peut-être en désespoir de cause, faisait manifestement pivoter toute sa défense pour tenter de faire sortir les tirs impeccables de Ted Williams.

En d'autres termes, il réfléchissait à la manière d'un expert en données. Il avait analysé les paramètres bruts du problème, tirés pour l'essentiel de son observation : Ted Williams frappait habituellement sa balle en direction du champ droit. Il s'était donc adapté... et la manœuvre avait fonctionné. Ses défenseurs interceptèrent ensuite bien plus souvent les foudroyants coups en flèche de Williams (même s'ils ne pouvaient rien contre les *home runs* qui leur filaient au-dessus la tête).

Si vous assistez aujourd'hui à un match de la Major League de base-ball, vous constaterez que la stratégie de défense est désormais d'aborder pratiquement tous les joueurs de la même façon que Ted Williams. Alors

que Boudreau s'était contenté d'observer dans quelle direction Williams avait coutume de frapper, les entraîneurs savent maintenant précisément à quel endroit chaque joueur a envoyé chacune de ses balles au cours de la dernière semaine, du dernier mois, tout au long de sa carrière, face à des lanceurs gauchers, lorsqu'il a déjà manqué deux fois auparavant, et ainsi de suite. Ces historiques leur permettent d'analyser la situation et de calculer le positionnement correspondant à la plus grande probabilité de réussite. Ce qui suppose parfois de déplacer les joueurs très loin de leur position initiale sur le terrain.

Modifier sa défense n'est que l'une des facettes d'une problématique beaucoup plus vaste : quelles mesures les équipes peuvent-elles prendre pour avoir le maximum de chances de gagner ? Dans leur quête de réponses, les statisticiens du base-ball ont examiné chaque variable susceptible d'être quantifiée et lui ont attribué une valeur. Combien vaut un double comparé à un simple ? À quel moment est-il intéressant – si ça l'est jamais – d'amortir un tir pour faire avancer un joueur de la première à la deuxième base ?

Les réponses à toutes ces questions sont assemblées et combinées sous forme de modèles mathématiques. Ce sont, pour le monde du base-ball, des univers parallèles qui dessinent chacun un canevas complexe de probabilités. Ils intègrent toutes les relations mesurables entre chaque élément de ce sport, des différents types de buts jusqu'aux joueurs eux-mêmes. L'objectif est de dérouler différents scénarios pour chaque phase critique du jeu, à la recherche des combinaisons optimales. Si les Yankees remplacent leur lanceur actuel par un droitier, quelle probabilité supplémentaire ont-ils d'éliminer le puissant batteur des Angels, Mike Trout ? Et quel en sera l'impact global sur leurs chances de victoire ?

Le base-ball constitue un espace idéal pour la modélisation mathématique prédictive. Comme l'a écrit Michael Lewis dans son best-seller de 2003, *Moneyball*, ce sport a de tout temps attiré les passionnés de données. Il y a de cela plusieurs dizaines d'années, les fans se plongeaient dans les statistiques inscrites au dos des cartes à collectionner, analysant la fréquence des *home runs* de Carl Yastrzemski ou comparant le nombre total de *strikeouts*<sup>2</sup> de Roger Clemens et de Dwight Gooden. Mais à partir des années 1980, des statisticiens sérieux ont pris l'initiative d'étudier ce que ces chiffres, et avec eux toute une avalanche d'informations nouvelles, signifiaient vraiment : comment ils se traduisaient en victoires, et comment

les dirigeants d'équipes pouvaient obtenir le maximum de réussite en dépensant le moins possible.

Le terme « *moneyball* » sert désormais de raccourci pour désigner toute démarche statistique dans des domaines longtemps régis par l'instinct. Mais le base-ball demeure une étude de cas idéale – et fournit un bon exemple d'alternative aux modèles toxiques, ou ADM, qui envahissent tant d'aspects de notre existence. L'honnêteté des modèles utilisés dans le base-ball résulte en partie de leur transparence. N'importe qui peut accéder aux statistiques et comprendre plus ou moins comment elles sont interprétées. Le modèle d'une équipe pourra certes accorder plus de valeur aux auteurs de *home runs*, tandis qu'une autre leur en attribuera un peu moins, car les batteurs puissants ont souvent tendance à se faire éliminer. Mais dans les deux cas, le nombre de *home runs* et de *strikeouts* est visible de tous.

Le base-ball a également pour lui sa rigueur statistique. Ses gourous disposent d'un immense volume de données, dont la quasi-totalité concerne directement les performances des joueurs. Elles s'avèrent en outre extrêmement pertinentes en regard des résultats qu'ils s'efforcent de prédire. Aussi évident que cela puisse paraître, nous verrons tout au long de ce livre que les créateurs d'ADM manquent en revanche souvent d'informations relatives aux comportements qui les intéressent directement. Ils leur substituent alors des données indirectes, ou supplétives<sup>3</sup>. Ils établissent des corrélations statistiques entre le code postal ou les caractéristiques linguistiques d'un individu, et la probabilité qu'il rembourse un emprunt ou fasse correctement son travail. Ces corrélations sont discriminatoires, et pour certaines illégales. La plupart du temps, les modèles utilisés dans le base-ball n'emploient pas de données supplétives ; ils exploitent des renseignements appropriés comme le nombre de balles, le nombre de renvois de balle ratés, ou réussis.

L'élément le plus déterminant est qu'il arrive de nouvelles statistiques chaque année, issues en moyenne de douze à treize matchs quotidiens entre avril et octobre. Les statisticiens peuvent comparer les résultats de ces rencontres aux prédictions de leurs modèles, et voir où ils se sont trompés. Peut-être avaient-ils prédit qu'un lanceur de relève gaucher concéderait beaucoup de coups sûrs aux batteurs droitiers – alors qu'il les a au contraire écrasés. Il leur faut dans ce cas ajuster leur modèle et rechercher également d'où vient leur erreur. Son nouveau lancer « en tire-bouchon » a-t-il eu un impact sur ses statistiques ? Lance-t-il mieux en nocturne ? Quels que

soient les enseignements qu'ils en tirent, ils peuvent réinjecter ces informations dans le modèle et ainsi l'affiner. C'est de cette manière que les modèles dignes de confiance fonctionnent. Ils entretiennent un va-et-vient permanent avec les réalités de tout ordre qu'ils essaient de comprendre ou de prédire. Quand les conditions changent, les modèles doivent eux aussi changer.

L'idée même d'envisager une comparaison entre le modèle du base-ball, avec ses milliers de variables en constante évolution, et celui utilisé pour évaluer les enseignants des écoles de Washington semble étrange. Le premier modélise toute une discipline sportive dans ses moindres détails, en procédant à des mises à jour continuelles. Le second, empreint de mystère, semble s'appuyer majoritairement sur une poignée de résultats aux tests de passage dans les classes supérieures d'une année sur l'autre. Peut-on réellement parler de modèle ?

La réponse est oui. Un modèle n'est après tout rien de plus que la représentation abstraite d'un processus quelconque, qu'il s'agisse d'un match de base-ball, de la chaîne d'approvisionnement d'une compagnie pétrolière, des actions menées par un gouvernement étranger, ou de la fréquentation d'une salle de cinéma. Qu'il soit mis en œuvre dans un programme informatique ou dans notre cerveau, le modèle exploite ce que nous savons pour prédire des réponses à diverses situations. Nous avons tous dans la tête des milliers de modèles. Ils nous disent à quoi s'attendre et orientent nos décisions<sup>4</sup>.

Voici l'exemple d'un modèle informel que j'utilise au quotidien. Mère de trois enfants, je m'occupe de la cuisine – mon mari, pauvre de lui, est incapable de se rappeler qu'il faut mettre du sel dans l'eau des pâtes. Tous les soirs, quand je m'attelle à la préparation du repas familial, je modélise intérieurement et intuitivement l'appétit de chacun. Je sais que l'un de mes fils adore le poulet (mais déteste les hamburgers), alors qu'un autre mangera uniquement les pâtes (avec un supplément de parmesan râpé). Mais je dois aussi considérer que l'appétit varie d'un jour sur l'autre, et que mon modèle risque donc d'être pris au dépourvu en cas de changement. Une certaine part d'incertitude reste inévitable.

Pour alimenter mon modèle de préparation des repas, j'utilise les informations dont je dispose à propos de ma famille, les ingrédients que j'ai sous la main ou que je sais pouvoir obtenir, ainsi que l'énergie, le temps et

mon degré d'ambition. Le résultat est donc ce que je décide de préparer, et comment. J'évalue ma réussite à l'aune de la satisfaction que manifestent mes proches une fois le repas terminé, des quantités qu'ils ont mangées, et de la qualité diététique des aliments. Observer comment le repas est accueilli et à quel point il est apprécié me permet de mettre à jour mon modèle pour la fois suivante. Ces mises à jour et ces ajustements en font ce que les statisticiens appellent un « modèle dynamique ».

Au fil des années, je dois dire non sans fierté que je suis devenue assez douée pour la préparation de nos repas familiaux. Mais que se passerait-il si nous nous absentions, mon mari et moi, pendant une semaine, et que je devais expliquer mon système à ma mère pour qu'elle me remplace ? Ou si une amie, qui a elle aussi des enfants, souhaitait connaître mes méthodes ? C'est à ce moment-là que j'entreprendrais de formaliser mon modèle, en lui conférant un caractère beaucoup plus systématique et, d'une certaine façon, mathématique. Et si j'en avais le courage, je pourrais même le coder dans un logiciel.

Ce programme contiendrait idéalement tous les choix d'aliments possibles, leur valeur nutritionnelle et leur coût, ainsi qu'une base de données complète des goûts de ma famille, avec les préférences et les aversions de chacun. J'aurais néanmoins du mal à rassembler de mémoire la totalité de ces informations. J'ai des tonnes de souvenirs, dans lesquels je vois quelqu'un reprendre deux fois des asperges ou éviter soigneusement les haricots verts. Mais ils sont tout embrouillés et difficiles à formaliser en une liste exhaustive.

La meilleure solution consisterait à améliorer le modèle au fil du temps, en saisissant tous les jours de nouvelles données sur ce que j'aurais acheté et cuisiné, et en notant les réactions de chaque membre de la famille. J'y intégrerais aussi certains paramètres ou certaines contraintes. Je pourrais par exemple limiter les fruits et légumes aux produits de saison et n'autoriser les friandises qu'au compte-gouttes, juste assez pour éviter une rébellion. J'ajouterais également un certain nombre de règles : untel apprécie la viande, un autre aime manger du pain avec les pâtes, un autre encore boit beaucoup de lait et tient absolument à tartiner de Nutella tout ce qui lui tombe sous la main.

Si je faisais de ce travail l'une de mes grandes priorités, j'obtiendrais peut-être au bout de plusieurs mois un excellent modèle. J'aurais converti le système de gestion alimentaire que j'ai dans la tête, mon modèle informel

interne, en un modèle externe formalisé. En créant ce modèle, j'étendrais mon pouvoir et mon influence sur le monde. J'aurais réussi à mettre au point une version automatisée de moi-même dont les autres pourraient se servir, même quand je ne suis pas là.

Il y aurait cependant toujours des erreurs, car les modèles sont par nature des simplifications. Aucun ne peut embrasser toute la complexité du monde réel ni toutes les nuances de la communication humaine. On laisse inévitablement de côté certaines informations utiles. Je pourrais ainsi avoir négligé d'indiquer à mon modèle que les règles sur la malbouffe sont moins strictes lors des anniversaires, ou que les carottes ont plus de succès crues que cuites.

Lorsqu'on crée un modèle, il faut tout d'abord choisir les données que l'on juge pertinentes pour l'alimenter. Ce faisant, nous simplifions donc le monde réel à une sorte de modèle réduit facile à comprendre, et dont on déduit des actions et des faits essentiels. Nous attendons de ce modèle qu'il accomplisse une seule et unique tâche et nous résignons au fait qu'il puisse agir de temps à autre comme une machine désorientée, comportant d'importants angles morts.

Ces angles morts sont parfois sans conséquence. Quand on lui demande un itinéraire, Google Maps modélise l'environnement en une succession de routes, de tunnels et de ponts. Il ne s'occupe pas des bâtiments, parce qu'ils sont sans rapport avec la tâche demandée. Quand un logiciel d'avionique guide un appareil, il modélise le vent, la vitesse de l'avion, et la piste d'atterrissage, mais laisse bien évidemment de côté les rues, les tunnels, les bâtiments ou les passants.

Les angles morts d'un modèle reflètent les jugements et les priorités de ses concepteurs. Si les choix opérés dans Google Maps ou dans un logiciel d'avionique paraissent clairs et évidents, d'autres sont bien plus problématiques. Le modèle de la valeur ajoutée utilisé dans les écoles de Washington, pour en revenir à cet exemple, évalue principalement les enseignants sur la base des scores obtenus par leurs élèves lors des tests de fin d'année, sans regarder s'ils les font participer en cours, s'ils s'emploient à améliorer certaines compétences spécifiques, s'ils savent gérer une classe, ou s'ils aident les enfants en cas de difficultés personnelles ou familiales. Ce modèle se montre excessivement simpliste, et sacrifie une perception juste et profonde des situations sur l'autel de l'efficience. Mais du point de vue de l'administration, il offre un outil efficace permettant de débusquer

des centaines d'enseignants apparemment médiocres, fût-ce au risque d'en méjuger certains.

On voit ici que les modèles, malgré leur réputation d'impartialité, sont le reflet d'une idéologie et d'objectifs bien précis. En supprimant la possibilité de consommer des friandises à chaque repas, j'ai imposé mon idéologie au modèle culinaire. On le fait sans hésiter, sans même réfléchir. Nos propres valeurs et nos propres désirs influencent nos choix, des données que nous décidons de recueillir jusqu'aux questions que nous posons. Les modèles sont un ensemble d'opinions inséré dans un système mathématique.

Dire qu'un modèle fonctionne ou pas est également affaire d'opinion. L'une des composantes clés de tout modèle, qu'il soit formalisé ou non, réside en définitive dans sa conception du succès. Il s'agit là d'un point important sur lequel nous reviendrons en explorant le sombre univers des ADM. Il faut à chaque fois se demander non seulement qui a conçu le modèle, mais aussi ce que cette personne (ou cette entreprise) cherche à accomplir. Si le gouvernement nord-coréen élaborait un modèle pour composer les repas de ma famille, par exemple, il pourrait l'optimiser afin de nous maintenir juste au-dessus du seuil de la famine, à moindre coût, et en fonction des réserves alimentaires disponibles. Nos préférences compteraient peu, voire pas du tout. À l'opposé, si la conception du modèle était confiée à mes enfants, le succès se mesurerait peut-être à la présence de crème glacée à tous les repas. Mon propre modèle s'efforce d'associer un soupçon nord-coréen de gestion des ressources, le bonheur de mes enfants, tout en incluant mes priorités en matière de santé, de commodité, de diversité et de développement durable. Il est en conséquence beaucoup plus complexe. Mais il n'en demeure pas moins le reflet de ma propre réalité. Un modèle conçu pour aujourd'hui fonctionnera en outre un peu moins bien demain. S'il n'est pas constamment mis à jour, il finira par se périmé. Les prix évoluent, au même titre que les préférences des individus. Un modèle conçu pour un enfant de 6 ans ne fonctionnera pas avec un adolescent.

C'est également vrai pour nos modèles internes. Un moment de gêne est vite arrivé lorsque des grands-parents rendent visite à leur petite-fille qu'ils n'ont pas vue depuis longtemps. La fois précédente, ils avaient recueilli des informations sur ce qu'elle savait, ce qui la faisait rire, les émissions de télé qu'elle aimait, et avaient créé (de façon inconsciente) un modèle relationnel spécifiquement adapté à cette fillette de 4 ans. En la retrouvant un an plus



tard, ils risquent quelques heures d'inconfort, car leur modèle ne sera plus à jour. *Thomas le petit train* est semble-t-il passé de mode. Il leur faudra un peu de temps pour réunir de nouvelles données à propos de l'enfant et ajuster leur modèle.

Cela ne veut pas dire qu'un modèle rudimentaire est nécessairement un mauvais modèle. Il en existe de très efficaces qui ne s'appuient que sur une seule variable. Le modèle le plus courant pour détecter un incendie dans un logement ou un bureau se borne à évaluer une seule variable : la présence de fumée. C'est en général suffisant. Les concepteurs de modèles rencontrent en revanche des problèmes – ou nous les font subir ! – lorsqu'ils appliquent aux êtres humains des schémas aussi simples que celui d'un détecteur de fumée.

Au niveau individuel, le racisme peut s'envisager comme un modèle prédictif, opérant au sein de milliards de cerveaux à travers le monde. Il repose sur des données erronées, incomplètes ou généralisées. Qu'elles soient le résultat d'expérience ou de rumeurs, ces données indiquent que certains types d'individus ont un mauvais comportement. Il en résulte la prédiction binaire que toutes les personnes de même origine se comporteront de manière identique.

Il va sans dire que les individus qui adoptent une pensée raciste ne consacrent guère de temps à dénicher des informations fiables pour éprouver leur modèle défectueux. Une fois ce modèle transformé en conviction intime, il s'enracine définitivement dans leur esprit. Il génère des hypothèses pernicieuses mais les met rarement à l'épreuve, se contentant au contraire d'alimenter le modèle en données qui semblent confirmer et renforcer les préjugés de départ. Le racisme est par conséquent le plus bâclé de tous les modèles prédictifs. Il est alimenté par des données collectées de façon approximative et par des corrélations fallacieuses, renforcé par les injustices institutionnelles et pollué par un biais de confirmation<sup>5</sup>. Curieusement, il fonctionne donc de façon assez similaire aux ADM que je décrirai dans ce livre.

\*\*\*

En 1997, Duane Buck, un Afro-Américain reconnu coupable de meurtre, fut présenté devant un jury populaire du comté de Harris, au Texas. Buck avait tué deux personnes, et les jurés devaient décider s'il serait condamné à

mort ou à la prison à perpétuité avec possibilité de libération conditionnelle. La procureure requérait la peine capitale, faisant valoir que si Buck était libéré, il risquait de tuer à nouveau.

L'avocat de Buck appela comme témoin un expert psychologue du nom de Walter Quijano, qui n'arrangea pas le moins du monde les affaires de son client. Ayant étudié les taux de récidive dans le système carcéral texan, Quijano fit référence aux origines de Buck et lors de son contre-interrogatoire, la procureure sauta sur l'occasion.

« Vous avez établi que le... le *facteur racial*, le fait d'être noir, augmentait pour diverses raisons assez complexes la dangerosité des individus. Est-ce exact ? », lui demanda-t-elle.

« Oui », répondit Quijano. La procureure insista sur ce témoignage dans son réquisitoire, et le jury condamna Buck à la peine de mort.

Trois ans plus tard, le procureur général du Texas, John Cornyn, découvrit que le psychologue avait fourni des témoignages similaires, fondés sur un argument racial, dans six autres procès ayant abouti à la peine capitale – la plupart d'entre eux alors qu'il était employé par le ministère public. Cornyn, qui serait élu sénateur en 2002, ordonna de nouvelles audiences non discriminatoires pour les sept détenus. « Il n'est pas normal, déclara-t-il dans un communiqué de presse, que l'on permette à la notion de race d'entrer en considération dans notre système pénal. [...] Les citoyens du Texas attendent et méritent un système qui garantisse à tous la même impartialité. »

Les six prisonniers concernés obtinrent de nouvelles audiences mais leur condamnation fut confirmée. Selon les conclusions du tribunal, le témoignage préjudiciable de Quijano n'avait pas été décisif. Buck ne fut jamais réentendu, du fait peut-être que la question raciale avait été soulevée par son propre témoin. Il est encore aujourd'hui dans le couloir de la mort.

Indépendamment de son évocation explicite lors d'un procès, la question raciale constitue de longue date un facteur majeur dans la détermination des peines. Une étude de l'université du Maryland a montré que dans le comté de Harris, dont fait partie la ville de Houston, les procureurs étaient trois fois plus enclins à réclamer la peine de mort pour des prévenus afro-américains, et quatre fois plus pour des Hispaniques, que pour des Blancs déclarés coupables des mêmes faits. Cette tendance n'est pas propre au Texas. D'après l'American Civil Liberties Union (Union américaine pour les libertés civiles), les peines infligées aux hommes noirs dans le système

fédéral sont près de 20 % plus longues que pour les Blancs convaincus de crimes similaires. Et bien qu'ils représentent seulement 13 % de la population, les Noirs représentent aux États-Unis 40 % de la population carcérale.

On pourrait donc penser que des outils de calcul de risque entièrement informatisés et alimentés par des données pourraient réduire l'influence des préjugés et favoriser un traitement plus équitable en matière de condamnation pénale. C'est dans cet espoir que les tribunaux de vingt-quatre États se sont tournés vers ce qu'on appelle des « modélisations du récidivisme ». Elles aident les juges à évaluer le danger posé par chaque condamné, et sont synonymes de progrès à bien des égards. Elles assurent une plus grande cohérence des peines, qui sont moins impactées par les humeurs et partis pris des magistrats. Elles permettent également de réaliser des économies, en réduisant légèrement la durée moyenne d'incarcération (l'hébergement d'un détenu coûte en moyenne 31 000 dollars par an, et deux fois plus dans des États dispendieux comme le Connecticut ou celui de New York).

La question est cependant de savoir si l'on a bien éliminé les préjugés humains ou si on les a simplement camouflés sous la technologie. Les nouveaux modèles de prédiction du récidivisme sont complexes et fondés sur les mathématiques, mais ils renferment tout un tas d'hypothèses dont certaines sont chargées d'a priori. Et tandis que les propos de Walter Quijano ont été transcrits dans les minutes du procès, permettant leur lecture et leur contestation ultérieure devant la cour, les mécanismes d'un modèle de prédiction du récidivisme sont eux dissimulés au sein d'algorithmes, uniquement intelligibles pour une petite élite.

L'un des modèles les plus prisés, baptisé LSI-R, comporte un long questionnaire que l'on fait remplir au prisonnier. L'une des questions – « De combien de condamnations avez-vous déjà fait l'objet ? » – se rapporte directement au risque de récidive. D'autres présentent elles aussi un lien évident : « Quel rôle d'autres personnes ont-elles joué dans le crime commis ? Drogues et alcool y ont-ils été pour quelque chose ? »

Mais à mesure que le questionnaire se poursuit, fouillant de plus en plus profondément dans la vie de l'individu, on imagine sans mal que les détenus issus d'un milieu privilégié répondront d'une certaine manière et ceux des quartiers défavorisés d'une autre. Demandez à un criminel ayant grandi dans une banlieue aisée « Quelle est la première fois où vous avez eu

affaire à la police ? » et il pourrait bien n'avoir aucun incident à signaler, hormis celui qui l'a mené en prison. Les jeunes hommes noirs, en revanche, auront sans doute été arrêtés des dizaines de fois, même quand ils n'avaient rien à se reprocher. Alors que les hommes blancs et d'origine latino-américaine âgés de 14 à 24 ans représentaient seulement 4,7 % de la population de la ville, une étude réalisée en 2013 par l'Union pour les libertés civiles de New York a montré qu'ils totalisaient 40,6 % des contrôles avec fouille spontanée. Plus de 90 % des individus interpellés étaient innocents. Les autres avaient peut-être consommé de l'alcool avant l'âge autorisé ou s'étaient fait prendre avec un joint de cannabis. Et contrairement à la plupart des enfants de familles fortunées, cela leur avait valu des problèmes. Si des « relations » précoces avec la police constituent un indice de récidivisme, les pauvres et les minorités ethniques semblent donc présenter bien plus de risques.

Les questions sont loin de s'arrêter là. On demande également aux détenus si leurs amis et leurs proches ont un casier judiciaire. Là encore, posez cette question à un condamné élevé au sein de la classe moyenne, et la réponse aura nettement plus de chances d'être négative. Certes, le questionnaire évite de demander l'origine ethnique des gens, puisque la loi l'interdit. Mais compte tenu de la masse de détails fournie par chaque prisonnier, la seule question illégale devient presque superflue.

Depuis son invention en 1995, des milliers de détenus ont été soumis au questionnaire LSI-R. Les statisticiens en ont exploité les résultats pour concevoir un système dans lequel les réponses fortement corrélées au récidivisme pèsent davantage et comptent pour plus de points. Une fois qu'ils ont répondu au questionnaire, les détenus sont classés selon trois niveaux de risque – élevé, moyen ou faible –, en fonction du nombre de points accumulés. Dans certains États, comme celui de Rhode Island, ces tests servent uniquement à cibler ceux qui obtiennent un score élevé, afin qu'ils suivent un programme de lutte contre la récidive durant l'incarcération. Mais ailleurs, notamment dans l'Idaho et le Colorado, les magistrats utilisent ces scores pour décider de la peine prononcée.

C'est injuste. Le questionnaire révèle dans quel environnement l'individu inculpé est venu au monde et a été éduqué, et comporte aussi des renseignements sur sa famille, son quartier et ses amis. Ces détails ne devraient pas intervenir dans une procédure criminelle, ni dans la détermination d'une peine. De fait, si un procureur cherchait à nuire à un

prévenu en mentionnant le casier judiciaire de son frère ou le taux de criminalité élevé de son quartier, un avocat digne de ce nom s'écrierait : « Objection, votre honneur ! » Et un juge sérieux la lui accorderait. C'est la base de notre système juridique. Nous sommes jugés pour ce que nous faisons, pas pour ce que nous sommes. Et bien que l'on ignore l'importance exacte attachée à ces paramètres dans le test, il n'est pas normal qu'ils aient le moindre poids.

Beaucoup feraient remarquer que les systèmes statistiques comme le LSI-R sont efficaces pour jauger du risque de récidive – ou du moins plus précis que les suppositions aléatoires d'un juge. Mais même en mettant de côté, ne serait-ce qu'une seconde, la question cruciale de l'équité, on se trouve néanmoins entraîné dans la pernicieuse boucle de rétroaction propre aux ADM. Un individu classé, de par son score, dans la catégorie « à haut risque » aura de grandes chances d'être au chômage et de venir d'un quartier où nombre de ses amis et de ses proches auront eu des démêlés avec la justice. À cause en partie de ce score élevé, il écope d'une peine plus longue et passe davantage d'années en prison entouré d'autres criminels – ce qui augmente la probabilité d'une nouvelle condamnation. On le relâche finalement au cœur du même quartier défavorisé, lesté désormais d'un casier judiciaire qui rend d'autant plus difficile l'obtention d'un emploi. S'il commet un autre crime, la modélisation du récidivisme pourra se prévaloir d'un nouveau succès. Mais c'est en réalité le modèle lui-même qui alimente un cycle malsain et qui contribue à entretenir cette réalité. Cette caractéristique constitue la signature d'une ADM<sup>6</sup>.

Nous avons analysé trois sortes de modèles dans ce chapitre. Ceux utilisés dans le base-ball sont en majorité sains. Transparents et continuellement mis à jour, leurs hypothèses aussi bien que leurs conclusions sont claires et visibles de tous. Ces modèles se nourrissent de statistiques tirées du sport en question, et non de données de remplacement. Les joueurs qui font l'objet d'une telle modélisation comprennent en outre le processus et partagent l'objectif visé : remporter les compétitions (ce qui ne signifie nullement qu'un grand nombre d'entre eux n'iront pas discuter leur évaluation au moment de signer leur contrat : « J'ai été sorti deux cents fois, d'accord, mais regardez combien de *home runs* j'ai à mon actif ! »).

De mon point de vue, il n'y a assurément rien à redire au second modèle que nous avons évoqué, à savoir mon modèle de préparation des repas familiaux. Si mes enfants devaient mettre en doute les hypothèses qui le

sous-tendent, aussi bien économiques que diététiques, je serais on ne peut plus ravie de les éclairer en argumentant. Et bien qu'il leur arrive de râler face à une assiette de légumes verts, ils admettraient sans doute, en insistant un peu, que nous partageons les mêmes objectifs de commodité, d'économie, d'alimentation saine et de qualité gustative – auxquels leurs modèles personnels, qu'ils seront libres d'élaborer quand ils commenceront à acheter leur propre nourriture, attribueraient néanmoins peut-être un poids différent.

Je me dois d'ajouter qu'il est fort peu probable que mon modèle connaisse une diffusion à grande échelle. Je vois mal une chaîne d'hypermarchés, le ministère de l'Agriculture américain ou n'importe quelle autre institution d'ampleur adopter mon application et l'imposer à des centaines de millions de personnes, à l'image de certaines des ADM dont nous allons parler. Mon modèle est au contraire inoffensif, et ce d'autant plus qu'il a peu de chances de sortir un jour de mon cerveau pour être codifié.

Il en va tout autrement de l'exemple de la récidive, qui clôt ce chapitre. Celui-ci dégage des relents familiers et néfastes. Livrons-nous donc à un rapide exercice de taxinomie des ADM, afin de voir où il se range.

Première question : même si le sujet a conscience d'être modélisé ou qu'il connaît la finalité du modèle, a-t-on affaire à un modèle opaque, voire invisible ? La plupart des détenus à qui l'on fait remplir des questionnaires obligatoires ne sont pas stupides. A minima, ils suspectent que les renseignements fournis seront utilisés à leurs dépens, pour les contrôler durant leur détention et peut-être les garder plus longtemps enfermés. Ils connaissent les règles du jeu. Mais c'est aussi le cas de l'administration pénitentiaire. Aussi passe-t-elle sous silence le véritable objectif du questionnaire LSI-R. Sinon, elle le sait bien, beaucoup de prisonniers tenteront de le déjouer en fournissant des réponses qui les feront passer pour des citoyens exemplaires le jour de leur sortie. Les détenus sont donc laissés autant que possible dans l'ignorance, et on ne leur communique pas le résultat du test : leur taux de risque de récidive.

Ils sont loin d'être les seuls. Les modèles opaques et invisibles sont légion, tandis que les modèles transparents font figure d'exception. On nous modélise en tant qu'acheteurs, en tant que (sur)consommateurs de télévision, en tant que patients, candidats à un prêt – y compris dans le cadre d'applications numériques auxquelles nous souscrivons gaiement –

mais tout cela est invisible à nos yeux. Même quand de tels modèles sont correctement utilisés, leur opacité peut provoquer un sentiment d'injustice. Si un ouvrier vous annonçait, à l'entrée d'un concert en plein air, que les dix premières rangées de sièges vous sont interdites, vous trouveriez peut-être cela insensé. En revanche, si on vous expliquait que ces dix premières rangées sont réservées aux spectateurs en fauteuil roulant, cela changerait la donne. La transparence est un élément déterminant.

Pourtant, beaucoup d'entreprises s'efforcent par tous les moyens de dissimuler les résultats de leurs modèles, voire le fait même qu'ils existent. Elles se justifient souvent en objectant que l'algorithme utilisé relève d'une « recette secrète » vitale pour leur activité. C'est leur propriété intellectuelle et elle doit être défendue, mobilisant au besoin des légions d'avocats et de lobbyistes. Dans le cas de géants du Web comme Google, Amazon et Facebook, ces algorithmes sur mesure et ajustés avec précision valent à eux seuls des centaines de milliards de dollars. Les ADM constituent par principe d'impénétrables boîtes noires. Il n'en est donc que plus ardu d'offrir une réponse définitive à la seconde question : le modèle agit-il contre l'intérêt du sujet ? En un mot, est-il déloyal ? A-t-il pour effet de nuire ou de détruire des vies ?

Sur ce plan-là, le LSI-R mérite à nouveau largement la qualification d'ADM. Ceux qui l'ont mis au point dans les années 1990 y voyaient sans doute un outil qui garantirait l'efficacité et l'impartialité de la justice pénale. Il pourrait aussi contribuer à infliger des peines plus légères aux criminels non dangereux, ce qui se traduirait par un gain de liberté pour eux et par d'énormes économies pour le contribuable américain, qui règle chaque année une facture de 70 milliards de dollars pour financer le système carcéral. Mais puisqu'il juge le détenu sur la base de détails qui seraient irrecevables devant un tribunal, ce questionnaire est injuste. Bien qu'il puisse profiter à beaucoup de monde, il est vecteur de souffrance pour d'autres.

Cette boucle de rétroaction néfaste y est pour beaucoup. Comme nous l'avons vu, les modèles utilisés pour déterminer les peines dressent le profil d'un individu en fonction de sa situation personnelle, encourageant de fait la perpétuation d'un environnement qui justifie leurs hypothèses de départ. Cette boucle destructrice se répète indéfiniment, rendant le modèle de plus en plus injuste.

La troisième question concerne la capacité ou non d'un modèle à s'étendre de manière exponentielle. Comme le dirait un statisticien : peut-il changer d'échelle ? Ce n'est pas une chicane de mathématicien : si les ADM (à l'origine de simples nuisances localisées) peuvent devenir de véritables tsunamis, capables de définir notre existence et d'en fixer les limites, c'est bien à cause de leur échelle. Comme nous l'observerons par la suite, les ADM qui se développent aujourd'hui dans les secteurs des ressources humaines, de la santé et de la banque, pour ne citer que ceux-là, sont rapidement en train d'établir des normes généralisées qui exercent une autorité proche de celle de la loi. Si l'algorithme d'une banque vous attribue par exemple le profil d'un emprunteur à haut risque, le monde entier vous traitera comme un mauvais payeur – même si cela découle d'un horrible malentendu. Et quand ce modèle change d'échelle, à l'instar de celui des scores de crédit, toute votre vie s'en trouve alors affectée – qu'il s'agisse d'obtenir un appartement, un emploi, ou une voiture pour aller de l'un à l'autre.

En termes d'échelle, la modélisation du récidivisme a un potentiel croissant. Aux États-Unis, elle est d'ores et déjà appliquée dans la majorité des États et le LSI-R en constitue l'outil le plus courant, utilisé au moins dans vingt-quatre d'entre eux. Au-delà du LSI-R, les établissements pénitentiaires abritent un marché dynamique et foisonnant pour les experts du Big Data. Le système pénal regorge de données, d'autant que le droit à la vie privée d'un détenu est encore moins protégé que celui du reste de la population. Dans les institutions pénitentiaires, connues pour leur indigence, le phénomène de la surpopulation et leur caractère souvent inefficace, coûteux et inhumain, le besoin d'améliorations est criant. Qui refuserait une solution bon marché comme celle-ci ?

Dans le paysage politique polarisé que nous connaissons aujourd'hui, la réforme pénale fait figure d'exception : c'est une question sur laquelle progressistes et conservateurs parviennent à trouver un terrain d'entente. Début 2015, les frères Charles et David Koch, conservateurs, se sont alliés avec un groupe de réflexion progressiste, le Center for American Progress, afin de promouvoir une réforme des prisons et de faire baisser la population carcérale. Mais je crains que leur effort bipartite, comme de multiples initiatives similaires, ne soit quasiment assuré de déboucher sur une solution alimentée par des données, dont tout le monde semble louer l'efficacité et la prétendue équité. Notre époque est ainsi faite. Même si



d'autres ADM supplantent le modèle du LSI-R, le système pénitentiaire a toutes les chances de représenter pour ces dernières un incubateur puissant et à grande échelle.

En résumé, nous avons là les trois éléments d'une ADM : opacité, échelle et nocivité. Tous seront présents, à des degrés divers, dans les exemples que nous allons traiter. Il y aura certes matière à ergoter. On pourrait soutenir par exemple que les scores de récidivisme ne sont pas totalement opaques, puisqu'ils produisent des évaluations dont les prisonniers peuvent dans certains cas prendre connaissance. Ils sont néanmoins pleins de zones d'ombre, car les détenus n'ont aucun moyen d'observer comment leurs réponses génèrent le score obtenu. L'algorithme d'évaluation demeure mystérieux. Quelques autres ADM pourraient donner l'impression de ne pas remplir les prérequis en matière d'échelle. Elles n'occupent pas une place majeure, du moins pour l'instant. Mais elles constituent des espèces dangereuses qui sont en bonne voie de développement, exponentielle. Je les comptabilise donc dans le lot. Vous noterez enfin peut-être que toutes ces ADM ne sont pas universellement nocives. Après tout, elles envoient des gens à Harvard, accordent à d'autres l'accès à des prêts avantageux ou à des emplois intéressants, et allègent les peines de prison de quelques criminels chanceux. Mais le problème n'est pas qu'elles bénéficient à certains. Le problème, c'est qu'elles fassent souffrir tant de monde. Ces modèles, gouvernés par des algorithmes, claquent la porte au nez de millions de personnes, souvent pour les motifs les plus minces et sans possibilité d'appel. Ils sont injustes.

Une chose encore à propos des algorithmes : ils peuvent être appliqués d'un domaine à un autre, et c'est souvent le cas. Les recherches en épidémiologie sont ainsi susceptibles d'éclairer les prédictions d'entrées dans les cinémas, tandis qu'on recycle les filtres antispam afin d'identifier le virus du Sida. C'est tout aussi vrai pour les ADM. Si les modèles mathématiques utilisés dans les prisons donnent l'impression de réussir dans leur tâche – qui se résume en réalité à gérer des individus au mieux –, ils pourraient par conséquent se propager au reste de l'économie avec d'autres ADM, dont nous serions les dommages collatéraux.

C'est précisément là où je veux en venir. Cette menace est de plus en plus concrète. Et l'univers de la finance nous en apporte un exemple édifiant.

---

1. Coup réussi au base-ball : le batteur frappe la balle de telle façon qu'un joueur de son équipe a le temps de passer par toutes les bases du terrain et revenir à son point de départ avant qu'un joueur de l'équipe adverse ait pu récupérer la balle.
2. Élimination d'un batteur à la suite de plusieurs échecs.
3. On appelle ces données de rechange, ou données supplétives, des « *proxy datas* ».
4. Ce type de modèle est appelé « modèle interne » dans l'ouvrage.
5. En psychologie, le « biais de confirmation » désigne le processus cognitif qui pousse un individu à privilégier les informations qui viennent confirmer des convictions ou hypothèses préexistantes, et rejeter celles qui pourraient les contredire.
6. L'utilisation de ces modèles de prévention de la récidive a depuis été largement critiquée, notamment via la publication par le média ProPublica d'un rapport à charge contre la systématisation de la justice prédictive.

## Comment j'ai perdu mes illusions

1001Ebooks

Imaginez qu'avant de prendre votre train de banlieue depuis Joliet jusqu'à LaSalle Street (dans le quartier financier de Chicago), vous ayez pris l'habitude de glisser chaque matin deux dollars dans un distributeur de café. La machine vous délivre un gobelet et vous rend deux pièces de vingt-cinq *cents*. Mais un beau jour, elle vous en rend quatre. Au cours du mois suivant, la même machine produit trois fois le même résultat. Un schéma récurrent est en train d'apparaître.

Si cette minuscule anomalie survenait non pas dans une gare de banlieue mais sur les marchés financiers, un ou une analyste quantitatif employé par un fonds spéculatif – quelqu'un comme moi – pourrait en faire sa cible. Il s'agirait pour lui ou elle d'étudier plusieurs années voire plusieurs décennies de données, puis d'entraîner un algorithme afin qu'il prédise cette erreur récurrente – une fluctuation de prix de cinquante *cents* – et qu'il parie dessus. Les récurrences même les plus ténues peuvent rapporter des millions au premier investisseur qui les déniche. Et elles continueront de générer des gains jusqu'à ce que l'un des deux événements suivants se produise : soit le phénomène s'arrête, soit le reste du marché finit par s'en rendre compte, et l'opportunité s'évanouit. À ce stade, un bon analyste sera déjà sur la piste de dizaines d'autres infimes irrégularités.

La quête des « inefficiences de marché », comme les appellent les analystes quantitatifs, s'apparente à une chasse au trésor. Elle peut être amusante. Et à mesure que je m'accoutumais à mon nouveau travail chez D. E. Shaw, le changement par rapport au milieu universitaire me sembla appréciable. Même si j'avais adoré enseigner à Barnard et mener des recherches sur la théorie algébrique des nombres, je m'étais aperçue que je progressais de manière atrocement lente. Je voulais faire partie du monde réel, et m'inscrire dans son rythme trépidant.

À cette époque, je tenais les fonds spéculatifs pour moralement neutres – sortes d’« écumeurs » du système financier, au pire. J’étais fière d’entrer chez Shaw, connu comme le Harvard des fonds spéculatifs, et de leur montrer que ma matière grise pouvait produire de l’argent. J’allais en outre gagner trois fois plus qu’en étant professeure. En débutant dans ce nouvel emploi, difficile de suspecter qu’il m’offrirait une place au premier rang lors de la crise financière, et que j’y apprendrais une terrifiante leçon quant au caractère parfois insidieux et destructeur des mathématiques. Au sein de ce fonds spéculatif, j’eus l’occasion pour la première fois d’observer de près une ADM.

Au départ, j’avais de quoi être comblée. Chez Shaw, les mathématiques étaient souveraines. Dans beaucoup de firmes, ce sont les traders qui mènent le bal ; ils concluent de grosses affaires, aboient leurs ordres et décrochent des primes de plusieurs millions de dollars. Les analystes sont leurs sous-fifres. Chez Shaw, en revanche, les traders sont à peine plus que des fonctionnaires. Ils portent le nom d’« exécutants ». Les mathématiciens, eux, règnent en maîtres. Je travaillais au sein du groupe « *futures* », une équipe de dix personnes, chargée des instruments à terme<sup>1</sup>. Dans une activité où tout dépend de ce qui arrivera demain, qu’y avait-il de plus important ?

Nous étions au total une cinquantaine d’analystes quantitatifs – dans un premier temps exclusivement des hommes, moi exceptée. La plupart étaient nés à l’étranger. Beaucoup venaient des mathématiques abstraites ou de la physique ; quelques-uns, comme moi, de la théorie des nombres. On ne me laissa toutefois guère l’occasion de discuter boulot avec eux. Puisque nos idées et nos algorithmes constituaient le socle même de l’activité du fonds, il était évident que nous autres analystes représentions aussi un risque : si nous quittions l’entreprise, nous pourrions rapidement utiliser notre savoir pour fournir des munitions à un concurrent déterminé.

Pour éviter des fuites à grande échelle, susceptibles de mettre la firme en péril, Shaw nous interdisait presque de parler de ce que nous faisions avec des confrères issus d’autres groupes – et parfois même avec nos propres collègues de bureau. L’information était en quelque sorte cloisonnée à l’aide d’un réseau de cellules individuelles qui n’était pas sans rappeler la structure d’Al-Qaïda. De cette façon, si jamais une cellule tombait – si l’un de nous s’en allait chez Bridgewater ou J.P. Morgan, ou se lançait à son compte –, nous n’emportions que ce que nous savions. Le fonds pouvait

poursuivre le reste de son activité sans être affecté. Comme on l'imagine, cela ne favorisait pas tellement la camaraderie.

Dans le groupe « *futures* », les nouvelles recrues se voyaient imposer une période d'astreinte toutes les treize semaines. Il fallait se tenir prêt à réagir aux dysfonctionnements informatiques pendant les heures d'ouverture de tous les marchés mondiaux, du dimanche soir (quand les places asiatiques s'éveillaient) jusqu'à la clôture de la Bourse de New York, le vendredi à 16 heures. Le manque de sommeil était un vrai problème. Mais pire encore était le manque de partage des informations, qui rendait difficile la résolution des problèmes. Si un algorithme semblait dysfonctionner, je devais le localiser puis trouver à toute heure du jour ou de la nuit la personne (c'était d'ailleurs toujours un homme) qui en était responsable, et lui demander de le corriger. Le contact n'était pas forcément amical.

Et puis il y avait les moments de panique. Durant les vacances, quand nous étions peu nombreux à travailler, des choses étranges avaient tendance à se produire. Notre vaste portefeuille contenait toutes sortes d'instruments, dont des contrats *forward* sur devises, sorte de promesse d'acquisition d'un gros montant de devise étrangère, à une échéance de quelques jours. Au lieu d'acheter réellement la devise en question, le trader « basculait » quotidiennement sa position, afin que cette promesse d'achat soit repoussée d'une journée. Nous maintenions ainsi notre pari sur l'orientation du marché sans jamais avoir à sortir de fortes sommes. Un jour, pendant les fêtes de Noël, je remarquai qu'une importante position en yens arrivait à échéance. Il fallait faire basculer ce contrat. Cette tâche revenait habituellement à un collègue qui travaillait depuis l'Europe, et qui se trouvait alors sans doute chez lui avec sa famille. Je m'aperçus que si nous ne réalisions pas très vite l'opération, quelqu'un devrait théoriquement se présenter à Tokyo avec une somme de 50 millions de yens. La résolution de ce problème agrémenta les vacances de quelques heures de frénésie.

Toutes ces difficultés pouvaient entrer dans la catégorie des aléas professionnels. Mais le vrai souci provenait du sentiment désagréable qui commençait à me nouer le ventre. Je m'étais habituée à jouer avec ces montagnes de devises, d'actions et d'obligations, avec les milliers de milliards de dollars qui circulaient sur les marchés internationaux. Or, contrairement aux nombres utilisés dans mes modèles universitaires, les chiffres qui alimentaient ceux du fonds spéculatif représentaient quelque chose de concret. Il s'agissait de l'épargne retraite et des hypothèques

d'individus bien réels. Rétrospectivement, cela paraît d'une évidence flagrante. Et je l'ai certainement toujours su, mais n'avais pas véritablement mesuré la nature de chacune des petites sommes que nous extirpions à l'aide de nos outils mathématiques. Ce n'était pas de l'argent trouvé, comme des pépites extraites d'une mine d'or ou des pièces de monnaie remontées d'un galion espagnol englouti. Cette richesse sortait de la poche des gens. Pour les fonds spéculatifs, qui sont les acteurs les plus prétentieux de Wall Street, c'était de la « *dumb money* » (de l'argent « stupide », celui des investisseurs inexpérimentés).

C'est en 2008, au moment où les marchés se sont effondrés, que l'horrible vérité m'a frappée de plein fouet. Plus grave même que de s'emparer de l'argent « stupide » des particuliers, l'industrie de la finance œuvrait à la création d'ADM, et j'y contribuais un peu.

En réalité, les ennuis avaient démarré un an plus tôt. Au mois de juillet 2007, les taux d'intérêt interbancaires étaient montés en flèche. Au lendemain de la récession qui avait suivi les attaques terroristes de 2001, la faiblesse des taux avait nourri la flambée de l'immobilier. N'importe qui, semblait-il, pouvait alors obtenir un prêt. Les promoteurs transformaient des zones périphériques, des étendues désertes et des prairies entières en de vastes ensembles de logements, pendant que les banques misaient des milliards sur toutes sortes d'instruments financiers liés à cette aubaine.

Mais la hausse des taux révélait un problème. Les banques perdaient confiance dans la capacité des autres établissements à rembourser leurs prêts au jour le jour. Elles prenaient peu à peu conscience du volume dangereux d'actifs à risque détenus dans leurs propres portefeuilles et jugeaient à raison que d'autres banques étaient assises sur un risque tout aussi important, sinon plus. On pourrait dire avec le recul que les pics des taux d'intérêt étaient en fait la marque d'un retour à la raison, même s'ils arrivaient manifestement trop tard.

Chez Shaw, ces soubresauts refroidirent quelque peu l'ambiance. Beaucoup d'entreprises allaient se retrouver en difficulté, c'était évident. Le secteur financier allait subir un choc, peut-être même d'une très grande ampleur. Pour autant, ce ne serait pas forcément notre problème. Nous ne nous étions pas rués tête baissée sur des marchés à risque. Les fonds spéculatifs portent en anglais le nom de *hedge funds*, soit littéralement « fonds à couverture<sup>2</sup> ». Et c'est en définitive ce que nous faisions : nous couvrir. C'était notre nature. À l'origine, nous avons qualifié les turbulences

du marché de « désordre ». Elles risquaient pour Shaw de provoquer quelques désagréments, voire une ou deux situations embarrassantes, comme lorsqu'un homme fortuné voit sa carte de crédit refusée dans un restaurant chic. Mais nous avions de bonnes chances de nous en sortir.

Après tout, ce n'étaient pas les fonds spéculatifs qui animaient ces marchés. Ils se contentaient, eux, de parier dessus. Quand le marché s'effondrerait, ce qui allait arriver, de belles opportunités émergeraient des décombres. Pour les fonds spéculatifs, le jeu ne consistait pas tant à profiter de la hausse des marchés qu'à prédire leurs mouvements. La baisse pouvait s'avérer tout aussi lucrative.

Pour comprendre comment les fonds spéculatifs opèrent aux marges du marché, imaginez un match des World Series de base-ball au stade Wrigley Field de Chicago. Grâce à un spectaculaire *home run* en fin de neuvième manche, les Chicago Cubs remportent leur premier championnat depuis 1908 (à l'époque où Teddy Roosevelt était président des États-Unis). Le stade explose et célèbre la victoire. Seule une rangée de supporters reste assise, analysant sereinement un ensemble de résultats. Ceux-là n'ont pas placé de traditionnels paris gagnants/perdants. Ils ont peut-être parié en revanche que les releveurs des Yankees concéderaient davantage de buts sur balles que de *strikeouts*, que le match compterait au moins un amorti mais pas plus de deux, ou que le premier lanceur des Cubs tiendrait au minimum six manches. Ils ont même parié sur le fait que d'autres parieurs gagneraient ou perdraient leurs propres paris. Ces gens-là misent sur un grand nombre d'événements liés au match, mais pas tellement sur la rencontre elle-même. Ils se comportent, en ce sens, à la manière d'un fonds spéculatif.

Nous nous sentions donc à l'abri, ou du moins plus que d'autres. Je me souviens d'un gala organisé en l'honneur des architectes de ce système qui allait bientôt s'écrouler. L'entreprise accueillait Alan Greenspan, l'ancien président de la Réserve fédérale<sup>3</sup>, et Robert Rubin, ancien secrétaire au Trésor et cadre dirigeant de la banque d'investissement Goldman Sachs. En 1999, Rubin avait fait campagne pour une révision de la loi Glass-Steagall (qui remontait à la Grande Dépression des années 1930). La cloison entre les activités bancaires et d'investissement était donc tombée, facilitant l'orgie spéculative de la décennie suivante. Les banques étaient auparavant libres de monter des prêts (pour beaucoup frauduleux) et de les vendre aux clients sous forme de titres financiers. Cela n'avait rien d'inhabituel, et pouvait même être considéré comme un service rendu à leur clientèle. Mais

avec l'abrogation de la loi Glass-Steagall, elles avaient désormais la possibilité – et le faisaient parfois – de parier contre ces titres qu'elles avaient elles-mêmes vendus aux clients. Les risques engendrés étaient astronomiques, et il en résultait pour les fonds spéculatifs un potentiel d'investissement infini. Car nous pariions en définitive sur les mouvements de marché, à la hausse comme à la baisse, et ces marchés-là connaissaient des fluctuations frénétiques.

Lors de la soirée chez D. E. Shaw, Greenspan nous alerta sur les problèmes des titres adossés à des créances hypothécaires. Ce souvenir reviendrait me hanter quand je me rendrais compte, quelques années plus tard, que Rubin (alors en poste chez Citigroup) avait joué un rôle déterminant dans l'assemblage d'un gigantesque portefeuille composé précisément de contrats toxiques de ce genre – l'une des principales raisons pour lesquelles il avait ensuite fallu renflouer Citigroup aux frais du contribuable.

Aux côtés de ces deux-là se trouvait le protégé de Rubin, et notre associé à temps partiel, Larry Summers. Il avait suivi Rubin au Trésor, avant d'être nommé président de l'université de Harvard. Mais Summers avait eu des rapports difficiles avec le corps enseignant. Les professeurs s'étaient notamment révoltés parce qu'il avait suggéré que le faible nombre de femmes en mathématiques et dans les sciences exactes venait peut-être d'une infériorité génétique – ce qu'il qualifiait de répartition inégale des « aptitudes intrinsèques ».

Après avoir quitté la présidence de Harvard, Summers avait atterri chez Shaw. Et au moment de s'adresser au prestigieux trio, je me rappelle que notre fondateur, David Shaw, avait indiqué en plaisantant qu'il s'était agi pour Summers d'une « promotion ». Les marchés avaient beau gronder, Shaw était encore le roi du monde.

À mesure que la crise s'aggravait, les associés perdirent toutefois un peu de leur arrogance. Ces marchés perturbés étaient en fin de compte tous imbriqués. Des rumeurs sur la vulnérabilité de Lehman Brothers, qui détenait 20 % de D. E. Shaw et gérait une grande partie de nos transactions, circulaient déjà. Comme les marchés continuaient d'être secoués, l'ambiance en interne devint tendue. Personne n'était plus habile que nous avec les chiffres. Mais qu'advviendrait-il si l'effrayant lendemain qui pointait à l'horizon ne ressemblait en rien à ce que nous avions connu hier ?



Si nous étions confrontés à quelque chose d'entièrement nouveau et de complètement différent ?

C'était une vraie inquiétude, car les modèles mathématiques s'appuient par nature sur le passé, et sur l'hypothèse de la répétition de certains schémas. Le groupe en charge des placements boursiers ne tarda pas à liquider ses participations, pour un coût substantiel. Et le recrutement frénétique de nouveaux analystes quantitatifs, qui m'avait conduit dans cette entreprise, s'interrompit. Les gens s'efforçaient de dédramatiser, mais la peur ne cessait de croître. Tous les regards étaient braqués sur les produits titrisés, en particulier ceux adossés à des créances hypothécaires, contre lesquels Greenspan nous avait mis en garde.

Pendant des décennies, les titres hypothécaires n'avaient jamais fait peur à personne, bien au contraire. Il s'agissait d'instruments financiers assez ennuyeux, utilisés aussi bien par les particuliers que par les fonds de placement pour diversifier leur portefeuille. L'idée sous-jacente était que la quantité permettait de compenser le risque. Chaque prêt présentait individuellement une possibilité de défaut : le propriétaire du logement pouvait se déclarer en faillite, ce qui signifiait que la banque ne serait jamais en mesure de recouvrer tout l'argent prêté. À l'extrême opposé, l'emprunteur pouvait également rembourser son prêt de manière anticipée, mettant alors un terme au flux des intérêts dus.

Dans les années 1980, les banquiers d'investissement avaient donc commencé à acquérir des milliers de prêts et à les combiner sous forme de titres mobiliers – des sortes d'obligations, autrement dit des instruments qui versaient des dividendes à intervalles réguliers, souvent tous les trimestres. Quelques propriétaires allaient bien entendu faire défaut. Mais la plupart des gens se maintiendraient à flot et continueraient à rembourser leur emprunt, générant un flux de revenu stable et prévisible. Au fil du temps, ces titres obligataires se muèrent en une véritable industrie, un pilier des marchés de capitaux. Des experts regroupaient les prêts en différentes classes, ou tranches. Certaines étaient considérées comme parfaitement sûres. D'autres comportaient davantage de risque – et offraient des taux d'intérêt supérieurs. Les investisseurs avaient de bonnes raisons d'être confiants, car les agences de notation Standard & Poor's, Moody's et Fitch avaient étudié ces produits et leur avaient attribué un score de risque. Ils constituaient selon elles des investissements raisonnables. Mais que penser de leur opacité ? Les investisseurs restaient ignorants de la qualité des

emprunts titrisés. L'unique moyen d'entrevoir ce qui se cachait à l'intérieur venait des évaluations d'analystes. Et ces analystes recevaient des commissions versées justement par les entreprises dont ils évaluaient les produits. Les titres de créances hypothécaires représentaient, cela va sans dire, un tremplin idéal pour la fraude.

S'il vous faut une métaphore, on utilise couramment dans ce domaine celle de la saucisse. Représentez-vous les emprunts comme des petits morceaux de viande de qualité variable, et les créances hypothécaires titrisées comme les lots de saucisses obtenues en mélangeant le tout, et en y ajoutant une poignée d'épices bien fortes. Ces saucisses peuvent évidemment être de qualité inégale, et il est difficile de dire (vu de l'extérieur) ce qui est entré dans leur composition. Mais puisqu'elles portent le sceau des services vétérinaires indiquant qu'elles sont bonnes à consommer, nous mettons de côté nos inquiétudes.

Durant le boom du secteur immobilier, comme le monde entier l'a appris plus tard, les sociétés de crédit hypothécaire réalisèrent d'abondants profits en octroyant des prêts pour des habitations que les gens n'avaient pas les moyens de s'acheter. La stratégie consistait simplement à émettre des prêts intenables, à empocher les commissions, puis à se défaire des valeurs mobilières résultantes – les saucisses – sur un marché des créances hypothécaires en pleine croissance. Dans un cas particulièrement célèbre, un cueilleur de fraises du nom d'Alberto Ramirez, qui gagnait 14 000 dollars par an, était parvenu à financer une maison de 720 000 dollars à Rancho Grande, en Californie. Son courtier lui avait apparemment dit qu'il pourrait renégocier le prêt au bout de quelques mois puis revendre ultérieurement le bien et en tirer un joli bénéfice. Quelques mois après, il n'était bien entendu plus en capacité d'honorer son emprunt.

Alors qu'elles pressentaient l'effondrement du marché du logement, les banques de crédit hypothécaire ne se contentaient pas de proposer des accords intenables, mais prospectaient activement, en quête de victimes, dans les quartiers pauvres et au sein des minorités ethniques. Dans une procédure intentée au niveau fédéral, la municipalité de Baltimore accusa la banque Wells Fargo d'avoir ciblé les quartiers noirs avec ce qu'on appelait des « prêts ghetto ». Selon Beth Jacobson, une de ses anciennes chargées de crédits, l'unité « marchés émergents » de la banque se focalisait sur les églises de la communauté noire. L'idée était que les pasteurs, personnes de confiance, orienteraient leurs fidèles vers ces produits. Il s'avéra qu'il était

question de prêts *subprimes*, supportant les taux d'intérêt les plus élevés. La banque vendait ces prêts y compris aux emprunteurs dotés d'une solvabilité sans faille, qui auraient dû être éligibles à de bien meilleures conditions. En 2009, lorsque Baltimore porta l'affaire en justice, plus de la moitié des propriétés faisant l'objet d'une saisie dans le cadre d'un prêt Wells Fargo étaient vides, et 71 % d'entre elles se situaient dans des quartiers majoritairement afro-américains (en 2012, Wells Fargo régla le procès à l'amiable et accepta de verser 175 millions de dollars à 30 000 victimes réparties dans tout le pays).

Soyons clair : les emprunts *subprimes* qui se sont accumulés durant la flambée du marché immobilier, qu'ils soient détenus par des cueilleurs de fraises californiens ou par des paroissiens noirs de Baltimore en grande difficulté, n'étaient pas des ADM. Il s'agissait non pas de modèles mais d'instruments financiers, et ils avaient peu de rapport avec les mathématiques (d'ailleurs, les courtiers faisaient tout leur possible pour ignorer les chiffres gênants).

Mais lorsque les banques ont commencé à titriser des emprunts comme celui d'Alberto Ramirez, à les réunir en différentes classes et à les mettre en vente, elles se sont appuyées pour ce faire sur des modèles mathématiques défectueux. Le modèle de risque associé aux créances hypothécaires était une ADM. Les banques étaient bien conscientes que certains de ces emprunts n'avaient aucune chance d'être remboursés. Mais elles s'accrochaient à deux hypothèses fausses, qui maintenaient leur confiance dans le système.

La première voulait que des mathématiciens d'élite, au sein de toutes ces entreprises, analysaient les chiffres et équilibraient le risque avec une extrême minutie. Les obligations étaient présentées comme des produits dont le niveau de risque était évalué par des spécialistes au moyen d'algorithmes de pointe. Hélas, ce n'était absolument pas le cas. À l'instar de tant d'ADM, les mathématiques servaient d'écran de fumée, dirigé contre les consommateurs. Elles n'avaient pour but que d'optimiser les profits à court terme des vendeurs. Et ces vendeurs étaient persuadés qu'ils réussiraient à liquider leurs titres avant que tout n'explose. Les gens futés sortiraient gagnants. Et ceux qui l'étaient moins – les pourvoyeurs d'argent « stupide » – hériteraient de milliards et de milliards de dettes irrécouvrables. Les mathématiciens rigoureux – et il en existait un certain nombre – travaillaient eux-mêmes à partir de chiffres fournis par des

individus qui menaient une fraude à grande échelle. Très peu de gens disposaient de l'expertise et des informations requises pour savoir ce qui se passait réellement d'un point de vue statistique, et la plupart de ceux qui savaient n'avaient pas l'intégrité nécessaire pour dénoncer la fraude à grande échelle. Les notes de risque attribuées aux titres financiers étaient conçues pour être opaques et intimidantes (sur le plan mathématique), afin notamment que les acheteurs ne puissent percevoir le véritable niveau de risque associé aux contrats qu'ils possédaient.

La seconde hypothèse erronée soutenait que les emprunteurs ne feraient pas massivement défaut de manière simultanée. Elle reposait sur une théorie, bientôt réfutée, selon laquelle les défauts de paiement constituaient des événements largement aléatoires et indépendants les uns des autres. D'où la conviction que les emprunts solides compenseraient dans chaque tranche les contrats défaillants. Les modèles de risque présupposaient que l'avenir serait une répétition du passé.

Pour vendre ces titres de créances hypothécaires, les banques avaient besoin d'une note AAA. Elles se tournèrent pour cela vers les trois agences chargées de noter la qualité de crédit. À mesure qu'il se développait, l'évaluation d'un marché des titres hypothécaires de plusieurs milliards de dollars se transforma en mine d'or pour les agences de notation, leur rapportant des commissions lucratives. Elles en devinrent dépendantes. Et si elles donnaient des notes à peine inférieures au triple A, elles comprenaient fort bien que les banques confieraient le travail à la concurrence. Elles jouèrent donc le jeu. Elles accordaient plus d'attention à la satisfaction de leurs clients qu'à l'exactitude de leurs modèles. Ces modèles de risque généraient en outre leur propre boucle de rétroaction pernicieuse. Les notes AAA attribuées à des produits déficients se transformaient en dollars. Cette valeur financière donnait à son tour confiance dans les produits et dans la façon dont ils avaient été « assemblés », en l'occurrence de façon frauduleuse. L'ensemble de cette machine sordide a fonctionné de la sorte, dans un échange cyclique de faveurs réciproques, jusqu'à ce que tout vole en éclats.

Parmi toutes les caractéristiques d'une ADM, ce qui a converti ces modèles de risque en une force monstrueuse d'envergure mondiale, c'est leur échelle. Les charlatans ont toujours existé ; lors des bulles immobilières antérieures<sup>4</sup>, des acheteurs inconscients s'étaient retrouvés en possession de terrains marécageux et d'une pile d'actes contrefaits. Mais

cette fois-ci, la puissance de l'informatique moderne avait alimenté une fraude d'une ampleur sans précédent. Les dégâts furent aggravés par le développement de vastes marchés autour des titres de créances hypothécaires : les couvertures de défaillances (*credit default swaps*, ou CDS) et les obligations synthétiques adossées à des actifs (*synthetic collateralized debt obligations*, ou CDO). Les CDS fonctionnaient comme de petits contrats d'assurance, qui transféraient le risque associé à un titre. Les *swaps* apportaient un sentiment de sécurité aussi bien aux banques qu'aux fonds spéculatifs, puisqu'ils pouvaient servir en apparence à équilibrer leurs risques. Mais si les entités détentrices de ces assurances faisaient faillite, comme ce fut le cas de beaucoup, une réaction en chaîne dévasterait alors toute l'économie mondiale. Les CDO synthétiques allaient un cran plus loin : il s'agissait de contrats dont la valeur dépendait de la performance des CDS et des titres de créances hypothécaires. Ils permettaient aux ingénieurs financiers d'accroître encore l'effet de levier de leurs paris.

Ce marché en surchauffe (et sur le point de s'effondrer) comptait 3 000 milliards de dollars d'emprunts *subprimes* en 2007, et celui qui l'entourait – englobant les CDS et les CDO synthétiques, qui amplifiaient les risques – était vingt fois plus gros. Aucune économie nationale ne pouvait s'y comparer.

Paradoxalement, les algorithmes prétendument puissants qui avaient créé le marché, ceux-là mêmes qui analysaient le risque des tranches de créances et les structuraient sous forme de titres financiers, se révélèrent incapables le moment venu de faire le ménage, et de calculer ce que valaient réellement tous ces bouts de papier. Les mathématiques pouvaient multiplier les inepties, mais pas les déchiffrer. Cette tâche revenait aux hommes. Eux seuls pouvaient passer les emprunts au crible, repérer les fausses promesses et les illusions, et attribuer aux prêts leur vraie valeur financière. C'était un processus fastidieux, car à l'inverse des ADM, les êtres humains n'ont pas les moyens d'étendre de façon exponentielle leur capacité de travail, et la plupart des entreprises du secteur n'en faisaient pas une priorité. Durant cette longue désintoxication, la valeur des créances – et des logements sur lesquelles elles reposaient – ne cessa bien entendu de chuter. Et avec le plongeon de l'économie, les propriétaires qui étaient en mesure d'honorer leur emprunt au début de la crise furent soudain menacés eux aussi d'un défaut de paiement.

Comme je l'ai indiqué, lors de l'effondrement du marché, Shaw se tenait à quelques pas de l'épicentre. Mais lorsque d'autres acteurs du secteur commencèrent à sombrer, ils se mirent à clôturer frénétiquement des transactions qui impactaient celles que nous avions en portefeuille. Par un effet de dominos, à l'entame du deuxième semestre 2008, nous perdions de l'argent de tous côtés.

Au cours des mois suivants, la catastrophe finit par toucher le grand public. C'est là que chacun s'aperçut enfin qu'il y avait des gens derrière tous ces algorithmes. Des propriétaires désespérés qui perdaient leur logement, et des millions d'Américains qui se retrouvaient sans emploi. Les défauts de paiement sur les cartes de crédit atteignirent des niveaux records. La souffrance humaine, dissimulée jusque-là derrière les chiffres, les feuilles de calcul et les scores de risque, devint palpable.

Chez Shaw, les conversations étaient empreintes d'angoisse. Après la chute de Lehman Brothers en septembre 2008, les gens envisagèrent des retombées politiques. Barack Obama semblait bien placé pour remporter l'élection en novembre. Allait-il s'en prendre au secteur financier, en instaurant de nouvelles dispositions réglementaires ? Augmenterait-il l'imposition sur l'intéressement ? Ces personnes-là n'étaient pas dépossédées de leur logement et n'avaient pas besoin de dépasser le solde de leur carte de crédit juste pour garder la tête hors de l'eau. Mais elles n'en trouvaient pas moins matière à s'inquiéter. La seule option était d'attendre que les choses se passent, de laisser les lobbyistes faire leur travail, et de voir si nous serions autorisés à continuer comme avant.

Dès 2009, il devint évident que les leçons de l'effondrement du marché n'avaient imprimé aucun changement de cap dans le monde de la finance, et ne lui avaient pas non plus insufflé de nouvelles valeurs. Les lobbyistes avaient pour l'essentiel atteint leur but, et le jeu restait le même : capter cet argent que l'on qualifiait de « stupide ». Hormis quelques mesures en matière de régulation, qui n'ajoutaient que de minimes contraintes, la vie suivit son cours.

Cette tragédie accéléra la perte de mes illusions. J'étais tout particulièrement déçue du rôle joué par les mathématiques. Je devais regarder en face l'horrible vérité : on avait délibérément brandi des formules dans le but d'impressionner, et non de clarifier. Pour la première fois, j'avais été directement confrontée à ce concept toxique, et j'eus alors

envie de m'échapper, de remonter le temps pour retrouver l'univers des démonstrations et des Rubik's Cubes.

Je quittai donc le fonds en 2009, avec la conviction que je m'attèlerais à corriger les ADM de la finance. De nouvelles règles obligeaient les banques à recruter des experts indépendants pour analyser leur risque. Je fus embauchée par l'une des entreprises qui fournissaient ce genre d'analyse, RiskMetrics Group, située au nord de Wall Street à une rue de distance de mon ancien employeur. Notre produit générait une avalanche de chiffres, dont chacun prédisait la probabilité qu'une certaine catégorie de valeurs mobilières ou de matières premières puisse s'écrouler au cours de la semaine, de l'année ou des cinq ans à venir. Quand tout le monde parie sur tous les actifs fluctuants du marché, une lecture judicieuse du risque vaut de l'or.

Pour calculer ce risque, notre équipe employait la méthode Monte-Carlo. Pour vous en faire une idée, imaginez simplement que vous lanciez dix mille fois la roulette au casino, en notant scrupuleusement tout du long le résultat des tirages. Avec la méthode Monte-Carlo, on partait en général des données historiques du marché et on réalisait des milliers de scénarios-tests. Comment le portefeuille étudié se serait-il comporté lors de chaque journée de cotation depuis 2010, ou depuis 2005 ? Aurait-il surmonté les jours les plus sombres du krach ? Quelle chance avait-on de voir apparaître un danger fatal d'ici une ou deux années ? Pour établir ces pronostics, les experts effectuent des milliers et des milliers de simulations. Cette méthode était largement critiquable, mais elle offrait un moyen simple d'appréhender votre risque.

Je servais d'agent de liaison, entre notre activité de gestion du risque et ceux qui en étaient les premiers et les plus fins spécialistes – les fonds spéculatifs quantitatifs. Je les appelais, ou ils m'appelaient, et nous examinions toutes les questions qu'ils se posaient au sujet de nos chiffres. Mais la plupart du temps, ils m'informaient uniquement lorsque nous avions commis une erreur. Le fait est que les fonds spéculatifs s'estimaient toujours plus perspicaces que tout le monde, et puisque la compréhension du risque représentait une condition fondamentale de leur existence, ils ne se fiaient jamais entièrement à des intervenants extérieurs comme nous. Ils possédaient leurs propres équipes pour l'étude du risque, et achetaient surtout notre produit pour être bien vus des investisseurs.

Je m'occupais aussi de l'assistance téléphonique, et il m'arrivait parfois de répondre à des clients travaillant au sein de grands établissements bancaires. Soucieux de restaurer une image grandement ternie par les événements récents, ils souhaitaient être perçus comme des gens responsables, et c'est d'abord pour cette raison qu'ils nous appelaient. Mais à l'inverse des fonds spéculatifs, ils ne montraient guère d'intérêt pour nos analyses. Le niveau de risque de leurs portefeuilles était un élément dont ils semblaient presque se moquer. Durant tout le temps que j'ai passé à l'assistance téléphonique, j'ai eu le sentiment que les personnes qui mettaient en garde contre le risque étaient considérées comme des rabat-joie, ou pire comme une menace pour le bilan de la banque. C'était encore le cas même après l'effondrement cataclysmique de 2008, et il n'est pas difficile de comprendre pourquoi. Si ces établissements avaient pu y échapper – parce qu'ils étaient trop importants pour faire faillite –, pourquoi iraient-ils s'inquiéter aujourd'hui du risque de leur portefeuille ?

Le refus d'admettre le risque imprègne en profondeur le monde de la finance. La culture de Wall Street est définie par ses traders, et le risque est un aspect qu'ils cherchent vigoureusement à sous-estimer. Cela vient de la manière dont nous déterminons le talent d'un trader, à savoir en fonction de son « ratio de Sharpe », qui se calcule en divisant les profits qu'il génère par le niveau de risque de son portefeuille. Ce ratio revêt une importance cruciale pour sa carrière, pour son bonus annuel, pour la perception même de ce qu'il est. Si l'on envisageait les traders comme un ensemble d'algorithmes désincarnés, on verrait que ces algorithmes se focalisent sans relâche sur l'optimisation du ratio de Sharpe. Dans l'idéal, celui-ci augmentera toujours, ou du moins ne descendra jamais trop bas. Si un rapport de risque sur les CDS faisait grimper l'évaluation du risque associé à l'une des positions clés d'un trader, son ratio de Sharpe se mettrait donc à chuter. À l'heure de calculer son bonus de fin d'année, il pourrait lui en coûter des centaines de milliers de dollars.

Je me rendis vite compte que mon travail était une mascarade. En 2011, ce fut à nouveau le moment d'aller voir ailleurs. Je constatai qu'il existait un marché en forte croissance pour les mathématiciens comme moi : le temps d'ajouter deux mots à mon curriculum vitae, je venais de me proclamer « *data scientist* » – experte en données –, et me retrouvais prête à m'immerger dans l'économie d'Internet. Je décrochai alors un poste dans une start-up new-yorkaise baptisée Intent Media.



Je fus d'abord chargée de construire des modèles pour anticiper le comportement des visiteurs sur différents sites de voyage. La question centrale était de savoir si les personnes qui arrivaient sur le site d'Expedia se contentaient de le parcourir ou si elles comptaient y dépenser de l'argent. Celles qui ne prévoyaient pas d'achat présentaient très peu de valeur en termes de revenus potentiels. Nous leur montrions donc des publicités comparatives pour des services concurrents comme Travelocity ou Orbitz. Si elles cliquaient dessus, cela nous rapportait quelques centimes, ce qui était toujours mieux que rien. Mais nous ne souhaitions pas diffuser ces annonces auprès des acheteurs sérieux. Dans le pire des cas, nous y gagnerions dix centimes de revenus publicitaires tout en envoyant de potentiels clients chez nos rivaux, où ils dépenseraient peut-être des milliers de dollars pour réserver un hôtel à Londres ou à Tokyo. Il faudrait plusieurs milliers de publicités visionnées pour compenser ne serait-ce que quelques centaines de dollars de commissions perdues. Il était donc crucial d'encourager ces gens-là à rester sur notre site.

Le défi consistait à concevoir un algorithme capable de distinguer les clients simplement venus faire du lèche-vitrines des vrais acheteurs. Il existait quelques signaux évidents. Étaient-ils connectés au site avec leur compte ? Avaient-ils déjà été clients ? Mais j'explorais aussi d'autres indices. Quelle heure était-il, et quel jour étions-nous ? Certaines semaines sont en effet particulièrement propices aux achats. Aux États-Unis, un pic se produit par exemple au milieu du printemps lors du Memorial Day (dernier lundi férié de mai), au moment où un grand nombre de gens organisent presque à l'unisson leurs vacances d'été. Mon algorithme attribuait aux visiteurs une valeur supérieure pendant ces périodes-là, car ils avaient davantage de chances de procéder à un achat.

D'un fonds spéculatif au commerce en ligne, le travail sur les statistiques se révélait en définitive aisément transposable – la plus grande différence étant qu'au lieu de prévoir les mouvements des marchés, je prédisais désormais les clics des internautes.

Je décelais en réalité toutes sortes de parallèles entre la finance et le Big Data. Ces deux industries exploitent avidement le même vivier de talents, issus pour la plupart d'universités d'élite telles que MIT, Princeton ou Stanford. Ces jeunes recrues sont assoiffées de succès et se sont focalisées toute leur vie sur des indicateurs externes – comme les scores au SAT<sup>5</sup> ou les taux d'admission dans les établissements d'enseignement supérieur. Que

ce soit dans le secteur de la finance ou des technologies, on a fait comprendre à ces jeunes prometteurs qu'ils deviendraient riches et qu'ils dirigeraient le monde. En outre, leur productivité montre qu'ils sont sur la bonne voie, et se traduit en dollars. D'où la conclusion – si fallacieuse qu'elle soit – que tout ce qui peut rapporter davantage d'argent est forcément bon, et « ajoute de la valeur ». Sinon, pourquoi le marché récompenserait-il ce genre de comportements ?

Dans l'une et l'autre de ces deux cultures, la richesse ne constitue plus un moyen d'assurer son train de vie. Elle devient directement liée à l'estime de soi. Un jeune citadin ayant bénéficié de tous les avantages – les études en lycée privé, la préparation approfondie aux tests d'entrée à l'université, le semestre à l'étranger, à Paris ou Shanghai – imaginera malgré tout qu'il s'est hissé parmi les privilégiés grâce à son talent, à son travail assidu et à ses prodigieuses capacités à résoudre les problèmes. L'argent dissipe tous les doutes. Et les autres membres de son cercle social se prêtent au même jeu, formant de la sorte un club fondé sur une admiration mutuelle. Ils tiennent à nous convaincre tous que le darwinisme est ici à l'œuvre, alors qu'on a bien l'impression, vu de l'extérieur, qu'ils ont à la fois profité du système et d'un simple coup de chance.

Dans ces deux secteurs, le monde réel, avec tout son désordre, est mis à l'écart. La tendance consiste à remplacer les gens par des flux de données, et à en faire des acheteurs, des électeurs ou des employés plus efficaces afin de remplir de manière optimale un objectif quelconque. C'est aussi simple à faire qu'à justifier, dès l'instant où la réussite de l'opération vous revient sous la forme d'un score anonyme, et où les individus concernés restent tout aussi abstraits que les chiffres qui s'agitent sur un écran.

Parallèlement à mon emploi dans la science des données, j'avais déjà commencé à tenir un blog et m'impliquais également de plus en plus dans le mouvement Occupy Wall Street. La rupture entre les modèles techniques et les personnes réelles, et les répercussions morales de cette rupture, m'inquiétaient davantage de jour en jour. J'observais en fait l'émergence d'un schéma identique à celui dont j'avais été témoin dans le monde de la finance : un faux sentiment de sécurité conduisait à la mise en œuvre généralisée de modèles imparfaits, à des définitions intéressées du succès, et au développement des boucles de rétroaction. Ceux qui s'y opposaient étaient considérés comme des réactionnaires nostalgiques.

Je me demandais à quoi pourrait ressembler l'équivalent de la crise du crédit dans le domaine du Big Data. Au lieu d'un effondrement, j'entrevois une dystopie croissante, avec un accroissement des inégalités. Les algorithmes s'assureraient que les gens désignés comme des « ratés » le restent. Une minorité de chanceux acquerrait toujours plus d'emprise sur l'économie des données et amasserait des fortunes scandaleuses tout en se persuadant de l'avoir mérité.

Après deux ans à travailler et à m'instruire dans la sphère du Big Data, j'avais plus ou moins perdu toutes mes illusions, et le mauvais usage des mathématiques s'amplifiait. Même en bloguant presque tous les jours, j'arrivais à peine à me tenir au fait de toutes les méthodes dont j'entendais parler, visant à manipuler, à contrôler et à intimider les gens au moyen d'algorithmes. Il y avait pour commencer ces enseignants, que je connaissais, et qui se débattaient sous le joug du modèle de la valeur ajoutée. Mais ça ne s'arrêtait pas là. Franchement alarmée, je quittai mon poste pour étudier à fond la question.

- 
1. Les instruments à terme désignent, entre autres, les contrats qui permettent de fixer des variables financières à l'avance afin d'engager de manière ferme plusieurs contractants à se positionner (achat ou vente) sur des valeurs (Source : Universalis).
  2. Les *hedge funds* sont des fonds spéculatifs non cotés en Bourse et généralement connus pour la prise de risque qu'ils représentent. Les gestionnaires de ces fonds effectuent des placements de protection contre les fluctuations du marché.
  3. La Banque centrale des États-Unis.
  4. Hausse des prix ponctuelle causée par une forte spéculation sur le marché immobilier.
  5. Scholastic Assessment Test, test d'aptitude qui détermine l'entrée dans les universités américaines.

# LA COURSE À L'ARMEMENT

Entrer à l'université

1001Ebooks

Dans certaines villes – à San Francisco ou à Portland, pour ne citer que ces deux-là –, si vous dînez avec des amis, vous découvrirez sans doute qu'il est impossible de partager une assiette. Il n'y a pas deux personnes qui mangent la même chose. Tout le monde suit un régime différent, allant du véganisme à diverses variantes du « régime paléo », et chacun ne jure que par le sien (même si ce n'est que pour un mois ou deux). Imaginez à présent que l'un d'entre eux, disons le « régime Cro-Magnon », devienne aux États-Unis la norme nationale – et que 330 millions de personnes se mettent toutes à suivre ses préceptes.

Les conséquences seraient dramatiques. Tout d'abord, un régime national unique mettrait l'économie agricole à rude épreuve. La demande pour les viandes et les fromages approuvés par ce régime grimperait en flèche, entraînant une hausse des prix. Dans le même temps, les nourritures proscrites, comme le soja et les pommes de terre, ne trouveraient plus preneur. La diversité déclinerait. Les producteurs de légumineuses, en crise, convertiraient leurs champs à l'élevage bovin et porcin, même sur des terres impropres à cet usage. Ce bétail supplémentaire consommerait d'immenses quantités d'eau. En outre, inutile de dire qu'un régime unique rendrait beaucoup d'entre nous extrêmement malheureux.

Quel rapport avec les ADM ? L'échelle. Qu'il s'agisse d'un régime alimentaire ou d'une législation fiscale, toute formule présente en théorie un caractère parfaitement inoffensif. Mais si elle acquiert la dimension d'une norme nationale ou mondiale, elle crée alors sa propre économie, dystopique et dénaturée. C'est ce qui s'est passé dans l'enseignement supérieur.

L'histoire commence en 1983. Cette année-là, un magazine d'information en difficulté, *U.S. News & World Report*, avait décidé d'entreprendre un ambitieux projet. Il allait évaluer 1 800 facultés et universités dans tous les États-Unis, et les classer selon leur degré d'excellence. Ce serait un outil très utile qui aiderait en cas de succès des millions de jeunes à prendre la première grande décision de leur existence. Ce seul choix définirait pour beaucoup leur parcours professionnel, et leur permettrait de rencontrer les amis de toute une vie – y compris souvent celui ou celle qu'ils épouseraient. Les éditeurs espéraient de surcroît qu'un numéro spécial sur le classement des établissements universitaires puisse faire sensation dans les kiosques. Cette semaine-là au moins, *U.S. News* parviendrait peut-être à égaler ses grands rivaux, *Time* et *Newsweek*.

Mais quelles informations viendraient alimenter ce nouveau baromètre ? Au départ, les journalistes d'*U.S. News* basèrent uniquement leurs scores sur les résultats des enquêtes d'opinion qu'ils avaient envoyées aux présidents d'université. Stanford fut classée comme la meilleure du pays, et Amherst, meilleure faculté pluridisciplinaire (ce qu'on appelle aux États-Unis les « *liberal arts* »). Bien qu'appréciées des lecteurs, ces évaluations rendirent beaucoup de responsables d'établissement fous de rage. La rédaction fut submergée de plaintes affirmant que le classement était injuste. Beaucoup de présidents, d'étudiants et d'anciens élèves firent valoir avec insistance qu'ils méritaient une meilleure place. Le magazine n'avait qu'à regarder les données.

Les rédacteurs d'*U.S. News* s'efforcèrent les années suivantes de mieux comprendre ce qu'ils pouvaient mesurer. De nombreux modèles débutent ainsi, par une série d'intuitions. Le processus n'a rien de scientifique, et ne repose pas sur une analyse statistique très approfondie. Dans le cas présent, il était simplement question pour eux de se demander ce qui comptait le plus en matière d'enseignement, de déterminer ensuite parmi toutes ces variables celles qu'il était possible d'évaluer, et enfin de décider du poids que chacune se verrait attribuer dans leur formule.

Dans la plupart des disciplines, l'analyse sur laquelle on fonde un modèle exige beaucoup plus de rigueur. En agronomie, les chercheurs compareront par exemple les facteurs de production – le sol, l'ensoleillement, les engrais – et certains traits spécifiques des récoltes obtenues. Ils pourront alors expérimenter et optimiser les techniques en fonction de leurs objectifs, que ce soit en termes de prix, de saveur ou de valeur nutritionnelle. Cela ne

signifie pas que les agronomes soient incapables de créer des ADM. Ils peuvent en créer, et ils en créent (en particulier lorsqu'ils négligent les effets à long terme des pesticides et leur envergure). Mais parce qu'ils se concentrent dans la plupart des cas sur des résultats clairs et précis, leurs modèles se prêtent parfaitement à l'expérimentation scientifique.

Les journalistes d'*U.S. News*, eux, se débattaient avec l'« excellence universitaire », une valeur bien plus difficile à appréhender que le coût du maïs ou la teneur protéinique en microgrammes de chaque épi. Ils n'avaient aucun moyen direct de quantifier l'impact d'un cursus de quatre ans sur un étudiant en particulier, et encore moins sur des dizaines de millions d'entre eux. Impossible de mesurer ce que l'on apprenait, l'assurance acquise, le bonheur éprouvé, les amitiés nouées, ni aucun autre aspect d'une expérience estudiantine de quatre années. L'idéal du président Lyndon Johnson en matière d'enseignement supérieur – « le moyen de s'accomplir davantage, de devenir plus productif pour soi-même et de s'enrichir sur le plan personnel » – ne rentrait pas dans leur modèle.

Ils sélectionnèrent à défaut des données indirectes qui semblaient corrélées avec la réussite. Ils examinèrent les scores au test d'aptitude SAT, les ratios étudiants-professeurs et les taux d'admission. Ils analysèrent le pourcentage de nouveaux inscrits qui franchissaient le cap de la deuxième année et le pourcentage de ceux qui décrochaient leur diplôme. Ils calculèrent la proportion d'anciens élèves qui versaient une contribution financière à leur université <sup>1</sup>, présumant qu'ils avaient une bonne chance dans ce cas d'avoir apprécié la formation reçue. Les trois quarts du score utilisé pour classer l'établissement seraient générés par un algorithme – une opinion formalisée en un code informatique –, intégrant ces données de substitution. Le dernier quart tiendrait compte de l'appréciation subjective des responsables d'université à travers tout le pays.

*U.S. News* publia un premier classement fondé sur l'analyse de données en 1988, dont les résultats paraissaient tout à fait sensés. Mais en devenant une référence nationale, il créa une boucle de rétroaction particulièrement néfaste. Le problème venait du fait que le classement s'auto-renforçait. Si une université était mal notée dans *U.S. News*, sa réputation en pâtissait, et sa situation se dégradait. Les meilleurs étudiants, de même que les meilleurs professeurs, faisaient en sorte de l'éviter. Ses anciens élèves, indignés, réduisaient leurs contributions. Elle dégringolait par suite d'autant plus bas

dans le classement. La place qu'elle y occupait décidait en somme de son sort.

Les administrateurs d'université avaient eu par le passé toutes sortes de moyens, pour la plupart empiriques, de jauger la réussite de leur établissement. Les étudiants faisaient l'éloge de tels ou tels professeurs. Certains diplômés poursuivaient d'illustres carrières de diplomates ou d'entrepreneurs. D'autres publiaient des romans récompensés par la critique. Tous ces éléments généraient un excellent bouche-à-oreille, qui confirmait la réputation de l'université. Mais celle de Macalester était-elle meilleure que celle de Reed, ou celle de l'Iowa meilleure que celle de l'Illinois ? Difficile à dire. Les universités étaient comme autant de styles de musique ou de régimes alimentaires différents. Des opinions divergentes avaient le droit de s'exprimer, avec de bons arguments de part et d'autre. Désormais, le vaste écosystème réputationnel des établissements d'enseignement supérieur se trouvait éclipsé par une simple colonne de chiffres.

Si l'on envisage cette évolution du point de vue d'un président d'université, elle s'avère bien triste. La plupart d'entre eux chérissaient sans aucun doute leur propre expérience d'étudiant – c'est elle en partie qui les avait incités à gravir les échelons de la hiérarchie académique. Et voilà pourtant qu'au sommet de leur carrière ils devaient consacrer une énergie considérable à accroître leurs performances sur quinze aspects définis par une équipe de journalistes d'un magazine de seconde zone. Ils se retrouvaient presque à nouveau dans la peau d'un élève, cherchant à obtenir de bonnes notes de la part de son professeur. Ils étaient en réalité pris au piège d'un modèle inflexible – une ADM.

Si la liste d'*U.S. News* n'avait connu qu'un succès restreint, elle n'aurait pas posé de souci. Mais elle a gagné au contraire une influence gigantesque et s'est rapidement imposée comme une référence nationale. Depuis cette époque, elle met le système éducatif américain sous pression, en dressant une liste rigide d'impératifs pour les administrateurs d'université autant que pour les étudiants. Le classement publié par *U.S. News* a une énorme portée, provoque des dégâts étendus et génère une spirale presque infinie de multiples boucles de rétroaction destructrices. Tout en n'étant pas aussi opaque que beaucoup d'autres modèles, il n'en demeure pas moins une authentique ADM.

Certains administrateurs se sont montrés prêts à tout pour obtenir une meilleure place. L'université Baylor a pris en charge les frais de l'examen SAT pour que les étudiants admis repassent le test, dans l'espoir qu'une seconde tentative améliorerait leurs scores – et donc son classement. De petits établissements élitistes, notamment les universités Bucknell, en Pennsylvanie, et Claremont McKenna, en Californie, ont envoyé de fausses données à *U.S. News*, en gonflant les scores SAT des nouveaux inscrits. L'université d'Iona, dans l'État de New York, a quant à elle reconnu en 2011 que ses employés avaient à peu près truqué tous les chiffres : scores aux tests, taux d'admission et de diplômés, nombre d'étudiants conservés après la première année, ratio étudiants-professeurs et générosité des anciens élèves. Le mensonge avait payé, du moins pour un temps. *U.S. News* a estimé que ces données incorrectes avaient fait passer Iona de la quinzième à la treizième place parmi les établissements de l'aire nord-est des États-Unis.

L'immense majorité des administrateurs d'université cherchaient des moyens moins flagrants d'améliorer leur classement. Au lieu de tricher, ils s'évertuaient à renforcer chacun des indicateurs qui entraient dans le calcul de leur score. Ils pouvaient ainsi affirmer faire un usage le plus rentable possible des ressources dont ils disposaient. Après tout, s'ils s'employaient à satisfaire l'algorithme d'*U.S. News*, ils récolteraient alors davantage de fonds, attireraient des étudiants et des professeurs plus brillants, et continueraient de monter dans la liste. Avaient-ils vraiment le choix ?

Robert Morse, qui travaille chez *U.S. News* depuis 1976 et supervise le classement des universités, a soutenu dans des entretiens que ce dernier poussait les établissements à se fixer des objectifs constructifs. S'ils parvenaient à augmenter le taux de diplômés ou à réduire le nombre d'étudiants par classe, c'était une bonne chose. L'enseignement tirait profit de l'accent mis sur ces aspects. Morse a néanmoins admis que les données les plus pertinentes – ce que les étudiants avaient appris dans chaque école – restaient impossibles à quantifier. Mais le modèle utilisé par *U.S. News*, basé sur des mesures indirectes, constituait faute de mieux le meilleur outil disponible.

Un modèle créé à partir de données de substitution est bien plus aisé à duper. La raison est que ces données supplétives sont plus faciles à manipuler que la réalité complexe qu'elles représentent. Prenons un exemple. Imaginons qu'un site web cherche à recruter un virtuose des



médias sociaux. Beaucoup de gens postulent, et fournissent des renseignements sur les diverses campagnes marketing qu'ils ont dirigées. Mais retracer et évaluer l'ensemble de leurs travaux demande beaucoup trop de temps. Le responsable du recrutement se contente donc à la place d'accorder un intérêt particulier aux candidats qui possèdent le plus grand nombre de *followers* sur Twitter. N'est-ce pas là le signe d'une forte implication dans les médias sociaux ?

Il s'agit certes d'un critère de rechange acceptable. Mais qu'arrivera-t-il quand l'information sera divulguée – elle le sera à coup sûr –, et que l'on saura que la clé pour obtenir un poste dans cette entreprise consiste à réunir une communauté sur Twitter ? Les candidats mettront vite tout en œuvre pour faire grimper leur compteur d'abonnés. Certains paieront un service à 19,95 dollars qui remplira leur fil Twitter de milliers de *followers*, générés pour l'essentiel par des robots. À mesure que les gens leurreront le système, la donnée indirecte utilisée perdra de son efficacité et les tricheurs engendreront de faux positifs.

Dans le cas du classement établi par *U.S. News*, le score fut rapidement accepté par tous – des futurs aux anciens étudiants et jusqu'aux départements des ressources humaines – comme une mesure de la qualité d'enseignement. Les universités jouèrent donc le jeu. Elles firent de gros efforts pour s'améliorer dans chacun des domaines évalués. Beaucoup étaient en réalité particulièrement frustrées par les 25 % de la note sur lesquels elles n'avaient aucun pouvoir – le score lié à leur réputation, qui résultait des questionnaires remplis par les présidents et les doyens de facultés.

Cette partie de l'analyse, comme toute collecte d'opinions humaines, était assurée de succomber à l'ignorance et aux bons vieux préjugés. Elle avait tendance à protéger les écoles célèbres classées en tête de liste, car c'étaient celles que les gens connaissaient. Et elle rendait les choses plus difficiles pour les établissements en devenir.

En 2008, la Texas Christian University (TCU), située à Fort Worth, était en chute dans le classement *U.S. News*. Elle était passée d'un score de 97, trois ans plus tôt, à 105, puis 108 et maintenant 113. Ce recul suscitait une certaine agitation parmi ses soutiens et anciens élèves, et mettait son président, Victor Boschini, sur la sellette. « C'est très frustrant pour moi », expliqua Boschini sur TCU 360, le site d'information du campus. Il souligna que l'université était en progrès sur tous les indicateurs. « Le taux

de conservation des effectifs s'améliore, nos levées de fonds également, tous les aspects dont on nous parle. »

L'analyse de Boschini posait deux problèmes. D'une part, le modèle de classement mis en œuvre par *U.S. News* ne jugeait pas les universités isolément les unes des autres. Une école aurait beau améliorer ses résultats, elle perdrait tout de même des places si d'autres progressaient plus vite. Pour employer la terminologie académique, le modèle d'*U.S. News* notait les établissements sur une courbe en cloche. Ce qui nourrissait une sorte de course aux armements de plus en plus intense.

L'autre problème concernait le score réputationnel, les 25 % que la TCU était incapable de maîtriser. Raymond Brown, doyen des admissions, fit remarquer que la réputation était la variable la plus fortement prise en compte, « ce qui est absurde puisqu'elle est totalement subjective ». Wes Waggoner, directeur des admissions en première année, ajouta que les universités se vendaient chacune auprès des autres afin d'accroître leur score de réputation : « Je reçois des courriers d'autres établissements, qui cherchent à [nous] convaincre de leurs qualités. »

Malgré ces récriminations, la TCU entreprit d'améliorer les 75 % qu'elle pouvait contrôler. Après tout, si le score de l'université grimpait, sa réputation finirait par suivre. Avec le temps, ses pairs prendraient note des progrès accomplis et lui accorderaient une meilleure appréciation. La clé consistait à orienter les choses dans la bonne direction.

La TCU lança une campagne de financement qui visait à recueillir 250 millions de dollars. Son objectif fut largement dépassé, si bien qu'en 2009 elle lui avait déjà rapporté 434 millions. Ce seul élément fit remonter le classement de l'université, puisque la levée de fonds faisait partie des indicateurs mesurés. Elle consacra l'essentiel de cet argent à réaménager le campus, avec notamment 100 millions de travaux pour la place centrale et la construction d'une nouvelle maison des étudiants, dans le but d'en faire une destination plus séduisante pour les élèves. S'il n'y a rien de mal à cela, cet aspect sert fort à propos l'algorithme d'*U.S. News*. Plus les étudiants sont nombreux à postuler, et plus l'école peut se montrer sélective.

Plus important peut-être, la TCU s'équipa d'un centre d'entraînement sportif dernier cri et injecta davantage de ressources dans son programme de football américain. Au cours des années suivantes, les Horned Frogs de TCU s'imposèrent comme une équipe de choc au niveau national. Ils furent invaincus en 2010, battant le Wisconsin au Rose Bowl Stadium.

Ce succès permit à la TCU de bénéficier de ce qu'on appelle l'« effet Flutie ». En 1984, au terme d'un des matchs universitaires les plus palpitants de l'histoire, un *quarterback* de l'université de Boston, Doug Flutie, avait réalisé à la dernière seconde une longue passe désespérée pour battre l'université de Miami. Flutie était devenu une légende. En l'espace de deux années, les candidatures à l'université de Boston avaient grimpé de 30 %. La même hausse s'était produite pour l'université de Georgetown quand son équipe de basket, menée par Patrick Ewing, avait disputé trois finales de championnats nationaux. Un programme sportif couronné de succès se révèle auprès de certains postulants comme l'outil de promotion le plus efficace. Pour des légions de lycéens portés sur le sport, qui assistent aux rencontres universitaires à la télévision, les établissements dotés de grandes équipes paraissent attirants<sup>2</sup>. Les étudiants arborent fièrement le nom de leur école. Ils se peignent le visage et célèbrent ses victoires. Les demandes d'inscription connaissent par suite une forte augmentation. Avec un nombre accru de candidatures, les administrateurs peuvent placer la barre plus haut et relever ainsi la moyenne des scores obtenus aux tests d'entrée, ce qui contribue à faire monter la cote de l'université. Plus elle refuse de candidats, plus son taux d'admission baisse (et donc s'améliore, en termes de classement).

La stratégie de la TCU fonctionna. En 2013, elle était devenue la deuxième université la plus sélective du Texas, devancée seulement par la prestigieuse université Rice de Houston. Cette même année, elle avait enregistré sur les tests d'aptitude SAT et ACT les scores les plus élevés de son histoire. Elle gagna des places dans la liste d'*U.S. News*. En 2015, elle terminait au soixante-seizième rang, soit trente-sept de mieux en à peine sept ans.

En dépit de mes griefs vis-à-vis du modèle employé par *U.S. News*, et de son statut d'ADM, il importe de noter que cette spectaculaire progression pourrait bien avoir été profitable à la TCU, en tant qu'établissement universitaire. En effet, la plupart des données de substitution exploitées par le modèle reflètent en fin de compte dans une certaine mesure la qualité globale d'une école. De la même manière, beaucoup d'adeptes du régime Cro-Magnon s'en portent très bien. Ce n'est pas le modèle lui-même qui pose problème, mais sa portée. Il contraint tout le monde à viser exactement les mêmes objectifs, générant de la sorte une compétition acharnée – et sans qu'on le veuille un grand nombre de conséquences nocives.

Avant que les classements n'existent, par exemple, les futurs étudiants pouvaient ainsi dormir tranquilles en sachant qu'ils avaient candidaté à un établissement dit « de secours », avec des critères d'entrée plus faibles. Si les universités placées en tête de leurs vœux ne les acceptaient pas – celles où ils auraient rêvé d'aller, sans grand espoir d'y parvenir (les « *stretch schools* »), mais aussi d'autres plus sûrs (les écoles « cibles ») –, ils bénéficieraient au sein de leur établissement de secours d'une formation tout à fait correcte et rejoindraient peut-être un de leurs premiers choix au bout d'un an ou deux.

Le concept de l'établissement de secours a désormais pratiquement disparu, et on le doit en grande partie au palmarès publié par *U.S. News*. Comme l'exemple de la TCU nous l'a montré, la sélectivité favorise un meilleur classement. Si le bureau des admissions d'une université est submergé de demandes, c'est le signe qu'elle fonctionne bien. L'afflux de candidatures témoigne de sa réputation. Et si un établissement peut se permettre de refuser l'immense majorité de ces postulants, il se retrouvera alors probablement avec des étudiants d'un meilleur calibre. Comme beaucoup de données indirectes, cet indicateur semble pertinent. Il épouse les mouvements du marché.

Mais ce marché peut être manipulé. Une école traditionnellement choisie comme établissement de secours pourra par exemple examiner les données historiques et s'apercevoir qu'une petite fraction seulement des meilleurs candidats ont fini par y être accueillis. La plupart sont parvenus à intégrer une de leurs écoles cibles ou de leurs « *stretch schools* », et n'ont pas eu besoin de ce qui s'apparentait à une garantie prévoyance. Dans l'optique d'augmenter son score de sélectivité, elle pourra désormais rejeter les excellents candidats, qui ont les plus grandes chances, d'après son propre algorithme, de ne pas s'inscrire. Mais ce processus est loin d'être infallible. Et l'université, malgré le travail des experts en données employés au bureau des admissions, laisse sans doute échapper ainsi un certain nombre d'étudiants de haut niveau qui auraient choisi d'y entrer – ceux qui découvrent avec consternation que les prétendues écoles de secours ne constituent plus un pari assuré.

Ce mécanisme alambiqué n'apporte rien à l'enseignement. L'université en pâtit. Elle perd les meilleurs étudiants – les « vedettes » qui enrichissent l'expérience de chacun, y compris des professeurs. Un ancien établissement de secours risque maintenant de devoir utiliser en partie une aide financière

précieuse pour attirer certaines de ces « vedettes » sur son campus. Ce qui peut signifier moins d'argent pour les étudiants qui en ont le plus besoin.

\*\*\*

C'est là que se situe le plus gros défaut du palmarès d'*U.S. News*. Les données indirectes choisies par les journalistes pour évaluer l'excellence universitaire sont après tout parfaitement sensées. Leur spectaculaire échec résulte plutôt de ce qu'ils ont choisi de ne *pas* prendre en compte : les frais de scolarité. Dans leur modèle, le financement des études a été laissé de côté <sup>3</sup>.

Ce qui nous amène à une question cruciale, que nous affronterons à maintes reprises dans cet ouvrage : quel objectif l'auteur d'un modèle poursuit-il ? Dans le cas présent, mettez-vous à la place des rédacteurs d'*U.S. News*, en 1988. Lors de l'élaboration de leur premier modèle statistique, à partir de quel moment pouvaient-ils savoir qu'il fonctionnait ? Eh bien, ils estimèrent que ce modèle serait dès le départ beaucoup plus crédible s'il reflétait la hiérarchie établie. Si Harvard, Stanford, Princeton et Yale arrivaient en tête, cela confirmerait leurs prédictions, reproduisant les modèles internes que les journalistes et leurs clients avaient eux-mêmes à l'esprit. Pour construire un tel modèle, il leur suffisait d'étudier ces universités de prestige et de prendre en considération ce qui les distinguait tant des autres. Qu'avaient-elles en commun, par opposition à l'établissement de secours d'une ville voisine ? Leurs étudiants affichaient des scores stratosphériques à l'examen SAT, et décrochaient leur diplôme avec une régularité de métronome. Les anciens élèves étaient riches, et reversaient d'importantes donations. En analysant les vertus des universités de marque, l'équipe d'évaluation créait un point de référence – celui de l'élite – pour mesurer l'excellence.

Si elle incorporait en revanche dans sa formule le coût des études, cela pouvait avoir d'étranges répercussions. Des universités bon marché étaient susceptibles de faire irruption dans la hiérarchie des meilleurs établissements, au risque de susciter des surprises et de semer le doute. Le public n'accorderait peut-être pas vraiment foi, dans ce cas, au palmarès d'*U.S. News*. Il était bien plus sûr de faire figurer les vénérables universités championnes en tête de classement. Bien sûr, elles coûtaient très cher. Mais c'était peut-être justement le prix de l'excellence.

En n'incluant pas le coût des études dans la formule, c'est un peu comme si *U.S. News* avait tendu un chèque en blanc aux présidents d'université. On les enjoignait de maximiser leurs performances sur quinze aspects donnés, et la maîtrise des coûts n'en faisait pas partie. S'ils augmentaient leurs tarifs, ils dégageraient même davantage de ressources à réinjecter dans les domaines sur lesquels ils étaient évalués.

Depuis, les frais de scolarité n'ont cessé de s'envoler. Entre 1985 et 2013, le coût des études supérieures a augmenté de plus de 500 %, près de quatre fois plus vite que l'inflation. Pour attirer les meilleurs élèves, les universités – comme on l'a vu avec la TCU – se sont engagées dans de grands chantiers, comprenant la construction de maisons d'étudiants modernes aux grandes baies vitrées, de dortoirs luxueux, et de gymnases équipés de murs d'escalade ou de bains à remous. Tous ces aménagements pourraient être merveilleux pour les étudiants, et serviraient à améliorer leur expérience universitaire, si ce n'était pas eux qui en payaient le prix, sous forme de prêts dont ils portent ensuite le fardeau sur des dizaines d'années. Le classement d'*U.S. News* ne peut être tenu pour unique responsable de cette tendance. Notre société tout entière a adopté non seulement l'idée qu'une formation supérieure revêtait une importance essentielle, mais aussi qu'un diplôme décroché dans un établissement de haut rang pouvait propulser un étudiant vers une vie de pouvoir et de privilèges. L'ADM d'*U.S. News* s'est nourrie de ces convictions, de ces peurs et de ces névroses. Elle a engendré de puissantes incitations qui ont poussé à la dépense, tandis qu'on fermait les yeux sur l'explosion des frais de scolarité.

Lorsqu'elles se mettent dans l'optique de progresser au palmarès d'*U.S. News*, les universités gèrent leur population étudiante presque à la manière d'un portefeuille d'investissements. De la publicité à la politique, on rencontrera souvent ce type d'approche dans l'univers des données. Pour les administrateurs d'universités, chaque étudiant potentiel possède une série d'atouts et présente aussi en général un ou deux inconvénients. Une athlète très douée constitue par exemple un atout pour l'université, mais peut avoir de mauvais scores aux tests d'entrée, ou une place moyenne parmi les élèves de sa classe. Il est aussi possible qu'elle ait besoin d'être aidée financièrement – autre inconvénient. Pour équilibrer le portefeuille, on ira trouver dans l'idéal d'autres candidats capables de payer leurs études et affichant des scores élevés aux tests. Or, ces candidats idéaux, une fois acceptés, choisiront peut-être d'aller ailleurs. C'est un risque, qui doit être

quantifié. Ce processus est effroyablement complexe, et toute une industrie du conseil s'est développée afin d'« optimiser le recrutement ».

Le cabinet Noel-Levitz, spécialisé dans l'enseignement, propose un outil d'analyse prédictive baptisé ForecastPlus, qui permet aux administrateurs de classer les recrues potentielles en fonction de leur région, de leur sexe, de leurs origines, de leur domaine d'étude, de leurs résultats scolaires ou de « toute autre caractéristique souhaitée ». Un autre cabinet de conseil, RightStudent, collecte et vend des données aux universités pour les aider à cibler les candidats les plus prometteurs – notamment les étudiants en mesure de régler l'intégralité de leurs frais de scolarité, ou d'autres susceptibles d'être éligibles à des bourses externes (paradoxalement, présenter un trouble de l'apprentissage peut donc devenir un atout<sup>4</sup>).

Toute cette activité s'inscrit dans le cadre d'un vaste écosystème entourant le classement d'*U.S. News*, dont le modèle fait office de loi fondamentale. Si les rédacteurs révisent les pondérations du modèle, en accordant par exemple moins d'attention aux scores SAT ou davantage aux taux de diplômés, l'écosystème éducatif dans son ensemble doit alors s'adapter. Cela va des universités aux consultants, en passant par les services d'orientation des lycées, et même jusqu'aux étudiants.

Les classements eux-mêmes constituent naturellement un fonds de commerce en pleine croissance. Le magazine *U.S. News & World Report*, longtemps unique affaire de sa maison d'édition, a périclité avant de disparaître des kiosques en 2010. Mais l'activité de notation continue de prospérer, s'étendant aux facultés de médecine, aux écoles de dentisterie, ainsi qu'aux cursus pluridisciplinaires et d'ingénierie. *U.S. News* s'occupe même de classer les lycées.

À mesure que les classements gagnent en importance, les efforts visant à les berner s'intensifient eux aussi. Dans un palmarès mondial publié par *U.S. News* en 2014, le département de mathématiques de l'université du roi Abdulaziz, en Arabie Saoudite, était arrivé à la septième place, immédiatement derrière Harvard. Il n'existait que depuis deux ans, mais avait réussi d'une manière ou d'une autre à devancer plusieurs géants de la discipline, dont Cambridge et le MIT.

De prime abord, cela semblait être une évolution positive. Le MIT et Cambridge s'étaient peut-être reposés sur leurs lauriers, pendant qu'un vaillant nouveau venu se hissait parmi l'élite. Avec un classement basé exclusivement sur la réputation, un tel retournement aurait pris des dizaines

d'années. Les données, en revanche, peuvent faire très vite émerger des surprises.

Mais il est également possible de tromper les algorithmes. Lior Pachter, bio-informaticien à Berkeley, s'est penché sur la question. Il a découvert que l'université saoudienne avait contacté de nombreux mathématiciens dont les travaux étaient fréquemment cités, et leur avait offert 72 000 dollars pour occuper une fonction de professeur adjoint. L'accord, selon les termes d'un courrier de recrutement posté par Pachter sur son blog, stipulait que ces mathématiciens devaient travailler trois semaines par an en Arabie Saoudite. L'université les faisait venir par avion en classe affaires et les logeait dans un hôtel cinq étoiles. On peut concevoir que leur travail sur place apportait au plan local une certaine valeur ajoutée. Mais l'université avait aussi exigé qu'ils modifient leur affiliation sur le site de citations académiques Thomson Reuters, une référence essentielle pour le classement *U.S. News*. L'université saoudienne pouvait de la sorte revendiquer comme siennes les publications de ses nouveaux professeurs adjoints. Et puisque le nombre de citations constituait l'une des sources principales de l'algorithme, l'université du roi Abdulaziz était montée en flèche dans le classement.

\*\*\*

Les étudiants de la ville chinoise de Zhongxiang avaient la réputation de réussir brillamment le test national normalisé, ou *gaokao*, et de gagner ainsi des places dans les meilleures universités du pays. Ils réussissaient en fait si bien que les autorités commencèrent à les soupçonner de tricherie. Selon un article en ligne du quotidien britannique *The Telegraph*, les soupçons s'accrochèrent en 2012, après que les autorités provinciales eurent découvert quatre-vingt-dix-neuf copies identiques lors d'un seul et unique test.

L'année suivante, en arrivant pour passer l'examen, les étudiants de Zhongxiang virent avec désarroi qu'on leur faisait franchir des détecteurs de métaux et qu'on les obligeait à se séparer de leurs téléphones mobiles. Certains abandonnèrent de minuscules transmetteurs déguisés en gommes. Une fois à l'intérieur, les étudiants se retrouvèrent accompagnés de cinquante-quatre enquêteurs issus de différentes académies scolaires. Quelques-uns de ces enquêteurs se rendirent en face, dans un hôtel, et



découvrirent là des groupes d'individus qui se tenaient prêts à communiquer avec les étudiants par le biais de leurs transmetteurs.

Cette répression contre la triche suscita une éruption de colère. Deux milliers de manifestants se rassemblèrent dans la rue à l'extérieur de l'école et se mirent à lancer des pierres. « Nous réclamons l'équité, scandaient-ils. Il n'y a pas d'équité si vous ne nous laissez pas tricher. »

Cela ressemble à une plaisanterie, mais ils étaient tout à fait sérieux. Les enjeux pour les étudiants étaient extrêmement élevés. Tel qu'ils voyaient les choses, la chance leur était donnée soit de suivre une formation d'élite et de mener une carrière fructueuse, soit de rester enlisés dans leur ville de province, un coin relativement perdu. Et – que ce soit le cas ou pas –, ils avaient l'impression que les autres académies trichaient. Empêcher les étudiants de Zhongxiang de faire de même était donc bel et bien inéquitable. Dans un système où la tricherie tient lieu de norme, suivre les règles équivaut à un handicap. Demandez donc aux coureurs du Tour de France, battus à plate couture sept années d'affilée par Lance Armstrong et ses coéquipiers dopés.

Dans un scénario pareil, le seul moyen de gagner consiste à acquérir un avantage et à s'assurer que d'autres n'en obtiennent pas un plus grand. C'est ce qui se passe non seulement en Chine mais aussi aux États-Unis, où les responsables des admissions en lycée, les parents et les élèves se trouvent pris dans un déploiement d'efforts désespérés pour manipuler le système enfanté par le modèle de *U.S. News*.

La boucle de rétroaction du modèle et l'angoisse qu'il génère font prospérer toute une industrie du tutorat. Beaucoup de ces services impliquent une dépense conséquente. Un « stage d'entraînement pour l'inscription à l'université », organisé par une entreprise du nom de Top Tier Admissions, coûte ainsi 16 000 dollars pour une durée de quatre jours (gîte et couvert non compris). Lors des sessions proposées, les élèves de première travaillent leur « *essay*<sup>5</sup> », apprennent à réussir haut la main leurs entretiens, et élaborent une « fiche d'activité » résumant l'ensemble des distinctions reçues, des sports pratiqués, des activités associatives et des actions de bénévolat effectuées, toutes choses que les responsables des admissions tiennent à connaître.

On pourrait penser que 16 000 dollars représentent beaucoup d'argent. Mais tout comme les manifestants chinois de Zhongxiang, de nombreuses familles américaines s'inquiètent davantage de la réussite et

l'épanouissement futurs de leurs enfants, qui dépendent à leur sens de l'admission dans une université d'élite.

Les tuteurs de ces « stages d'entraînement » les plus efficaces possèdent une bonne compréhension des modèles d'admission de chaque établissement, qui leur permet de déterminer dans quelle mesure un éventuel futur étudiant pourra trouver sa place au sein de leurs portefeuilles respectifs. Un entrepreneur basé en Californie, Steven Ma, pousse à l'extrême cette approche fondée sur le concept de marché. Fondateur de ThinkTank Learning, Ma évalue les étudiants potentiels au travers de son propre modèle et calcule leur probabilité d'entrer dans les universités visées. Interrogé par *Bloomberg BusinessWeek*, il a par exemple indiqué qu'un lycéen de terminale né sur le sol américain, affichant un GPA<sup>6</sup> de 3,8, un score SAT de 2 000 et huit cents heures d'activités parascolaires, avait 20,4 % de chances d'entrer à l'université de New York et 28,1 % d'être accepté à l'université de Californie-Sud. ThinkTank propose alors des modules de conseil avec garantie de résultat. Si cet hypothétique étudiant suit la préparation dispensée et entre à l'université de New York, elle lui coûte 25 931 dollars, ou 18 826 dollars pour l'université de Californie-Sud. S'il est refusé, elle ne lui coûte rien.

Le modèle d'admission de chaque université dérive au moins en partie du modèle d'*U.S. News*, et chacun constitue une mini-ADM. Ces modèles plongent les étudiants et leurs parents dans une grande panique et les conduisent à dépenser des sommes indécentes. Ils présentent en outre un caractère opaque, laissant ainsi la plupart des participants (ou victimes) dans une ignorance totale. Ils s'avèrent en revanche très rentables pour des consultants comme Steven Ma, qui parviennent à percer leurs secrets soit en cultivant certaines relations au sein des universités, soit par rétro-ingénierie de leurs algorithmes.

Les victimes sont bien entendu cette vaste majorité d'Américains, les familles pauvres et de la classe moyenne, qui n'ont pas des milliers de dollars à dépenser pour payer des cours ou des conseillers. Elles se voient privées d'un précieux savoir, réservé aux initiés. Le résultat, c'est un système éducatif qui favorise les privilégiés et qui dessert les élèves nécessiteux, fermant la porte à une grande majorité d'entre eux – et les poussant sur la voie de la pauvreté. Il accroît la fracture sociale.

Néanmoins, ceux qui se fraient un chemin jusqu'à une université de haut niveau sont eux aussi perdants. Si l'on y réfléchit bien, le jeu des

admissions dans l'enseignement supérieur, quoique lucratif pour certains, n'apporte pratiquement aucune valeur éducative. Cette mécanique complexe et stressante se borne à réordonner et à reclasser d'une nouvelle façon le même vivier de jeunes post-adolescents. Ce n'est pas en se pliant à de nombreuses contraintes supplémentaires, ni en rédigeant des présentations écrites méticuleusement ciblées sous l'œil vigilant de tuteurs professionnels, qu'ils acquièrent la maîtrise de compétences essentielles. Certains fouillent sur Internet pour y trouver ce genre de tutorat à prix cassé. Tous, qu'ils soient aisés ou qu'ils appartiennent à la classe populaire, sont simplement formés pour s'adapter à une gigantesque machine – pour contenter une ADM. Et une fois l'épreuve passée, nombre d'entre eux seront accablés de dettes qu'ils mettront des dizaines d'années à rembourser. Ce ne sont que des pions, dans une course à l'armement qui se révèle de surcroît particulièrement nuisible.

Peut-on corriger cela ? Durant son second mandat, le président Obama avait suggéré l'élaboration d'un nouveau modèle pour le classement des universités, plus en phase avec les priorités du pays et les ressources des classes moyennes que celui d'*U.S. News*. L'objectif était accessoirement de réduire le pouvoir des *for-profit schools*, les établissements à but lucratif (véritables sangsues, une calamité que nous étudierons dans le chapitre suivant). L'idée d'Obama visait à lui associer des indicateurs différents, parmi lesquels l'accessibilité financière, le pourcentage d'étudiants pauvres et issus des minorités, ainsi que l'insertion professionnelle en fin de cursus. À l'instar du palmarès *U.S. News*, il tiendrait également compte du taux de diplômés. Si les établissements tombaient en dessous des minima requis sur ces divers aspects, ils se verraient exclus du marché des prêts étudiants fédéraux qui représente 180 millions de dollars par an (et dont se repaissent les universités à but lucratif).

Tous ces objectifs paraissent à coup sûr louables, mais n'importe quel système de classement reste exposé aux manipulations. Et quand cela se produit, il engendre alors de nouvelles boucles de rétroaction, en même temps qu'une foule de conséquences indésirables.

Il est facile par exemple, en abaissant les critères, d'augmenter le taux de diplômés. Beaucoup d'étudiants peinent à atteindre les prérequis en maths et en sciences, et éprouvent des difficultés en langues étrangères. Modérons ces exigences, et un nombre plus important d'étudiants obtiendront alors leur diplôme. Mais si l'un des buts de notre système éducatif est de fournir

davantage de scientifiques et de technologues pour répondre aux besoins de l'économie mondiale, est-ce bien intelligent ? Gonfler les données sur le niveau de revenu des diplômés serait également un jeu d'enfant. Il suffirait que les universités restreignent leurs programmes pluridisciplinaires, et se débarrassent tant qu'à faire des unités dispensant des formations d'enseignants et de travailleurs sociaux, puisque ces derniers gagnent moins que les ingénieurs, les chimistes et les informaticiens. Ils n'en sont pourtant pas moins précieux pour la société.

Il ne serait pas non plus très compliqué d'abaisser les coûts. L'une des méthodes, qui gagne dès aujourd'hui en popularité, consiste à réduire le pourcentage de professeurs titulaires en remplaçant ces personnels onéreux, lorsqu'ils partent à la retraite, par des formateurs ou des assistants moins coûteux. Pour quelques départements de certaines universités, cette option pourrait être pertinente. Mais elle a un prix. Les professeurs titulaires dirigent avec les étudiants de troisième cycle d'importantes recherches et servent de référence pour leur département, tandis que des assistants surmenés, susceptibles de donner cinq cours dans trois établissements uniquement pour payer leur loyer, ont rarement le temps et l'énergie nécessaires pour fournir plus qu'un enseignement de base. Une autre démarche possible, celle de supprimer des postes administratifs inutiles, semble bien trop rarement mise en œuvre.

Le nombre d'étudiants « en situation d'emploi neuf mois après l'obtention de leur diplôme » peut également être manipulé. En 2011, un reportage du *New York Times* s'est intéressé aux facultés de droit, qui sont d'ores et déjà évaluées en fonction de l'aptitude à procurer aux étudiants des perspectives de carrière. Imaginons qu'un avocat tout juste sorti des bancs de l'école, portant la charge d'un prêt étudiant de 150 000 dollars, trouve un travail de barman. Pour certaines facultés sans scrupules visées par l'enquête du *Times*, il fera alors partie des diplômés « en situation d'emploi ». Certaines allaient même plus loin et embauchaient en intérim leurs propres diplômés, payés à l'heure, quand la période fatidique des neuf mois touchait à son terme. D'autres envoyaient des questionnaires à leurs anciens élèves récemment sortis de l'école et notaient « en situation d'emploi » tous ceux qui n'y répondaient pas.

\*\*\*

Peut-être valait-il mieux que l'administration Obama n'ait pas réussi pas à élaborer son système de classement revu et corrigé, tant les présidents d'université s'y sont violemment opposés. Ils avaient en effet passé des dizaines d'années à réaliser les optimisations nécessaires pour contenter l'ADM d'*U.S. News*. Une nouvelle formule basée, entre autres indicateurs, sur le taux de diplômés, la taille des classes, l'insertion professionnelle et le niveau de revenu des anciens élèves risquait de chambouler leur classement et de mettre à mal leur réputation. Sans doute avaient-ils aussi soulevé d'excellents arguments quant aux vulnérabilités inhérentes à un tout nouveau modèle et aux nouvelles boucles de rétroaction qu'il engendrerait.

Le gouvernement avait donc capitulé. Et le résultat n'en est peut-être que meilleur : au lieu d'un classement, le ministère de l'Éducation a mis en ligne une grande masse de données. Les étudiants peuvent ainsi poser leurs propres questions sur les aspects qui comptent pour eux – notamment la taille des classes, le taux de diplômés et la dette moyenne en sortie d'études. Ils n'ont aucun besoin de connaître les statistiques ou la pondération des variables. Le logiciel lui-même, à la manière d'un site de voyages en ligne, crée des modèles individuels à destination de chaque personne. Pensez donc : un système transparent, contrôlé par l'utilisateur, et individualisé... On pourrait appeler ça le contraire d'une ADM.

---

1. Il est courant que les universités américaines sollicitent et reçoivent des dons parfois très importants de la part de leurs anciens élèves. Ces donations ne sont pas obligatoires.

2. Des bourses importantes sont également attribuées aux étudiants qui se distinguent pour leurs aptitudes sportives à l'université, leur permettant parfois d'amortir le coût des études, synonyme de dettes importantes.

3. Aux États-Unis, la question du financement des études est au centre des préoccupations pour les futurs étudiants. L'accès à l'université est payant, et varie d'environ 15 000 à 20 000 dollars par an pour les universités « publiques », jusqu'à environ 100 000 dollars par an pour les universités privées les plus prestigieuses, ce qui donne lieu à un système de prêts étudiants largement généralisé et un endettement sur de longues durées, à moins de bénéficier d'une bourse. Les universités dites « à but non lucratif » (*non-profit schools*) peuvent être privées ou publiques, en fonction de chaque État, et restent majoritaires par rapport aux écoles à but lucratif (*for-profit schools*) qui fonctionnent comme des compagnies privées, avec des actionnaires et des revenus privés.

4. Par comparaison, en France, c'est l'algorithme de Parcoursup qui depuis début 2018 détermine le sort des étudiants et leur répartition dans les universités, au cours d'un processus qui demeure problématique (plus de 100 000 élèves sans affiliation à quelques semaines de la rentrée 2018).

5. Présentation écrite sur un sujet choisi par le candidat jouant un rôle important dans la sélection (*note du traducteur*).

6. Le GPA (Grade Point Average) est une note attribuée aux lycéens en fonction de la moyenne reçue dans chaque matière étudiée. 3,8 correspond environ à une note de 16/20.

# MACHINE

## DE PROPAGANDE

La publicité en ligne

1001Ebooks

Un jour, durant mon passage comme experte en données au sein de la start-up publicitaire Intent Media, un investisseur en capital-risque de premier plan est venu visiter nos bureaux. Il songeait apparemment à investir dans l'entreprise, et celle-ci voulait naturellement se montrer sous son meilleur jour. Nous fûmes donc tous convoqués pour écouter ce qu'il avait à dire.

Il décrivit l'avenir radieux de la publicité ciblée. Avec les torrents de données qu'ils allaient fournir, les gens offriraient aux publicitaires la capacité d'acquérir une connaissance approfondie de leurs clients potentiels. Les entreprises auraient de ce fait la possibilité de leur adresser les informations qu'elles jugeraient utiles de manière ciblée : au bon moment et au bon endroit. Une pizzeria, par exemple, pourrait non seulement savoir que vous êtes dans le quartier mais que vous avez aussi probablement envie de la même pizza double fromage pepperoni à pâte épaisse déjà dégustée la semaine précédente à la mi-temps du match de football des Dallas Cowboys. Son système pourrait détecter que les individus dont les habitudes sont similaires aux vôtres ont davantage de chances de cliquer sur un bon de réduction durant cette fenêtre de vingt minutes.

La plus grande faiblesse de l'argument me semblait résider dans sa justification. L'investisseur soutenait que le déferlement prochain de publicités personnalisées se montrerait tellement utile et opportun que les clients l'accueilleraient à bras ouverts. Ils en réclameraient encore plus. Selon lui, la plupart des gens se plaignaient des publicités car elles ne les concernaient pas. Demain, ce ne serait plus le cas. Les personnes dépeintes

dans sa démonstration exclusive apprécieraient vraisemblablement de recevoir des messages taillés sur mesure, qui leur proposeraient peut-être des villas aux Bahamas, des bouteilles d'huile d'olive vierge pressée à la main ou des jets privés en multipropriété. Il ajouta sur le ton de la plaisanterie qu'il ne serait quant à lui jamais plus importuné par une publicité pour l'université de Phoenix – une « usine » d'enseignement à but lucratif s'adressant principalement aux plus démunis qui cherchent désespérément à s'en sortir dans la vie (souvent les plus faciles à tromper).

J'ai trouvé étrange qu'il mentionne l'université de Phoenix. Pour une raison quelconque, il voyait ces publicités, et pas moi. À moins que je ne les aie pas remarquées. Quoi qu'il en soit, j'en savais long sur les universités à but lucratif, devenues à cette époque des organismes multimillionnaires. Ces fabriques à diplômes, comme on les appelle, étaient souvent financées par des prêts d'État, et les diplômes qu'elles délivraient n'avaient guère de valeur sur le marché du travail. Dans beaucoup de métiers, ils n'étaient pas plus valorisés qu'un baccalauréat.

Tandis que l'ADM du classement des meilleures universités publié par *U.S. News* empoisonnait la vie des étudiants aisés ou issus des classes moyennes (et de leurs familles), les établissements à but lucratif se concentraient sur une autre partie de la population, plus vulnérable. Et Internet leur fournissait pour cela l'outil parfait. Rien de bien étonnant, dès lors, à ce que la croissance spectaculaire de cette industrie ait coïncidé avec l'éclosion d'Internet, réseau de communication ininterrompu à destination des masses. L'université de Phoenix, qui dépensait pour sa publicité plus de 50 millions de dollars uniquement sur Google, ciblait les personnes pauvres en utilisant l'ascenseur social comme appât. Son accroche commerciale véhiculait une critique sous-jacente des classes en difficulté, accusées de ne pas en faire assez pour améliorer leur existence. Et le procédé a fonctionné. Entre 2004 et 2014, les inscriptions dans des établissements à but lucratif ont triplé, et ce secteur représente désormais aux États-Unis 11 % du nombre total d'étudiants.

Le marketing de ces universités se trouve très éloigné de la promesse initiale d'Internet, présenté comme une grande force égalisatrice et démocratisante. « Sur Internet, personne ne sait que vous êtes un chien<sup>1</sup> » : si l'adage était vrai au début de l'ère des start-up, c'est aujourd'hui tout le contraire. On nous classe, on nous catégorise, on nous note au travers de centaines de modèles, en fonction des préférences et des schémas de

comportement que nous dévoilons. Ces données procurent une base solide pour des campagnes publicitaires légitimes, mais elles nourrissent aussi leurs cousines prédatrices : des publicités qui visent les personnes en grand besoin et leur vendent des promesses mensongères ou hors de prix. Celles-ci dénichent les inégalités et s'en repaissent. Elles perpétuent en conséquence la stratification sociale existante, avec toutes ses injustices. Le fossé le plus large sépare d'un côté les gagnants du système – comme notre investisseur en capital-risque – et de l'autre les individus qui servent de proies à ses modèles.

Partout où le besoin et l'ignorance se combinent, des publicités prédatrices ont toutes les chances d'apparaître. Si les gens sont angoissés par leur vie sexuelle, des publicitaires rapaces leur promettent du Viagra ou du Cialis, ou même une « extension du pénis ». S'ils sont à court d'argent, ce sera un déluge d'offres pour des prêts sur salaire à taux élevé. Si leur ordinateur se met à ralentir, il s'agira peut-être d'un virus injecté par l'un de ces publicitaires, qui proposera alors de régler le problème. Et comme nous allons le voir, l'explosion des universités à but lucratif est elle aussi alimentée par ce genre de publicités.

En termes d'ADM, les publicités prédatrices constituent pratiquement la définition du genre. Elles s'attaquent avec une ampleur phénoménale aux plus désespérés d'entre nous. Dans le domaine de l'éducation, elles promettent ce qui n'est en général qu'une fausse route vers la prospérité, tout en calculant comment tirer de chaque prospect le maximum d'argent. Elles génèrent par leur action de gigantesques et abominables boucles de rétroaction, et laissent leurs clients sous des montagnes de dettes. Du fait de l'opacité des campagnes, leurs cibles n'ont en outre quasiment aucune idée de la manière dont elles se font escroquer. Ces publicités surgissent simplement sur l'écran de l'ordinateur, suivies plus tard d'un appel téléphonique. Les victimes apprennent rarement comment elles ont été choisies, ou comment les recruteurs sont parvenus à en savoir autant sur elles.

Prenons l'exemple du Corinthian College, qui était encore il y a peu l'un des géants du secteur. Ses diverses sections totalisaient plus de 80 000 étudiants, dont l'immense majorité bénéficiaient d'un prêt financé par le gouvernement. En 2013, cet établissement à but lucratif fut épinglé par le procureur général de Californie pour avoir menti sur ses taux d'insertion professionnelle, surfacturé les étudiants, et utilisé dans le cadre de



publicités prédatrices de prétendus « sceaux » de l'armée afin d'allécher les personnes vulnérables. La plainte déposée soulignait qu'une de ses sections, le Brandon Campus de l'université en ligne Everest, facturait des frais de scolarité de 68 800 dollars pour une licence d'assistant juridique par Internet (alors qu'un tel cursus en coûte moins de 10 000 dans beaucoup d'établissements traditionnels).

Selon les termes de la plainte, le Corinthian College ciblait de surcroît des individus « isolés » et « impatients » dotés d'une « faible estime de soi », dont « l'entourage se souciait peu », qui se trouvaient « dans une situation de blocage » et s'avéraient « incapables d'envisager et de planifier correctement leur avenir ». Les pratiques du Corinthian College étaient qualifiées d'« illégales, injustes et frauduleuses ». En 2014, suite à de nouveaux signalements, l'administration Obama suspendit l'accès de l'établissement aux prêts étudiants fédéraux – une source vitale de financement. Mi-2015, l'entreprise revendit la plupart de ses campus et se plaça sous la protection du chapitre 11 du Code des faillites.

Mais cette industrie va toujours plus loin. Le Vatterott College, institut de formation professionnelle, offre à cet égard un exemple particulièrement terrible. En 2012, un rapport d'une commission du Sénat sur les universités à but lucratif a décrit son manuel de recrutement, dont le contenu semble diabolique. Il ordonne en effet aux recruteurs de cibler les catégories suivantes : « Mères célibataires bénéficiant de l'aide sociale. Femmes enceintes. Personnes fraîchement divorcées, souffrant d'une faible estime de soi, mal payées, victimes d'un deuil récent, ayant subi des maltraitances physiques ou psychologiques, ou sortant d'un séjour en prison. Toxicomanes en réinsertion. Individus aux perspectives bouchées occupant un emploi sans avenir. »

Pourquoi visaient-ils spécifiquement ces gens-là ? Parce que la vulnérabilité vaut de l'or. Il en a toujours été ainsi. Imaginez un charlatan itinérant dans un vieux western. Il s'arrête en ville avec son chariot bringuebalant, plein de fioles et de bouteilles. Face à une cliente potentielle d'un certain âge, il cherche à identifier ses points faibles. Elle se couvre la bouche lorsqu'elle sourit, trahissant de la sorte que sa mauvaise denture est un sujet sensible. Elle fait nerveusement tourner sa vieille alliance qui, à en juger par l'articulation enflée de son doigt, restera coincée là jusqu'à la fin de ses jours. Arthrite. Au moment de lui vanter ses produits, l'homme se concentre donc sur ses dents abîmées et sur ses mains douloureuses. Il peut

promettre de lui rendre un beau sourire et de la débarrasser de ses douleurs articulaires. Avec les connaissances qu'il a réunies, il sait que la vente est déjà à moitié conclue avant même de s'éclaircir la gorge pour entamer son discours.

Les publicitaires prédateurs utilisent des stratégies similaires, qu'ils mettent en œuvre sur une échelle gigantesque, ciblant chaque jour des millions de personnes. L'ignorance du client représente bien entendu une pièce cruciale du puzzle. Il y a parmi les étudiants visés beaucoup d'immigrés qui arrivent aux États-Unis en s'imaginant que les universités privées sont plus prestigieuses que les établissements publics. Si l'on parle de Harvard ou de Princeton, l'argument peut se défendre. Mais personne, hormis de nouveaux venus connaissant mal le système, ne pourrait jamais croire que DeVry ou l'université de Phoenix soient préférables à n'importe quelle université d'État (sans parler de joyaux publics comme les universités de Berkeley, du Michigan ou de Virginie).

Une fois l'ignorance de la cible établie, la clé pour le recruteur, à l'instar du marchand de potions miracles, consiste à repérer les gens les plus vulnérables et à exploiter à leurs dépens leurs informations personnelles. Il faut découvrir ce qui les fait le plus souffrir – ce qu'on appelle en anglais le « *pain point* », le « point de douleur ». Il pourra s'agir d'un manque d'estime de soi, de l'angoisse d'élever ses enfants dans un quartier où des gangs rivaux se font la guerre, ou peut-être d'une addiction à la drogue. Beaucoup de personnes divulguent involontairement leurs « points de douleur » quand elles cherchent des réponses sur Google ou, plus tard, quand elles remplissent des questionnaires d'inscription à l'université. Cette précieuse pépite en main, il suffit aux recruteurs de promettre qu'un cursus coûteux au sein de leur établissement fournira la solution et éliminera la douleur. « Nous avons affaire à des gens qui vivent dans et pour le moment présent, expliquent les documents de formation de Vatterott. Leur décision d'entrer à l'école, d'y rester ou de la quitter repose davantage sur l'émotion que sur la logique. La douleur constitue à court terme le plus puissant facteur de motivation. » Chez ITT Technical Institute, une équipe de recrutement est allée jusqu'à dessiner un dentiste penché sur un patient en plein martyre, avec cette formule : « Découvrez où ils ont mal. »

Le premier clic d'un étudiant potentiel sur le site Internet d'une université à but lucratif intervient uniquement après qu'un vaste processus industriel a préparé le terrain. Le Corinthian College possédait à titre

d'exemple une équipe marketing de trente personnes qui dépensait chaque année 120 millions de dollars, dont une bonne partie servait à générer et approfondir 2,4 millions de contacts, débouchant sur l'inscription de 60 000 nouveaux étudiants pour un chiffre d'affaires annuel de 600 millions de dollars. Ces énormes équipes marketing approchent les étudiants potentiels par le biais d'un large éventail de canaux – publicités télévisées, panneaux en bord de route et sur les abribus, publipostages, liens sponsorisés sur Google, voire visites de recruteurs dans les écoles et prospections en porte-à-porte. Un analyste de l'équipe conçoit ensuite diverses offres promotionnelles avec comme but explicite d'obtenir un retour. Pour optimiser le recrutement – et le chiffre d'affaires –, il doit savoir qui a reçu ses messages et, si possible, l'impact qu'ils ont eu. Ce n'est qu'au moyen de ces données qu'il peut ensuite optimiser le dispositif.

L'élément clé de tout programme d'optimisation consiste naturellement à se fixer un objectif. Dans le cas d'usines à diplômes comme l'université de Phoenix, on peut dire, sans trop s'avancer, qu'il s'agit de recruter le plus grand nombre possible d'étudiants susceptibles de décrocher un prêt d'État pour couvrir l'essentiel de leurs frais de scolarité. Avec cet objectif en tête, les experts en données doivent trouver le moyen de gérer au mieux leurs différents canaux de communication afin qu'ils produisent tous ensemble un bénéfice maximal pour chaque dollar investi.

Les experts en données commencent par une approche bayésienne, un quasi-standard dans le domaine des statistiques. L'analyse bayésienne vise à classer les variables qui possèdent le plus d'impact par rapport au résultat désiré. Les liens sponsorisés, la télévision, les affiches et les autres canaux de promotion sont évalués en fonction de leur efficacité par dollar dépensé. Ils développent chacun une probabilité différente, exprimée sous forme d'une valeur, ou d'un poids.

L'opération se voit néanmoins compliquée par le fait que les diverses campagnes de communication interagissent toutes entre elles et que leur impact respectif est en grande partie impossible à mesurer. Les publicités sur les autobus augmentent-elles par exemple la probabilité qu'un prospect réponde à un appel téléphonique ? Difficile à dire. Les messageries en ligne autorisent un suivi plus aisé, et permettent aux universités à but lucratif de recueillir des informations capitales sur chaque prospect – l'endroit où il vit et les pages web qu'il a consultées.

Voilà pourquoi ces établissements consacrent l'essentiel de leurs budgets publicitaires à Google et Facebook. Ces deux plateformes permettent aux annonceurs de segmenter de manière extrêmement détaillée les publics ciblés. L'agence chargée de la promotion d'un film de Judd Apatow pourrait par exemple cibler les hommes de 18 à 28 ans situés dans les cinquante localités les plus riches, en se focalisant peut-être sur ceux qui ont cliqué ou « liké » des liens vers l'un de ses grands succès, *40 Ans, toujours puceau*, qui ont cité son nom sur Twitter, ou qui sont amis avec quelqu'un qui l'a mentionné. Mais les universités à but lucratif chassent dans la direction opposée. Elles viseront plus probablement les populations des localités les plus pauvres, en accordant une attention particulière à ceux qui ont cliqué sur une publicité pour des prêts sur salaire, ou qui semblent concernés par le syndrome du stress post-traumatique (elles recrutent beaucoup parmi les anciens combattants, notamment parce qu'il est plus facile de leur obtenir une aide financière).

La campagne met ensuite en œuvre une succession sans fin de publicités concurrentes, afin de déterminer celles qui ramènent le plus de prospects. Cette méthode, fondée sur ce qu'on appelle des tests A/B, est utilisée depuis des décennies par les spécialistes du publipostage. Ils expédient une multitude d'accroches marketing, mesurent les réactions et ajustent leurs campagnes. Chaque fois que vous découvrez une nouvelle offre de carte de crédit dans votre boîte aux lettres, cela signifie que vous participez à l'un de ces tests. En jetant le courrier sans même l'ouvrir, vous fournissez à l'entreprise une donnée précieuse : cette campagne n'a pas fonctionné sur vous. La prochaine fois, ils tenteront une approche légèrement différente. L'effort pourrait sembler vain, puisque la majorité de ces offres finissent à la poubelle. Mais pour beaucoup de praticiens du marketing direct, qu'ils opèrent sur Internet ou par courrier, un taux de réponse de 1 % est un eldorado. Ils travaillent en effet sur des quantités énormes : 1 % de la population des États-Unis représente plus de 3 millions de personnes.

Dès lors que ces campagnes ont basculé sur Internet, l'apprentissage s'en est trouvé accéléré. Le Web offre aux annonceurs le plus grand laboratoire de tous les temps pour l'étude des consommateurs et la production de contacts. Les remontées d'informations de chaque promotion leur parviennent en quelques secondes – bien plus vite qu'avec le courrier. En l'espace de quelques heures (au lieu de plusieurs mois), chaque campagne peut se resserrer sur les messages les plus efficaces et approcher

l'étincelante promesse de toute publicité : joindre un prospect au bon moment, en lui présentant précisément le meilleur message de nature à déclencher une décision, et réussir de la sorte à prendre dans ses filets un nouveau client rémunérateur. Ce processus d'ajustement ne s'arrête jamais.

Et de plus en plus, les machines qui traitent et analysent l'information passent toutes seules nos données au crible, à la recherche de nos habitudes et de nos espoirs, de nos peurs et de nos désirs. Grâce au *machine learning* (l'« apprentissage automatique », une branche de l'intelligence artificielle aujourd'hui en forte croissance), l'ordinateur plonge dans les données en suivant des instructions élémentaires. L'algorithme identifie par lui-même certains schémas, puis avec le temps les rattache à tels ou tels résultats. En un sens, il apprend.

Comparé au cerveau humain, le *machine learning* n'est pas particulièrement efficace. Lorsqu'un enfant pose son doigt sur la cuisinière, il ressent une souffrance et assimile pour le restant de sa vie la corrélation entre la chaleur du métal et la douleur lancinante dans sa main. Il apprend aussi le terme qui lui correspond : une brûlure. Pour élaborer des modèles statistiques reliant causes et effets, un programme de *machine learning* aura souvent besoin en revanche de plusieurs millions ou milliards de données distinctes. Mais pour la première fois dans l'histoire, ces pétaoctets de données sont désormais facilement accessibles, et on dispose aussi d'ordinateurs puissants permettant de les traiter. Pour nombre de tâches, le *machine learning* s'avère de surcroît plus souple et plus nuancé que les programmes traditionnels pilotés par un ensemble de règles.

Les linguistes se sont par exemple efforcés pendant des décennies, depuis les années 1960 jusqu'au début de notre siècle, d'apprendre à lire aux ordinateurs. Ils ont passé la majeure partie de ce temps à programmer des définitions et des règles grammaticales. Or, comme le découvre bien vite n'importe quel étudiant en langues étrangères, celles-ci regorgent d'exceptions. Elles contiennent des termes d'argot et autorisent le sarcasme. Selon l'époque et le lieu, le sens de certains mots évolue. La complexité de la langue est le cauchemar d'un programmeur. Son codage informatique n'a en définitive aucune chance d'aboutir.

Mais avec Internet, les gens ont produit aux quatre coins de la planète des millions de milliards de mots, évoquant leur vie, leur travail, leurs achats et leurs liens d'amitié. Sans le vouloir, nous avons ainsi créé pour le traitement du langage naturel le plus vaste corpus d'entraînement jamais vu. En

passant du papier aux e-mails et aux réseaux sociaux, nous avons permis aux machines d'étudier les termes employés, de les comparer à d'autres, et de s'informer de leur contexte d'utilisation. Des progrès rapides et spectaculaires ont été réalisés. En 2011, avec Siri, son « assistant personnel », Apple n'avait guère impressionné la majorité de la sphère technologique. Le système n'était au courant que de certains domaines, et commettait des erreurs ridicules. La plupart de mes connaissances le jugeaient à peu près inutile. Mais aujourd'hui, j'entends des gens s'adresser à leur téléphone en permanence, pour s'enquérir de la météo, des résultats sportifs ou du chemin à emprunter. Quelque part entre 2008 et 2015, plus ou moins, les compétences linguistiques des algorithmes ont progressé de la première année de maternelle au niveau collège, et bien davantage encore pour certaines applications.

Ces progrès en matière de langage naturel représentent une véritable manne pour les publicitaires. Les programmes « savent » ce qu'un mot signifie, du moins suffisamment pour l'associer dans certains cas à des comportements et à des résultats donnés. Grâce en partie à cette maîtrise linguistique croissante, les publicitaires peuvent explorer des schémas plus profonds. Un programme pourra s'appuyer au départ sur les habituels renseignements démographiques et géographiques. Mais au fil des semaines et des mois, il intégrera peu à peu les schémas comportementaux des personnes ciblées et commencera à prédire leurs prochains actes. Il apprendra à les connaître. Et si le programme présente un caractère prédateur, il évaluera leurs faiblesses et leurs vulnérabilités puis empruntera la voie la plus efficace pour les exploiter.

En complément des techniques informatiques de pointe, les publicitaires prédateurs travaillent souvent avec des intermédiaires, qui emploient dans leur ciblage des méthodes bien moins raffinées. En 2010, une publicité percutante a ainsi été publiée ; elle comportait une photo du président Obama, accompagnée du texte suivant : « Obama demande aux mères de retourner à l'école : terminez vos études ! – aide financière disponible pour les personnes éligibles. » L'annonce laissait entendre que le président avait signé une nouvelle loi visant à ramener les mères de famille sur les bancs de l'école. C'était un mensonge. Mais s'il incitait les gens à cliquer, l'objectif était donc atteint.

Derrière cet intitulé trompeur, toute une industrie répugnante était à l'œuvre. Selon une enquête de l'organisme ProPublica, lorsqu'une cliente

cliquait sur cette publicité, on lui posait quelques questions – âge et numéro de téléphone, notamment – et elle était immédiatement contactée par un établissement à but lucratif. Ses interlocuteurs ne lui fournissaient pas davantage d'informations sur la nouvelle loi d'Obama, puisque celle-ci n'avait jamais existé. Ils se proposaient plutôt de l'aider à emprunter de l'argent pour financer son inscription.

On appelle « génération de contacts » ou « lead-gen » ce genre de ciblage en ligne. Le but est de constituer des listes de prospects susceptibles d'être revendues – dans le cas présent, à des universités à but lucratif. D'après l'article de ProPublica, ces établissements y consacrent 20 % à 30 % de leurs budgets promotionnels, et déboursent pour les contacts les plus prometteurs jusqu'à 150 dollars pièce.

Selon David Helperin, chercheur en politiques publiques, l'un de ces générateurs de contacts – l'entreprise Neutron Interactive, basée à Salt Lake City – postait ainsi de fausses offres d'emploi sur des sites comme Monster.com et diffusait également des publicités promettant de faciliter l'obtention de bons d'alimentation et d'une couverture de santé Medicaid<sup>2</sup>. Faisant usage des mêmes méthodes d'optimisation, elle lançait tout un tas d'annonces différentes et mesurait leur efficacité auprès de chaque groupe démographique.

Ces publicités avaient pour objectif de leurrer les demandeurs d'emploi, en proie au désespoir, afin qu'ils fournissent leur numéro de portable. Lorsqu'on les rappelait, seuls 5 % des gens se montraient intéressés par une formation universitaire. Mais ces noms représentaient d'appréciables contacts, qui pouvaient valoir aux yeux des universités à but lucratif jusqu'à 85 dollars chacun. Et elles faisaient tout ce qui était en leur pouvoir pour rentabiliser cet investissement. D'après un rapport du Government Accountability Office (équivalent de la Cour des comptes aux États-Unis), une fois qu'ils s'étaient enregistrés, les étudiants potentiels pouvaient s'attendre à recevoir sous cinq minutes les premiers appels. En l'espace d'un mois, l'une de ces cibles marketing en avait reçu plus de 180.

Pour générer des contacts, les universités à but lucratif possèdent bien entendu leurs propres méthodes. Parmi les outils les plus précieux dont elles disposent, on trouve le site Internet du College Board, une ressource qu'utilisent beaucoup d'étudiants pour s'inscrire aux tests SAT et se renseigner sur l'étape suivante de leur vie scolaire. D'après Mara Tucker, conseillère d'orientation à l'Urban Assembly Institute of Math and Science

for Young Women (une école publique de Brooklyn), le moteur de recherche du site est conçu de manière à diriger les étudiants pauvres vers des universités à but lucratif. Dès que l'élève indique, en réponse à un questionnaire en ligne, qu'il aura besoin d'une aide financière, ces établissements émergent en tête de liste des choix recommandés.

Les universités à but lucratif offrent également des services gratuits en échange d'un face-à-face direct avec les étudiants. Cassie Magesis, une autre conseillère de l'Urban Assembly, m'a indiqué qu'elles proposent des ateliers gratuits pour guider les élèves dans la rédaction de leur curriculum vitae. Ces sessions leur sont certes profitables, mais les élèves pauvres qui fournissent leurs coordonnées font ensuite l'objet d'une véritable traque. Les établissements à but lucratif ne prennent pas la peine de cibler les étudiants riches. Ceux-là, tout comme leurs parents, en savent déjà trop.

Le recrutement sous toutes ses formes est au cœur de l'activité lucrative, et constitue dans la plupart des cas une part des dépenses bien plus grande que l'enseignement. Un rapport du Sénat portant sur trente organismes à but lucratif a établi qu'ils employaient un recruteur pour quarante-huit étudiants. Apollo Group, maison mère de l'université de Phoenix, a dépensé en 2010 plus de 1 milliard de dollars pour son marketing, focalisé presque en totalité sur le recrutement. Ce qui donnait 2 225 dollars par étudiant pour le marketing, et seulement 892 dollars pour l'enseignement. À comparer au Portland Community College, dans l'Oregon, qui dépense 5 953 dollars par étudiant pour l'enseignement et environ 1,2 % de son budget, soit 185 dollars par étudiant, pour le marketing.

\*\*\*

Les mathématiques, sous l'aspect de modèles complexes, alimentent la publicité prédatrice qui génère des prospects pour ces établissements. Mais dès lors qu'un recruteur harcèle des étudiants potentiels sur leur téléphone portable, on quitte l'univers des nombres. Les argumentaires de vente, promettant une scolarité abordable, de brillantes perspectives professionnelles et l'ascension sociale, ne sont pas si différents des promotions pour des élixirs magiques, des traitements contre la chute des cheveux, ou des ceintures vibrantes censées réduire la graisse abdominale. Ils n'ont rien de nouveau.



Cependant, l'un des aspects cruciaux d'une ADM tient aux dégâts qu'elle provoque dans la vie d'un grand nombre de personnes. Et avec ce genre de publicités prédatrices, les dégâts véritables n'apparaissent que lorsque les étudiants commencent à contracter d'énormes prêts pour payer leur scolarité.

La règle dite du 90/10, incluse dans la loi sur l'enseignement supérieur (Higher Education Act) votée en 1965, représente un indicateur essentiel en la matière. Elle stipule que les universités ne peuvent être financées à plus de 90 % par l'aide fédérale. L'idée était que les étudiants, du moment où une partie des coûts serait à leur charge, auraient tendance à prendre leurs études plus au sérieux. Mais les établissements à but lucratif ont rapidement intégré ce ratio dans leur business plan. Si les étudiants sont en mesure de rassembler quelques milliers de dollars, que ce soit par leur épargne ou au moyen de prêts bancaires, les universités peuvent alors solliciter neuf fois cette somme en prêts d'État, ce qui rend chaque élève incroyablement rentable.

Aux yeux de nombreux étudiants, les prêts ressemblent à de l'argent gratuit, et l'école ne fait rien pour corriger cette méprise. Or, il s'agit d'une dette, qui finit vite par submerger beaucoup d'entre eux. Pour les étudiants du Corinthian College, déclaré en faillite, la dette en suspens s'élevait à 3,5 milliards de dollars. Garantie presque entièrement par les contribuables, elle ne sera jamais remboursée.

Certaines personnes qui partent étudier dans une université à but lucratif en tirent sans aucun doute un savoir et des compétences qui s'avèrent très utiles. Mais réussissent-elles mieux que les diplômés d'un établissement de premier cycle (appelé aux États-Unis « *community college* »), dont les cursus ne coûtent qu'une petite fraction de ceux des universités à but lucratif ? En 2014, des enquêteurs de l'institut de recherche CALDER/American Institutes for Research ont rédigé près de 9 000 CV fictifs. Parmi ces faux candidats à l'embauche, certains étaient titulaires d'un titre de premier cycle délivré par une université à but lucratif, d'autres détenaient un diplôme similaire décerné par un *community college*, tandis qu'un troisième groupe n'avait jamais fait d'études supérieures. Les chercheurs ont envoyé ces CV en réponse aux offres d'emploi publiées dans sept grandes villes, et ont mesuré le taux de réponse. Ils ont découvert que les diplômes des universités à but lucratif possédaient moins de valeur sur le marché du travail que ceux des *community colleges*, et à peu près autant

qu'un baccalauréat. Et pourtant, ces établissements coûtaient en moyenne 20 % plus cher que les fleurons de l'université publique.

La boucle de rétroaction de cette ADM s'illustre davantage par son caractère destructeur que par sa complexité. Aux États-Unis, les 40 % les plus pauvres de la population sont aux abois. Beaucoup d'emplois industriels ont disparu, remplacés par la technologie ou délocalisés à l'étranger. Les syndicats n'ont plus d'impact. 20 % des gens (le quintile supérieur) contrôlent 89 % de la richesse nationale, tandis qu'en bas de l'échelle 40 % n'en possèdent rien. Leur patrimoine est négatif : dans cet immense sous-prolétariat en grande difficulté, les ménages entretiennent une dette nette moyenne de 14 800 dollars, dont l'essentiel par le biais de cartes de crédit aux taux exorbitants. C'est d'argent que ces gens ont besoin. Et la clé pour en gagner davantage, s'entendent-ils répéter sans cesse, consiste à suivre des études.

Surgissent alors les universités à but lucratif, qui ciblent et escroquent les plus démunis au moyen d'ADM extrêmement perfectionnées. Elles leur vendent la promesse d'une formation supérieure et l'alléchante vision de l'ascension sociale – mais les enfoncent plus encore dans l'endettement. Elles tirent profit du besoin pressant dans lequel se débattent les foyers pauvres, ainsi que de leur ignorance et de leurs aspirations, puis exploitent la situation. Et elles opèrent de la sorte sur une échelle considérable – ce qui engendre découragement et désespoir, en même temps qu'un sentiment de scepticisme quant à la valeur plus générale des études, et aggrave encore le vaste différentiel de richesse du pays.

Il convient de noter que ces usines à diplômes creusent les inégalités dans les deux sens. Les présidents des universités à but lucratif les plus importantes empochent chaque année des millions de dollars. Gregory W. Cappelli, PDG d'Apollo Education Group, la maison mère de l'université de Phoenix, a perçu par exemple en 2011 une rémunération totale de 25,1 millions de dollars. Dans les établissements publics, qui connaissent sur ce plan leurs propres exubérances, seuls les entraîneurs de football américain ou de basket peuvent espérer gagner autant.

Les universités à but lucratif sont hélas loin d'être seules à faire usage de publicités prédatrices. Elles ne manquent pas de compagnie. Là où les gens sont en souffrance, ou acculés au désespoir, vous verrez des publicitaires appliquer leurs modèles prédateurs. L'une des plus grosses aubaines concerne naturellement les prêts financiers. Tout le monde a besoin

d'argent, mais certains de manière plus urgente que d'autres. Ces gens-là ne sont pas difficiles à trouver. Les plus démunis ont davantage de chances d'habiter dans une localité pauvre. Et du point de vue d'un publicitaire prédateur, compte tenu de leurs requêtes sur les moteurs de recherche et des coupons de réduction sur lesquels ils cliquent, ils supplient même, pour ainsi dire, qu'on veuille bien leur accorder une attention particulière.

À l'instar des universités à but lucratif, l'industrie du prêt sur salaire, ou *payday loan*, utilise des ADM. Certaines sont mises en œuvre par des entreprises légales, mais c'est surtout un secteur d'activité foncièrement prédateur qui facture des taux d'intérêt scandaleux : 574 %, en moyenne, sur des prêts à court terme reconduits en moyenne huit fois – ce qui les transforme plutôt en prêts à long terme. Elles bénéficient de l'apport crucial d'une armée de courtiers en données et de générateurs de contacts, dont beaucoup sont de véritables escrocs. Leurs publicités jaillissent sur les écrans d'ordinateurs et de téléphones, et proposent d'obtenir rapidement de l'argent. Dès qu'ils remplissent ce genre de demande, en fournissant souvent leurs coordonnées bancaires, les prospects s'exposent au vol et aux abus.

En 2015, la Federal Trade Commission (Commission fédérale du commerce, chargée de la répression des fraudes) a poursuivi deux courtiers en données pour la revente des demandes de prêt de plus d'un demi-million de consommateurs. D'après l'acte d'accusation, les sociétés incriminées – Sequoia One, basée à Tampa en Floride, et Gen X Marketing Group, située non loin, à Clearwater – avaient dérobé leur numéro de téléphone, le nom de leur employeur, leur numéro de Sécurité sociale et les références de leur compte bancaire, et vendu ces informations environ 50 *cents* chacune. Les entreprises acheteuses s'étaient emparées, selon l'organe de régulation, d'« au moins » 7,1 millions de dollars puisés dans les comptes des clients. Nombre de leurs victimes s'étaient vu ensuite facturer des frais bancaires, au motif qu'elles avaient vidé leur compte ou émis des chèques sans provision.

Si l'on y réfléchit, les montants dont il est question sont ridiculement faibles. Répartis sur un demi-million de comptes, 7,1 millions de dollars représentent pour chacun un peu plus de 14 dollars. Même si beaucoup de comptes sont restés inaccessibles aux voleurs, l'essentiel de l'argent a sans doute été dérobé sous forme de toutes petites sommes – en ponctionnant les

derniers 50 ou 100 dollars que certaines personnes démunies conservent sur leur compte.

Les instances de régulation préconisent aujourd'hui l'adoption de nouvelles lois encadrant le marché des données personnelles – une source essentielle pour toutes sortes d'ADM. À ce jour, une poignée de textes fédéraux tels que le Fair Credit Reporting Act et le Health Insurance Portability and Accountability Act (ou HIPAA) fixent certaines limites sur la collecte des données ayant trait à la santé et au crédit. D'autres viendront peut-être s'y ajouter, avec en ligne de mire les générateurs de contacts.

Comme nous le verrons dans les chapitres suivants, certaines ADM parmi les plus efficaces et les plus nocives parviennent cependant à contourner ces obstacles. Elles étudient tout, du quartier où nous vivons jusqu'aux amis que nous côtoyons sur Facebook, et s'en servent pour prédire notre comportement – voire pour nous jeter derrière les barreaux.

---

1. « *On the Internet, nobody knows you're a dog.* » Adage des débuts d'Internet, tiré d'un dessin de presse de 1993 devenu viral, où un chien prononçait ces mots tout en surfant sur le Web.

2. Assurance maladie réservée aux personnes à faibles revenus.

## La justice à l'heure du Big Data

# VICTIMES CIVILES

1001Ebooks

**L**a petite ville de Reading, en Pennsylvanie, a connu des temps difficiles à l'ère postindustrielle. Nichée au milieu de collines verdoyantes à 80 kilomètres à l'ouest de Philadelphie, elle s'était enrichie grâce aux chemins de fer, à l'acier, au charbon et au textile. Mais au cours des dernières décennies, avec le déclin brutal de toutes ces industries, la ville a dé péri. En 2011, elle affichait le taux de pauvreté le plus élevé du pays (41,3 %), avant d'être dépassée de peu l'année suivante par Detroit. Quand la récession est venue frapper l'économie de Reading, à la suite du krach financier de 2008, ses recettes fiscales ont chuté, entraînant le licenciement de quarante-cinq agents de police, en dépit d'une criminalité persistante.

Malgré un effectif réduit, William Heim, le chef de la police de Reading, dut trouver le moyen d'assurer un maintien de l'ordre d'aussi bonne qualité, sinon meilleur. En 2013, il décida donc d'investir dans un logiciel de prédiction criminelle conçu par PredPol, une start-up du Big Data basée à Santa Cruz, en Californie. Le programme analysait les données historiques en matière de délinquance et calculait, heure par heure, les lieux dans lesquels des délits avaient la plus grande probabilité d'être perpétrés. Les policiers de Reading pouvaient visualiser les conclusions du logiciel qui se présentaient sous l'aspect d'un ensemble de zones carrées équivalant chacune à la surface de deux terrains de football. S'ils patrouillaient plus assidument dans ces secteurs-là, ils avaient de bonnes chances de prévenir les actes délictueux. Et de fait, au bout d'un an, le commandant Heim annonça que les cambriolages avaient reculé de 23 %.

Partout aux États-Unis, les programmes prédictifs du type PredPol font aujourd'hui fureur au sein de services de police étranglés par les restrictions budgétaires<sup>1</sup>. D'Atlanta à Los Angeles, les forces de l'ordre déploient leurs

agents dans les carrés qui apparaissent çà et là, puis signalent des taux de criminalité en baisse. La ville de New York met en œuvre un outil similaire baptisé CompStat. La police de Philadelphie utilise quant à elle un produit local nommé HunchLab, comprenant une analyse des zones à risque qui tient compte de certaines particularités, telles que la présence de distributeurs de billets ou de petits commerces, susceptibles d'attirer les délinquants. Comme n'importe où ailleurs dans l'industrie du Big Data, les développeurs des logiciels de prédiction criminelle s'empressent d'intégrer toutes les informations capables d'améliorer la précision de leurs modèles.

Si l'on y réfléchit bien, les outils de prédiction des « points chauds » fonctionnent de manière analogue aux modèles de déplacement défensif utilisés dans le base-ball, décrits plus haut. Ces systèmes examinent l'historique des coups de batte de chaque frappeur, et positionnent les joueurs de champ là où la balle a le plus de chances d'aller. Les logiciels de prédiction criminelle effectuent une analyse similaire, et positionnent les policiers là où les délits ont le plus de chances de survenir. Les deux types de modèles optimisent les ressources. Mais parmi les seconds, certains sont plus sophistiqués ; en fonction des recrudescences d'actes, ils prédisent d'éventuelles vagues de criminalité. PredPol, par exemple, s'inspire d'un logiciel de sismologie : il étudie tel ou tel délit perpétré dans une zone donnée, l'incorpore à ses schémas historiques, et prédit ensuite à quel moment et à quel endroit ce délit pourrait être à nouveau commis (l'une des corrélations très simples ainsi découvertes montre que si votre voisin se fait cambrioler, vous avez tout intérêt à verrouiller solidement vos portes et fenêtres).

Les modèles de prédiction criminelle comme PredPol ne sont pas dénués de vertus. Contrairement aux agents d'interpellation du film dystopique *Minority Report*, de Steven Spielberg (et à quelques initiatives inquiétantes dans le monde réel, que nous aborderons d'ici peu), les policiers ne poursuivent pas les gens avant qu'ils ne se soient rendus coupables d'un crime. Jeffrey Brantingham, professeur d'anthropologie à l'UCLA (Université de Californie à Los Angeles) et fondateur de PredPol, m'a bien précisé que son modèle était indifférent aux origines ethniques. En outre, à l'inverse d'autres programmes – notamment des modélisations du risque de récidive, dont nous avons parlé, qui servent de guide pour fixer les peines de prison –, PredPol ne se concentre pas sur l'individu mais sur la localisation géographique. Le type, le lieu, la date et l'heure de chaque

crime constituent les données essentielles à partir desquelles il travaille. Voilà qui semble de bonne guerre. Et si les agents de police passent plus de temps dans les zones à haut risque, afin de dissuader les cambrioleurs et les voleurs de voitures, on a de bonnes raisons de penser que la collectivité en bénéficiera.

Cela étant, la plupart des délits sont d'une gravité moindre, et c'est là que de sérieux problèmes apparaissent. Quand la police installe son système PredPol, elle doit faire un choix. Elle a la possibilité de se focaliser exclusivement sur les crimes dits « Part 1 ». Il s'agit des crimes violents, tels que les homicides, les incendies volontaires et les agressions physiques, qui lui sont généralement signalés. Mais elle peut aussi élargir le spectre en incluant les délits « Part 2 », qui englobent le vagabondage, la mendicité agressive et la vente ou la consommation de petites quantités de drogues. Si un policier n'était pas là pour les constater, ces délits de « nuisance publique » passeraient pour beaucoup inaperçus.

Les infractions de cet ordre, parfois qualifiées par les forces de l'ordre de « comportements antisociaux », présentent un caractère endémique dans nombre de quartiers déshérités. Malheureusement, le fait de les intégrer au modèle menace de fausser l'analyse. Lorsque des données relatives aux nuisances publiques viennent alimenter un modèle prédictif, davantage d'agents sont dirigés vers ces quartiers, où ils ont des chances d'effectuer plus d'arrestations. Car après tout, même si leur objectif est d'empêcher les cambriolages, les meurtres et les viols, il y a forcément des périodes creuses : c'est la nature même du travail de patrouille policière. Et si au cours d'une de ces rondes un policier aperçoit deux gamins, qui ne semblent pas avoir plus de 16 ans, en train de vider à grands traits une bouteille dissimulée dans un sac en papier kraft, il les arrêtera. Les petits délits de ce genre ajoutent de plus en plus de points à la cartographie des modèles, conduisant ces derniers à renvoyer les policiers dans les mêmes quartiers.

Il se crée ainsi une boucle de rétroaction pernicieuse. Le maintien de l'ordre génère par lui-même de nouvelles données, qui justifient en retour son renforcement dans certaines zones. Et nos prisons se remplissent de centaines de milliers d'individus déclarés coupables de crimes sans victime. La plupart viennent de quartiers déshérités, et sont en majorité noirs ou hispaniques. Même si un modèle se montre indifférent à la couleur de peau, il en va donc tout autrement du résultat obtenu. Aux États-Unis, où les

ville sont fortement marquées par la ségrégation, la localisation géographique se substitue très efficacement aux données raciales.

Si les modèles ont pour but de prévenir les crimes les plus graves, on pourrait se demander pourquoi les simples petits délits font l'objet du moindre suivi. La réponse, c'est que le lien entre comportement antisocial et délinquance tient lieu de dogme depuis qu'un criminologue du nom de George Kelling s'est associé en 1982 avec un expert en politiques publiques, James Q. Wilson, pour publier dans le magazine *Atlantic Monthly* un article déterminant sur la théorie dite des « fenêtres cassées ». L'idée était que dans un quartier les petits délits et les incivilités créaient un climat de désordre qui faisait fuir les citoyens respectueux des lois. Ils laissaient derrière eux des rues sombres et vides – un terreau propice à la grande délinquance. L'antidote consistait pour la collectivité à résister à la propagation du désordre, ce qui supposait notamment de réparer les fenêtres cassées, de nettoyer les rames de métro couvertes de graffitis, et de prendre des mesures afin de dissuader les délits mineurs.

Ce raisonnement entraîna dans les années 1990 l'adoption de campagnes de tolérance zéro, dont l'exemple le plus célèbre fut celui de la ville de New York. Les policiers arrêtaient les jeunes qui sautaient par-dessus les tourniquets du métro. Ils appréhendaient les individus surpris à partager un simple joint de cannabis, et leur faisaient faire le tour de la ville pendant des heures dans un fourgon cellulaire avant finalement de les verbaliser. Certains attribuèrent à ces énergiques campagnes le mérite d'une chute spectaculaire des crimes violents. D'autres furent d'un avis contraire. Les auteurs du best-seller *Freakonomics* allèrent jusqu'à tracer une corrélation entre le recul de la criminalité et la légalisation de l'avortement dans les années 1970. De nombreuses autres théories furent également avancées, depuis la baisse des taux d'addiction au crack jusqu'à l'essor économique des années 1990. Quoi qu'il en soit, le mouvement en faveur de la tolérance zéro s'attira un large soutien, et le système judiciaire jeta en prison des millions de gens – pour la plupart des jeunes hommes issus des minorités, dont beaucoup n'étaient auteurs que d'infractions légères.

Mais la tolérance zéro n'avait en réalité pas grand rapport avec la thèse des « fenêtres cassées » de Kelling et Wilson. Leur étude de cas s'était focalisée sur ce qui apparaissait en matière de police comme une initiative couronnée de succès, menée dans la ville de Newark, au New Jersey. Selon le programme mis en place, les agents qui arpentaient ses rues étaient



censés faire preuve d'une *extrême* tolérance. Ils avaient mission de s'adapter aux critères de maintien de l'ordre propres à chaque quartier, et de contribuer à les faire respecter. Ces critères variaient d'un secteur de la ville à l'autre. Dans un quartier donné, cela pouvait signifier que les ivrognes avaient obligation de conserver leurs bouteilles dans un sac et d'éviter les artères principales, mais que les petites rues adjacentes leur étaient permises. Ou que les toxicomanes avaient le droit de s'asseoir sur les perrons mais pas de s'y allonger. L'idée était uniquement de s'assurer que les critères ne baissent pas. Dans le cadre de ce plan, les policiers aidaient donc le quartier à maintenir son ordre à lui, sans toutefois imposer le leur.

Vous pensez peut-être que je m'éloigne un peu de PredPol, des mathématiques et des ADM. Chaque manière d'aborder l'activité policière, de la théorie des fenêtres cassées à la tolérance zéro, constitue pourtant un modèle. Au même titre que mon système de planification des repas ou que le classement des meilleures universités établi par *U.S. News*, chaque modèle de lutte contre la criminalité exige qu'on lui apporte certaines données, qui donnent lieu à une série de réponses, et chacun se trouve calibré dans l'optique d'atteindre un objectif précis. Envisager le travail de police sous cet angle s'avère important, car ces modèles mathématiques occupent désormais une place prépondérante dans le domaine du respect de l'ordre. Et certains d'entre eux sont des ADM.

Ceci étant posé, on peut alors comprendre que des services de police choisissent de prendre en compte les données concernant les nuisances. Nourris à l'orthodoxie de la tolérance zéro, beaucoup font la même corrélation entre petite et grande délinquance qu'entre fumée et feu. Quand la police du Kent, en Angleterre, a essayé PredPol en 2013, elle a intégré à son modèle les données sur les nuisances publiques. Le logiciel semblait bien fonctionner : elle constata que les carrés de PredPol étaient dix fois plus efficaces que des patrouilles aléatoires, et deux fois plus précis que les analyses livrées par le renseignement policier. Et quel type de délits le modèle prédisait-il le mieux ? Les petites nuisances. Voilà qui est on ne peut plus logique. Jour après jour, un ivrogne urindra en effet sur le même mur, et un toxicomane s'étendra sur le même banc, tandis qu'un voleur de voitures ou un cambrioleur se déplacera sans arrêt, en s'efforçant d'anticiper les mouvements des forces de l'ordre.

Alors même que les commandants de police mettent l'accent sur la lutte contre les crimes violents, il leur faudrait faire preuve d'une remarquable

retenue pour ne pas injecter dans leurs modèles prédictifs ces tonnes de données relatives aux nuisances. On croit volontiers que la quantité est synonyme de qualité : quand un modèle focalisé uniquement sur les crimes violents ne produit sur l'écran qu'une constellation de points épars, l'ajout des données touchant aux nuisances dresse en revanche un tableau plus frappant et plus complet du non-respect des lois dans la ville.

Et dans la plupart des circonscriptions, hélas, une telle cartographie criminelle se superpose à celle de la pauvreté. Le nombre élevé d'arrestations dans ces zones-là ne fait donc que confirmer la thèse, largement partagée dans les classes moyennes et supérieures de la société, selon laquelle les pauvres sont responsables de leurs propres échecs et commettent l'essentiel des délits.

Mais qu'en serait-il si la police s'intéressait à d'autres types d'actes ? L'idée pourrait sembler contre-intuitive, car la plupart des gens, y compris les policiers, envisagent la criminalité sous la forme d'une pyramide. En haut les homicides, suivis des viols et des agressions, plus courants, puis des vols à l'étalage, des menus larcins et même des infractions au stationnement, qui surviennent en permanence. Il paraît logique de s'attaquer en priorité aux délits situés au sommet de la pyramide. Selon l'avis majoritaire, la réduction des crimes violents constitue, et doit constituer, un aspect central de la mission des forces de l'ordre.

*Quid* cependant de certains délits très éloignés des cases prévues sur les cartes de PredPol – ceux perpétrés par les riches ? Dans les années 2000, les rois de la finance s'en sont donné à cœur joie. Ils ont menti, ils ont parié des milliards de dollars contre leurs propres clients, ils ont fraudé et graissé la patte aux agences de notation. Des crimes colossaux ont été commis, avec pour résultat de dévaster l'économie mondiale pendant près de cinq ans. Des millions de personnes ont perdu leur logement, leur emploi et leur couverture médicale.

Nous avons toute raison de croire que ce genre de délits se perpétue à l'heure actuelle dans le monde financier. S'il y a bien une chose que nous ayons apprise, c'est que l'objectif moteur de la finance consiste à réaliser d'énormes profits, les plus importants possible, et que tout ce qui peut ressembler à de l'autorégulation s'avère sans effet. Du fait en grande partie de sa richesse et de la puissance des lobbies, cette industrie est sous-contrôlée.

Imaginez seulement que la police applique à l'univers de la finance sa stratégie de tolérance zéro. Elle arrêterait les individus coupables même des infractions les plus légères, comme avoir arnaqué des investisseurs sur leur épargne-retraite, les avoir sciemment mal conseillés, ou avoir commis de petits truandages. Des groupes d'intervention d'élite policière débarqueraient peut-être en force à Greenwich, dans le Connecticut, capitale des fonds spéculatifs. Les forces de l'ordre infiltreraient aussi les bars autour du Chicago Mercantile Exchange, l'un des principaux marchés à terme des États-Unis.

C'est bien entendu peu probable. Les policiers ne possèdent pas l'expertise nécessaire pour ce genre de missions. De la formation reçue jusqu'au gilet pare-balles, tout dans leur métier est adapté à la rue. La répression de la criminalité des cols blancs réclamerait des personnels dotés de compétences et d'outils différents. Les équipes aux effectifs restreints et aux budgets insuffisants chargées de ce travail, du FBI aux enquêteurs de la Securities and Exchange Commission (le gendarme de la Bourse, équivalent de l'AMF française), ont appris au fil des décennies que les banquiers étaient pratiquement invulnérables. Ils subventionnent copieusement nos hommes politiques – ce qui est toujours utile –, et on les considère en outre comme d'une importance cruciale pour notre économie. Cette situation les protège. Si leurs établissements font faillite, l'économie des États-Unis risque de sombrer avec eux. Les pauvres, eux, ne disposent pas d'un tel argument. Ainsi, à l'exception d'un ou deux cas isolés comme Bernard Madoff, champion des chaînes de Ponzi<sup>2</sup>, les financiers ne se font jamais arrêter. En 2008, ils sont sortis tous ensemble à peu près indemnes de l'effondrement des marchés. Que pourrait-il bien leur arriver maintenant ?

Ce que je veux souligner, c'est que la police opère des choix et décide où porter son attention. Aujourd'hui, elle se focalise quasi exclusivement sur les pauvres. C'est son héritage, et c'est aussi la manière dont elle conçoit sa mission. Or, des experts en données agrègent désormais ce statu quo de l'ordre social au sein de modèles comme celui de PredPol, qui possèdent une emprise sans cesse croissante sur nos vies.

Si PredPol fournit un outil logiciel tout à fait utile, qui poursuit même de nobles ambitions, la conséquence en est qu'il s'agit aussi d'une ADM « clés en main ». Même avec les meilleures intentions du monde, PredPol procure aux services de police la faculté de cibler les pauvres, d'en interpellier

d'avantage, d'arrêter une partie d'entre eux et d'en envoyer un certain nombre en prison. Et dans beaucoup de cas, sinon la plupart, les officiers de commandement estiment qu'ils suivent de la sorte la seule voie judiciaire pour combattre la criminalité. C'est là qu'elle sévit, affirment-ils, pointant du doigt les ghettos mis en évidence sur la carte. Et ils disposent à présent d'une technologie de pointe (alimentée par le Big Data), qui leur permet de renforcer leur position sur le terrain en conférant au processus un apanage de précision et de « science ».

Le résultat, c'est que nous criminalisons la pauvreté, tout en étant convaincus que nos outils sont non seulement scientifiques, mais également justes.

\*\*\*

Un week-end du printemps 2011, j'ai participé à un hackathon organisé dans la ville de New York. Ce genre d'événement vise à réunir des « hackers », des mordus d'informatique, des mathématiciens et des programmeurs passionnés, et à mobiliser leurs capacités intellectuelles pour faire la lumière sur les systèmes numériques qui exercent tant de pouvoir dans nos vies. J'étais associée à l'Union pour les libertés civiles de New York, et notre travail consistait à mettre au jour les données relatives à l'une des principales stratégies anticriminalité de la police new-yorkaise – la pratique dite du *stop, question and frisk* (« interpellation, interrogatoire et fouille »), connue de la plupart des gens sous la simple appellation de *stop and frisk*. Son usage s'était spectaculairement développé à l'ère du système CompStat, piloté par les données.

La police voyait le *stop and frisk* comme un dispositif de filtrage de la criminalité. Le principe est simple. Les agents interpellent les individus qui leur paraissent suspects – au vu de leur démarche, de leur accoutrement, ou en raison de leurs tatouages. Ils les questionnent et les jaugent, maintenus souvent bras et jambes écartés contre un mur ou sur le capot d'une voiture. Ils leur demandent leurs papiers puis les fouillent. Le raisonnement est que l'interpellation d'un nombre suffisant de personnes permet à coup sûr de prévenir quantité de petits délits, voire quelques graves forfaits. Cette politique, mise en place sous l'administration du maire Michael Bloomberg, bénéficiait d'un fort soutien de la part de la population. Au cours de la décennie précédente, le nombre d'interpellations avait augmenté de 600 %,

jusqu'à près de 700 000. L'immense majorité des interpellés étaient innocents. Ces contacts avec la police provoquaient souvent l'exaspération voire la fureur des individus arrêtés dans la rue. Une grande partie des gens établissaient néanmoins un lien entre ce programme et le net déclin de la criminalité dans la ville. Beaucoup avaient le sentiment que New York était plus sûre. Et les statistiques le confirmaient. Le nombre d'homicides, qui avait atteint 2 245 en 1990, n'était plus que de 515 (et baisserait en dessous de 400 en 2014).

Tout le monde savait que les individus interpellés par la police comptaient une proportion démesurée d'hommes jeunes à la peau mate. Mais combien étaient-ils ? Et dans quelle mesure ces interpellations avaient-elles abouti à une arrestation ou empêché la commission d'un délit ? Bien que théoriquement publiques, les informations étaient stockées pour l'essentiel dans une base de données difficile d'accès. Le logiciel ne tournait pas sur nos ordinateurs et ne s'exportait pas non plus sous forme de classeur Excel. Dans le cadre du hackathon, nous avons la tâche d'en faire sauter les verrous et de libérer les données, afin que chacun puisse analyser la nature et l'efficacité du programme *stop and frisk*.

Ce que nous avons découvert, sans grande surprise, c'est qu'une écrasante majorité des confrontations avec la police – environ 85 % – concernaient de jeunes hommes d'origine hispanique ou afro-américaine. Dans certains quartiers, beaucoup avaient été interpellés à plusieurs reprises. 0,1 % seulement, soit 1 sur 1 000, se trouvait mêlé d'une manière ou d'une autre à un crime violent. Ce filtre en attrapait cependant beaucoup d'autres pour des délits plus légers, de la possession de stupéfiants à la consommation d'alcool avant l'âge légal, délits qui autrement seraient peut-être passés inaperçus. Certains des individus contrôlés s'emportaient, comme on peut s'y attendre, et bon nombre d'entre eux se voyaient inculpés de refus d'obtempérer.

L'Union pour les libertés civiles intenta un procès contre l'administration Bloomberg, accusant la politique du *stop and frisk* de présenter un caractère raciste. C'était l'exemple même d'un maintien de l'ordre inéquitable, qui conduisait davantage les personnes issues des minorités devant la justice, puis en prison. Les hommes noirs, souligna-t-elle, avaient six fois plus de chances d'être incarcérés que les blancs et vingt et une fois plus d'être abattus par la police – du moins selon les données disponibles (lesquelles sont notoirement sous-évaluées).

Le *stop and frisk* n'est pas à proprement parler une ADM, car il dépend du jugement humain et n'est pas formalisé sous forme d'algorithme. Mais il repose sur un calcul basique et destructeur. Si la police interpelle un millier de personnes dans certains quartiers, elle débusquera en moyenne un suspect d'envergure et quantité de petits malfrats. Voilà qui ne diffère guère des calculs de probabilité utilisés par les publicitaires prédateurs ou les spammeurs. Même si le ratio de succès est minuscule, il suffit de se donner suffisamment de chances pour parvenir tout de même à atteindre sa cible. Et cela permet de mieux comprendre pourquoi le programme s'est à ce point développé sous le mandat de Bloomberg. Si l'interpellation de six fois plus d'individus débouchait sur six fois plus d'arrestations, les désagréments et le harcèlement subis par des milliers et des milliers d'innocents en valaient la peine. Ces derniers ne souhaitaient-ils pas, eux aussi, que la criminalité soit endiguée ?

Certains aspects du *stop and frisk* l'apparentaient bien toutefois à une ADM. Il comportait par exemple une boucle de rétroaction néfaste. Il prenait au piège des milliers de Noirs et d'Hispaniques, dont un grand nombre pour des infractions mineures qui par ailleurs se commettent impunément tous les samedis soirs, dans les fêtes universitaires. Mais alors qu'une grande majorité d'étudiants se remettaient librement de leurs excès après une nuit de sommeil, les victimes du *stop and frisk* étaient arrêtées et expédiées pour certaines dans l'enfer du centre pénitencier de Rikers Island. Chaque arrestation générait qui plus est de nouvelles données, justifiant encore davantage la politique menée.

Avec le développement du *stop and frisk*, le vénérable concept juridique de motif légitime fut pratiquement vidé de son sens, puisque les policiers pourchassaient non seulement des personnes qui pouvaient avoir déjà commis un délit, mais aussi celles qui étaient susceptibles d'en commettre un à l'avenir. Bien entendu, le but était parfois indubitablement atteint. En arrêtant un jeune homme qui, sous une bosse suspecte à la ceinture, se révélait dissimuler une arme à feu non déclarée, ils évitaient peut-être un meurtre ou un vol à main armée dans le quartier, si ce n'est même toute une série d'actes de cet acabit. Ou peut-être pas. Quoi qu'il en soit, le *stop and frisk* répondait à une certaine logique, que beaucoup trouvaient convaincante.

Mais cette politique respectait-elle la Constitution ? En août 2013, la juge fédérale Shira A. Scheindlin statua le contraire. Elle nota que les agents

interpellaient couramment « des Noirs et des Hispaniques qui ne l'auraient pas été s'ils avaient été blancs ». Le *stop and frisk*, écrivit-elle, bafouait le quatrième amendement, qui protège les citoyens contre les perquisitions et saisies non motivées, et ne leur apportait pas non plus la protection égalitaire garantie par le quatorzième amendement. Elle réclama que la pratique soit profondément réformée, avec notamment l'usage accru de caméras corporelles par les policiers en patrouille. Ces dispositifs aideraient à établir l'existence d'un motif légitime – ou son absence – et retireraient au modèle du *stop and frisk* une partie de son opacité. Mais ils ne résoudraient en rien la question de la répartition inégale des contrôles.

Lorsqu'on examine des ADM, on est souvent confronté au choix entre équité et efficacité. Nos traditions juridiques penchent nettement en faveur de l'équité. La Constitution américaine présume par exemple l'innocence des gens, et a été conçue pour garantir cette présomption. Du point de vue d'un modélisateur, la présomption d'innocence est une contrainte, dont il résulte que certains coupables restent libres, surtout quand ils ont les moyens d'engager de bons avocats. Ceux-là mêmes qui se voient condamnés disposent du droit de faire appel de leur verdict, ce qui consomme du temps et des ressources. Le système sacrifie donc d'énormes gains d'efficacité en échange d'une promesse d'équité. La Constitution considère implicitement que la libération, par manque de preuve, d'un individu qui pourrait fort bien avoir commis un délit fait courir moins de danger à notre société que d'emprisonner ou d'exécuter une personne innocente.

Les ADM tendent en revanche à favoriser l'efficacité. Elles se nourrissent, par nature, de données qui peuvent être mesurées et comptabilisées. Mais l'équité est difficile à appréhender et à quantifier. C'est un concept. Et les ordinateurs, malgré tous leurs progrès en matière de langage et de logique, ont encore énormément de mal à traiter les concepts. Ils ne « comprennent » la beauté que sous l'aspect d'un terme associé au Grand Canyon, aux couchers de soleil sur l'océan, et aux conseils cosmétiques du magazine *Vogue*. Ils essaient en vain de mesurer « l'amitié » au nombre de *likes* et de relations sur Facebook. Et le concept d'équité leur échappe totalement. Les programmeurs ne savent pas comment le coder, et leurs supérieurs leur demandent rarement de le faire.

L'équité n'est donc pas prise en compte dans les ADM, et la conséquence en est une production massive et industrielle d'*iniquité*. Si l'on assimile une

ADM à une usine, l'iniquité correspond à la suie noire que vomissent ses cheminées. C'est une émission toxique.

La question est de savoir si notre société serait prête à sacrifier un peu d'efficience dans l'intérêt de l'équité. Devrions-nous handicaper les modèles, en laissant de côté certaines informations ? Il se peut par exemple que l'ajout de plusieurs gigaoctets de données sur les comportements antisociaux permette d'aider PredPol à prédire les coordonnées géographiques des actes de grande délinquance. Mais c'est au prix d'une boucle de rétroaction néfaste. Je suis donc d'avis qu'il conviendrait d'y renoncer.

C'est un dossier difficile à plaider, similaire à bien des égards aux batailles sur les écoutes de la National Security Agency. Les partisans de cette surveillance font valoir qu'elle est essentielle pour la sûreté du pays. Et ceux qui dirigent le vaste appareil sécuritaire des États-Unis réclameront toujours davantage d'informations pour accomplir leur mission. Ils continueront de violer la vie privée des individus, tant qu'ils n'auront pas reçu le message qu'il leur faut trouver un moyen d'effectuer leur mission dans le respect des limites fixées par la Constitution. Ce sera peut-être plus compliqué, mais c'est indispensable.

L'autre question qui se pose est celle de l'égalité. La société américaine serait-elle aussi disposée à sacrifier le concept de motif légitime si tout le monde devait endurer le harcèlement et l'humiliation du *stop and frisk* ? La police de Chicago applique en la matière son propre programme. Au nom de l'équité, qu'advierait-il si elle envoyait un peloton de patrouilleurs dans le quartier huppé de Gold Coast ? Ils arrêteraient peut-être les joggeurs qui traversent W. North Boulevard en dehors des clous à la sortie du parc, ou verbaliseraient les déjections canines le long de Lakeshore Drive. Cette présence policière accrue permettrait probablement d'intercepter plus de conducteurs ivres et mettrait éventuellement au jour quelques cas de fraude à l'assurance, de violences conjugales, ou de racket. À l'occasion, manière simplement d'offrir à chacun un aperçu sans fard de l'expérience, les agents pourraient jeter des citoyens aisés contre le coffre de leur véhicule, leur tordre les bras et leur passer les menottes, en proférant peut-être aussi des injures et en les traitant de tous les noms.

Au fil du temps, l'accent mis sur le quartier de Gold Coast produirait des données. Elles y signaleraient une hausse de la criminalité, qui conduirait à accentuer encore la pression de la police. Il en résulterait sans doute une



irritation croissante et une multiplication des affrontements. J'imagine bien un conducteur garé en double file, répondant insolemment aux policiers, refusant de sortir de sa Mercedes, et se retrouvant inculqué de rébellion : un délit de plus à Gold Coast !

On pourrait se dire que ce n'est pas très grave. Mais l'égalité constitue un aspect crucial de la justice. Ce qui suppose, parmi maintes autres choses, d'être tous traités de manière égale par le système judiciaire. Les gens favorables à des politiques comme le *stop and frisk* devraient en faire eux-mêmes l'expérience. La justice ne saurait se réduire aux sanctions qu'une partie de la société inflige à l'autre.

Les effets nocifs d'un maintien de l'ordre inégalitaire, qu'il découle du *stop and frisk* ou de modèles prédictifs comme PredPol, ne prennent pas fin au moment où les accusés sont arrêtés et poursuivis. Une fois entre les mains du système judiciaire, beaucoup affrontent une autre ADM, étudiée au premier chapitre : le modèle du récidivisme, utilisé comme guide pour la détermination des peines. Les données discutables issues d'un travail de police biaisé alimentent directement ledit modèle. Les magistrats examinent ensuite cette analyse prétendument scientifique, cristallisée sous forme d'un unique score de risque. Ceux qui prennent ce score au sérieux ont alors toute raison de prononcer des sentences plus longues contre les détenus qui paraissent davantage enclins à commettre de nouveaux délits.

Pourquoi y a-t-il une plus grande probabilité pour que des prisonniers de couleur, originaires des quartiers pauvres, commettent des délits ? D'après les données fournies aux modèles de récidivisme, c'est parce qu'ils ont plus de chances d'être sans emploi, de ne pas avoir le bac, et d'avoir déjà eu affaire à la justice – au même titre que leurs amis.

Une autre façon d'aborder les mêmes données consiste néanmoins à considérer que ces prisonniers habitent des quartiers déshérités, où les établissements scolaires sont de piètre qualité et qui ne leur offrent pas de grandes perspectives d'avenir. En outre, ils font l'objet de contrôles de police très fréquents. Les chances pour qu'un ex-détenu de retour dans un tel quartier connaisse de nouveaux démêlés judiciaires sont donc sans nul doute plus élevées que celles d'un fraudeur fiscal relâché au cœur d'une banlieue cossue. Dans ce système-là, on punit plus lourdement les pauvres et les gens de couleur, à cause de ce qu'ils sont et de l'endroit où ils vivent.

Pour des systèmes prétendument scientifiques, les modèles du récidivisme laissent aussi à désirer sur le plan logique. L'hypothèse, admise

sans discuter, soutient qu'une plus longue incarcération des prisonniers « à haut risque » rend la société plus sûre. S'il est exact, bien entendu, que les prisonniers ne commettent aucun délit lorsqu'ils se trouvent derrière les barreaux, leur séjour en détention ne risque-t-il pas d'influencer leur comportement une fois libérés ? Les années passées dans un environnement brutal, peuplé de criminels, n'auraient-elles pas plus de chances – et non moins – de les porter à commettre un nouveau méfait ? Un tel constat saperait les fondements mêmes des recommandations pénales basées sur la modélisation du récidivisme. Mais les systèmes carcéraux, inondés de données, ne réalisent pas ce genre d'études, pourtant extrêmement importantes. Les données leur servent trop souvent à justifier leur mode de fonctionnement, et non à se remettre en cause ou à s'améliorer.

Comparez cette attitude à celle que l'on observe chez Amazon. Le géant de la distribution, à l'instar du système judiciaire, se focalise intensément sur une certaine forme de récidivisme. Mais il poursuit un but diamétralement opposé. Il veut que les gens reviennent acheter sur son site, encore et encore. Son logiciel cible le récidivisme et l'encourage.

Si Amazon fonctionnait à la manière du système judiciaire, il commencerait par attribuer un score aux acheteurs selon leur potentiel de récidive. Sont-ils plus nombreux à habiter certaines localités, ou à posséder un diplôme universitaire ? Amazon intensifierait dans ce cas ses efforts marketing en direction de ces gens-là, en leur offrant par exemple des remises ; si c'était efficace, les détenteurs d'un score de récidive élevé reviendraient acheter davantage d'articles. Les résultats obtenus auraient alors tout l'air, à première vue, de confirmer son système de notation.

Mais à la différence des ADM utilisées par la justice pénale, Amazon ne s'en tient pas à des corrélations aussi légères. L'entreprise exploite en matière de données un véritable laboratoire. Et si elle souhaite découvrir ce qui suscite la récidive d'achats, elle effectue des recherches. Ses experts ne se contentent pas de regarder les codes postaux ou les niveaux d'études. Ils examinent également l'expérience vécue au sein de l'écosystème d'Amazon. Ils commenceront peut-être par observer les profils de tous ceux qui ont acheté une ou deux fois sur le site et ne sont jamais revenus. Ont-ils rencontré des difficultés au moment de payer ? Leurs colis sont-ils arrivés en temps et en heure ? Sont-ils plus nombreux, en proportion, à avoir posté une critique négative ? Les questions se succèdent ainsi sans fin, car

l'avenir de l'entreprise dépend d'un système capable d'apprendre en permanence et de comprendre ce qui motive les clients.

Si j'avais l'occasion de travailler comme experte en données pour le système judiciaire, je ferais mon possible pour étudier en profondeur ce qui se passe dans les prisons, et l'impact que l'expérience carcérale peut avoir sur le comportement des détenus. Je me pencherais d'abord sur le principe du cachot. Des centaines de milliers de prisonniers sont enfermés vingt-trois heures par jour dans ces cellules séparées, pour la plupart pas plus grandes qu'un box d'écurie. Des chercheurs ont établi que l'isolement prolongé générerait un profond sentiment de découragement et de désespoir. Cela pourrait-il avoir un impact sur le récidivisme ? Voilà un test que j'adorerais mener – mais je ne suis même pas sûre que les données qui s'y rapportent soient collectées.

Qu'en est-il du viol ? Dans *Unfair : The New Science of Criminal Injustice*, Adam Benforado écrit que certains types de détenus sont la cible de viols dans les établissements pénitentiaires. Les jeunes et les individus de petite stature s'avèrent particulièrement vulnérables, de même que les déficients mentaux. Certains servent pendant des années d'esclaves sexuels. Il s'agit là d'un autre sujet important à analyser ; toute personne munie des données et de l'expertise adéquates pourrait y parvenir, mais le catalogage des effets à long terme de ces sévices n'a éveillé jusqu'à présent aucun intérêt de la part des responsables au sein des systèmes carcéraux.

Un scientifique digne de ce nom se mettrait aussi en quête de signaux positifs associés à l'expérience carcérale. Quelle est l'influence d'une plus grande exposition à la lumière du jour, d'une pratique sportive accrue, d'une meilleure alimentation, d'un programme d'alphabétisation ? Ces facteurs bonifieront peut-être le comportement des détenus après leur libération. Ils exerceront plus probablement un impact variable. Un projet de recherche sérieux destiné à évaluer le système judiciaire examinerait les effets de chacun de ces éléments, comment ils se combinent, et quels sont les individus qu'ils ont le plus de chances d'aider. Le but, si les données étaient utilisées de manière constructive, serait d'optimiser les centres de détention – de la même façon que des entreprises comme Amazon optimisent leur site web ou leur chaîne logistique – au profit à la fois des prisonniers et de la société dans son ensemble.

Mais les responsables administratifs des prisons sont fortement incités à éviter cette approche fondée sur l'analyse de données. Les risques en termes

de relations publiques sont trop importants – aucune ville ne souhaite faire l’objet d’un article cinglant dans les pages du *New York Times*. Et bien entendu, le surpeuplement carcéral rapporte en outre beaucoup d’argent. Les établissements gérés par le privé, qui n’accueillent pas plus de 10 % des personnes écrouées, représentent une industrie de 5 milliards de dollars. Pour engranger des profits, ils doivent afficher, à l’instar des compagnies aériennes, un taux de remplissage élevé. Un remue-ménage excessif risquerait de mettre en péril cette source de revenus.

Au lieu d’analyser les prisons et de les optimiser, on les traite donc comme des boîtes noires. Les détenus y entrent et disparaissent de notre vue. Des choses sordides s’y déroulent certainement, mais à l’abri de murs épais. Que se passe-t-il dedans ? Inutile de demander. Les modèles actuels s’en tiennent obstinément à l’hypothèse douteuse et pourtant incontestée selon laquelle une détention plus longue des prisonniers supposés à risque nous assurerait une plus grande sécurité. Si des études semblent bousculer cette logique, elles peuvent être aisément laissées de côté.

Et c’est précisément ce qui se produit. Prenons l’exemple d’une étude sur le récidivisme réalisée par Michael Mueller-Smith, professeur d’économie à l’université du Michigan. Ayant examiné 2,6 millions de dossiers criminels dans le comté de Harris, au Texas, il en a déduit que plus les détenus de ce comté restaient longtemps enfermés, moins ils avaient de chances de trouver d’emploi à leur sortie, et plus ils en avaient de recourir à l’aide alimentaire et à d’autres dispositifs d’assistance publique, et de commettre de nouveaux délits. Mais pour transposer en une politique intelligente ce genre de conclusions, et rendre ainsi mieux justice, les responsables publics devraient prendre position au nom d’une minorité redoutée, que beaucoup d’électeurs (si ce n’est la plupart) préféreraient de très loin ignorer. Un engagement peu vendeur à l’heure actuelle.

Le *stop and frisk* paraît intrusif et injuste, mais dans peu de temps, il semblera également primitif. Car la police reprend en effet aujourd’hui des outils et des techniques provenant de la campagne mondiale menée contre le terrorisme, et les applique à la lutte contre la criminalité locale. À San Diego, par exemple, les agents ne se contentent pas de demander l’identité des personnes interpellées, ou de les fouiller. Il leur arrive également de les photographier avec un iPad puis d’envoyer ces clichés dans le *cloud* vers un service de reconnaissance faciale, qui les compare à une base de données de délinquants et de suspects. Selon un article du *New York Times*, la police de

San Diego a employé ce logiciel de reconnaissance faciale sur 20 600 individus entre 2011 et 2015. Beaucoup ont aussi été soumis à un frottis buccal afin de prélever leur ADN.

Les progrès technologiques en matière de reconnaissance faciale autoriseront bientôt un contrôle bien plus large. La municipalité de Boston envisageait par exemple l'utilisation de caméras pour scanner le visage de milliers de spectateurs lors des concerts en plein air. Les données auraient été transmises à un service capable en une seconde de comparer chacun de ces visages à un million d'autres. Au bout du compte, la ville avait renoncé à son projet. Les préoccupations liées au respect de la vie privée avaient ici triomphé de la soif d'efficience. Mais ce ne sera pas toujours le cas.

À mesure que la technologie avance, nous allons assurément assister à une intensification spectaculaire de la surveillance. La bonne nouvelle, si l'on peut dire, c'est qu'à partir du moment où des milliers de caméras implantées dans nos villes analyseront notre portrait, la police n'aura plus à discriminer autant. Et la technologie se montrera sans nul doute utile pour localiser des suspects, comme lors de l'attentat à la bombe du marathon de Boston. Mais cela signifie aussi que nous ferons tous l'objet d'une forme numérique de *stop and frisk*, nos visages comparés à ceux de criminels et de terroristes connus enregistrés dans des bases de données.

L'attention pourrait alors fort bien se tourner vers le repérage de contrevenants *potentiels* – non pas simplement l'identification de quartiers ou de zones sur une carte, mais celle d'individus précis. Ces campagnes préventives, d'un usage déjà courant dans la lutte contre le terrorisme, constituent un terreau fertile pour les ADM.

En 2009, la police de Chicago a obtenu une subvention de 2 millions de dollars de la part du National Institute of Justice afin de développer un programme de prédiction criminelle. La théorie qui avait mené à l'octroi de cette somme reposait sur l'idée qu'il serait possible, sous réserve de recherches et de données suffisantes, de démontrer que la propagation de la criminalité suivait certains schémas particuliers, à l'image d'une épidémie. On pouvait la prédire, et avec un peu de chance la prévenir.

Le responsable scientifique de cette initiative était Miles Wernick, directeur du Centre de recherche en imagerie médicale de l'Illinois Institute of Technology, ou IIT. Quelques décennies plus tôt, Wernick avait aidé l'armée américaine à sélectionner ses cibles sur le champ de bataille grâce à l'analyse de données. Il était passé depuis à l'analyse d'informations

médicales, portant notamment sur la progression de la démence. Mais comme la plupart des *data scientists*, il ne considérait pas que son expertise fût attachée à un secteur d'activité spécifique. Il repérait des schémas. Et à Chicago, il se focaliserait sur ceux qui touchaient à la criminalité et aux délinquants.

L'équipe de Wernick concentra ses premiers efforts sur la détection de « points chauds », à la manière de PredPol. Mais le groupe de Chicago alla bien plus loin. Il élaborait une liste des quelque 400 individus qui avaient le plus de risques de commettre un délit violent. Et il classa ces individus selon leur probabilité d'être impliqué dans un homicide.

À l'été 2013, l'une des personnes de la liste – Robert McDaniel, 22 ans, sorti sans diplôme du lycée –, répondit un beau jour à la porte et trouva face à lui un agent de police. Il indiquerait plus tard au *Chicago Tribune* qu'il n'avait jamais enfreint la législation sur le port d'armes et n'avait jamais été inculpé de crime avec violence. Comme la plupart des jeunes d'Austin, quartier mal famé du West Side, McDaniel avait eu des démêlés avec la justice et connaissait quantité de gens aux prises avec le système pénal. La policière lui fit savoir que les forces de l'ordre l'avaient à l'œil et qu'il ferait mieux de se tenir tranquille.

L'analyse qui avait conduit la police jusqu'à McDaniel s'appuyait en partie sur son réseau social. Il connaissait des délinquants. Or, on ne peut nier que les gens sont statistiquement plus enclins à se comporter comme ceux qu'ils fréquentent. Facebook a établi par exemple que les amis qui communiquent souvent ont bien plus de chances de cliquer sur la même publicité. Statistiquement parlant, le proverbe « qui se ressemble s'assemble » se vérifie bel et bien.

Et pour être tout à fait juste envers la police de Chicago, précisons qu'elle n'arrête pas les personnes telles que Robert McDaniel – du moins pas encore. Son objectif avec cette opération consiste à sauver des vies. Si les 400 individus qui semblent les plus susceptibles de commettre un délit violent reçoivent une visite et un avertissement, certains réfléchiront peut-être à deux fois avant de se munir d'une arme à feu.

Mais considérons le cas de McDaniel en termes d'équité. Il se trouve qu'il a grandi dans un quartier pauvre et dangereux. Sur ce plan-là, il n'a pas eu de chance. Il a été entouré par la délinquance et nombre de ses relations sont tombées dedans. Et du fait en grande partie de ces circonstances – et non de ses propres actes –, il a été lui-même jugé

dangereux. Désormais, la police le tient à l'œil. S'il commet une bêtise, comme le font régulièrement des millions d'autres Américains, s'il achète de la drogue, se bagarre dans un bar ou porte sur lui une arme de poing non déclarée, la justice ne lui fera aucun cadeau, et se montrera sans doute beaucoup plus sévère qu'avec la plupart d'entre nous. Après tout, il a été averti.

Le modèle qui a mené la police jusqu'à la porte de Robert McDaniel ne poursuit pas à mon sens le bon objectif. Au lieu d'essayer simplement d'éradiquer la criminalité, les forces de l'ordre devraient s'efforcer de nouer des liens au sein des quartiers. C'était l'un des axes fondamentaux de l'étude originelle dite des « fenêtres cassées » : les policiers se déplaçaient à pied, échangeaient avec les gens et tâchaient de les aider à maintenir les normes en vigueur dans leur collectivité. Mais cet objectif s'est souvent perdu, laminé par des modèles qui assimilent arrestations et sécurité.

Ce n'est pas le cas partout. Je me suis récemment rendue à Camden, dans le New Jersey, qui était en 2011 la capitale du meurtre aux États-Unis. J'y ai découvert que sa police, rénovée et placée sous contrôle de l'État en 2012, avait un double mandat : réduire la criminalité et susciter la confiance parmi la population. Si l'objectif fixé consiste à créer de la confiance, il est fort possible alors que l'arrestation devienne le dernier recours, plutôt que le premier. Cette approche plus empathique pourrait réchauffer les relations entre les forces de l'ordre et les citoyens, et limiter les tragédies que nous avons connues ces dernières années – la mort de jeunes hommes noirs tués par la police, et les émeutes qui se sont ensuivies.

D'un point de vue mathématique, néanmoins, la confiance est difficile à quantifier. C'est un vrai défi pour les modélisateurs. Il s'avère hélas bien plus simple de continuer à dénombrer les arrestations et de bâtir des modèles fondés sur l'hypothèse du « qui ressemble s'assemble », et qui nous traitent en conséquence. Les innocents entourés de criminels se font malmenés, tandis que les criminels entourés de bons citoyens respectueux des lois passent au travers. Et compte tenu de la forte corrélation entre les situations de pauvreté et le signalement de délits, ce sont les plus modestes qui continuent d'être pris au piège de ces filets numériques. Le reste de la population n'a guère besoin de s'en préoccuper.

---

1. Des logiciels prédictifs commencent également à être testés dans plusieurs préfectures en France, notamment dans l'Oise.

2. Système frauduleux consistant à rémunérer les souscripteurs d'un fonds d'investissement grâce à l'argent des nouveaux contributeurs.



## Trouver un emploi **INAPTE AU SERVICE**

1001  books

Il y a quelques années, un jeune homme du nom de Kyle Behm interrompit ses études à l'université Vanderbilt. Il souffrait de troubles bipolaires et avait besoin de temps pour se faire soigner. Un an et demi plus tard, Kyle se trouva suffisamment remis pour reprendre son cursus dans un autre établissement. À la même époque, un ami lui parla d'un travail à mi-temps au supermarché Kroger. Ce n'était qu'un emploi à salaire minimum dans la grande distribution, mais il semblait assuré de l'obtenir. Son ami, qui s'apprêtait à quitter le poste, pouvait le recommander à sa hiérarchie. Pour un brillant étudiant tel que Kyle, ce recrutement avait tout l'air d'une formalité.

Mais il ne fut jamais rappelé pour un entretien. Quand il voulut se renseigner, son ami lui expliqua que le test de personnalité qu'il avait rempli dans le cadre de sa candidature s'était révélé négatif. Ce test faisait partie d'un programme de sélection élaboré par Kronos, une société spécialisée dans la gestion d'effectifs située près de Boston. Après qu'il lui eut rapporté ce qui s'était passé, son père Roland, avocat, demanda à Kyle quel type de questions on lui avait posé. Il indiqua qu'elles ressemblaient beaucoup au test du « modèle des cinq facteurs » qu'il avait effectué à l'hôpital. Ce modèle évalue les individus selon cinq caractéristiques de leur personnalité ; la propension à être extraverti, agréable, consciencieux, névrotique, ou ouvert d'esprit.

Au départ, être privé d'un emploi à faible rémunération à cause d'un test certes discutable ne leur parut pas si grave. Roland Behm poussa son fils à postuler ailleurs. Mais Kyle ramenait à chaque fois la même mauvaise nouvelle. Les entreprises auprès desquelles il candidatait utilisaient toutes le même test, et il ne décrochait aucune offre. Comme Roland se le remémorerait plus tard : « Kyle m'a dit : "J'ai eu un score quasi parfait au

SAT et j'étais à Vanderbilt<sup>1</sup> il y a encore quelques années. Si je n'arrive même pas à obtenir un travail à temps partiel au salaire minimum, suis-je encore bon à quelque chose ?" J'ai répondu que j'en étais convaincu. »

Mais Roland Behm était abasourdi. Des questions relatives à sa santé mentale semblaient éjecter son fils du marché de l'emploi. Il décida de creuser le sujet et découvrit vite que l'utilisation de tests de personnalité en matière de recrutement était de fait très répandue parmi les grandes entreprises. Il constata pourtant très peu de recours juridiques contre cette pratique. Les gens qui postulent pour un poste et voient leur candidature rejetée apprennent rarement, m'expliqua-t-il, qu'ils ont été écartés en raison des résultats d'un test. Et même quand c'est le cas, il est peu probable qu'ils contactent un avocat.

Behm notifia par la suite sept entreprises – Finish Line, Home Depot, Kroger, Lowe's, PetSmart, Walgreen Co. et Yum Brands –, et les informa de son intention de déposer une action de groupe, au motif que l'usage de ces tests lors du processus de candidature présentait un caractère illégal.

À l'heure où j'écris ces lignes, le procès est toujours en attente<sup>2</sup>. Les débats se concentreront probablement sur le fait de savoir si le test de Kronos peut être considéré comme un examen médical, dont l'utilisation à des fins de recrutement se trouve frappée d'illégalité en vertu d'une loi contre les discriminations basées sur le handicap – le Americans with Disabilities Act – promulguée en 1990. S'il s'avère que c'est bien le cas, le tribunal devra déterminer si la responsabilité d'avoir enfreint l'ADA revient aux entreprises recruteuses elles-mêmes, ou s'il faut en blâmer Kronos.

Ce qui nous intéresse plus particulièrement ici, c'est la manière dont des systèmes automatiques nous jugent lorsque nous cherchons un emploi, et sur les critères qu'ils évaluent. Nous avons déjà vu que des ADM polluaient le processus d'admission à l'université, aussi bien pour les riches que pour les classes moyennes. Dans le même temps, les ADM mises en œuvre par la justice pénale piègent des millions de gens, dont une grande majorité d'individus pauvres qui pour la plupart n'ont jamais eu la chance de suivre des études. Les membres de chacun de ces groupes affrontent des défis radicalement différents. Mais ils ont aussi quelque chose en commun. Au bout du compte, ils ont tous besoin d'un emploi.

Autrefois, trouver du travail était essentiellement affaire de relations. En postulant chez Kroger, Kyle Behm avait en réalité suivi la voie traditionnelle. Son ami l'avait averti qu'un poste se libérait, et avait parlé de

lui en bien. Depuis des décennies, c'est de cette façon que les gens parvenaient à « glisser un pied dans la porte », que ce soit celle d'un commerce alimentaire, d'un port, d'une banque ou d'un cabinet d'avocats. S'ensuivait habituellement un entretien, lors duquel un manager cherchait à se faire une opinion du candidat – opinion qui se traduisait trop souvent par une simple appréciation sommaire : cette personne me ressemble-t-elle (ou est-ce qu'elle ressemble à d'autres avec qui je m'entends bien) ? Pour les demandeurs d'emploi qui n'avaient pas d'ami dans l'entreprise, les possibilités étaient par conséquent réduites, surtout s'ils étaient d'une origine ou d'une religion différente. Les femmes se retrouvaient elles aussi exclues de ce jeu d'initiés.

C'est en partie pour rendre le processus plus juste que des firmes comme Kronos ont apporté une dimension scientifique à la gestion des ressources humaines. Fondée dans les années 1970 par des diplômés du Massachusetts Institute of Technology (MIT), elle élabore une pointeuse d'un nouveau genre, équipée d'un microprocesseur, qui additionnait les heures de travail des employés et les déclarait automatiquement. Cela peut sembler banal, mais il s'agissait là des prémices d'une offensive électronique (qui se propage aujourd'hui à une vitesse fulgurante) visant à pister et à optimiser la main-d'œuvre.

Kronos développa au fil de son évolution une large gamme d'outils logiciels pour la gestion d'effectifs, avec entre autres un programme baptisé Workforce Ready HR qui promettait d'après sa page web d'éliminer les « approximations » en matière de recrutement : « Nous pouvons vous aider à sélectionner, à engager et à intégrer les candidats susceptibles d'être les plus productifs – les employés les mieux adaptés, qui se montreront plus performants et resteront plus longtemps en poste. »

Kronos fait partie d'une industrie en plein essor. Le recrutement s'automatise, et les nouveaux programmes mis au point incluent souvent des tests de personnalité semblables à celui qu'avait passé Kyle Behm. C'est une activité qui représente désormais, selon Hogan Assessment Systems Inc., société spécialisée dans ce type de tests, un chiffre d'affaires annuel de 500 millions de dollars, et qui affiche une croissance de 10 à 15 % par an. Josh Bertin, du cabinet de conseil Deloitte, estime que des tests de personnalité sont aujourd'hui effectués sur 60 à 70 % des travailleurs potentiels aux États-Unis, soit 30 à 40 % de plus qu'il y a cinq ans<sup>3</sup>.

Naturellement, ces programmes de recrutement ne peuvent collecter aucune information sur la manière dont le candidat se comporterait réellement au sein de l'entreprise. Impossible en effet de prédire l'avenir. À l'instar de tant d'autres programmes basés sur le Big Data, ils se contentent donc de données de rechange. Or, comme nous l'avons vu, ces données supplétives s'avèrent immanquablement inexactes et souvent injustes. En 1971, dans l'affaire *Griggs v. Duke Power Company*, la Cour suprême des États-Unis a statué du reste que l'utilisation de tests d'intelligence dans le cadre d'un processus d'embauche était discriminatoire et par suite illégale. On pourrait croire que ce procès aurait eu pour effet de provoquer un examen de conscience. Mais l'industrie a simplement décidé de leur en substituer d'autres, et notamment des tests de personnalité comme celui qui a rejeté la candidature de Kyle Behm.

Même en laissant de côté les questions d'équité et de légalité, des études semblent indiquer que les tests de personnalité sont de piètres prédicteurs des performances professionnelles. Frank Schmidt, professeur de sciences commerciales à l'université de l'Iowa, a analysé un siècle de données sur la productivité en entreprise, afin de mesurer la valeur prédictive de divers procédés de sélection. Les tests de personnalité se sont classés en bas de l'échelle – leur capacité de prédiction n'excédait pas le tiers de celle des examens cognitifs, et ils se situaient également très en deçà d'une simple vérification des références. C'est un constat particulièrement désolant, car les recherches montrent que certains de ces tests peuvent réellement aider les employés à mieux se connaître. Ils peuvent aussi servir à renforcer la cohésion d'une équipe et à améliorer la communication. Ils amènent en effet les gens à s'interroger ouvertement sur leur manière de collaborer. Cet état d'esprit serait en mesure à lui seul d'aboutir à l'instauration d'un meilleur environnement de travail. En d'autres termes, si le but fixé est de rendre le travailleur plus heureux, les tests de personnalité pourraient finalement constituer un outil précieux.

Mais à l'heure actuelle, ils font au contraire office de filtre pour éliminer les candidats. « L'objectif premier du test, expliqua Roland Behm, n'est pas de dénicher le meilleur employé, mais d'exclure le plus de monde possible de la façon la plus économique. »

On pourrait penser que les tests de personnalité sont faciles à manipuler. Si vous réalisez sur Internet le test des cinq facteurs, cela ressemble à un jeu d'enfant. L'une des questions posées demande si vous êtes sujet à de

fréquentes sautes d'humeur. Il sera probablement judicieux de répondre que ce n'est pas du tout le cas. Une autre question demande si vous vous emportez facilement. Là encore, on cochera « non » : peu d'entreprises souhaitent recruter des employés irascibles.

Trier les postulants sur la base de telles questions peut en réalité leur attirer des ennuis. Les autorités de régulation de Rhode Island ont ainsi établi que CVS Pharmacy écartait illégalement les candidats souffrant de maladie mentale ; en effet, leur test de personnalité demandait aux sujets de répondre oui ou non à des énoncés tels que : « Les gens font beaucoup de choses qui vous mettent en colère » ou « Il ne sert à rien d'avoir des amis proches ; ils vous laissent toujours tomber ». Des questions plus complexes, plus dures aussi à contourner, auront davantage de chances d'épargner aux entreprises ce genre de déboires. C'est pourquoi nombre des tests utilisés aujourd'hui imposent aux candidats des choix difficiles à trancher, qui leur laissent probablement l'angoissante impression qu'aucune réponse n'est la bonne.

Chez McDonald's, par exemple, les recrues potentielles devaient choisir, parmi les deux énoncés suivants, celui qui les décrivait le mieux : « Il est difficile d'être enthousiaste quand on a beaucoup de problèmes à résoudre » ou « J'ai parfois besoin qu'on me pousse un peu pour me mettre au travail ».

Le *Wall Street Journal* a demandé à Tomas Chamoro-Premuzic, psychologue du travail, d'analyser ce type de questions embarrassantes. Le premier énoncé rendait compte, indiqua-t-il, des « différences individuelles sur le plan du névrosisme et de la conscienciosité » ; le second signalait un « manque d'ambition et de motivation ». L'éventuel employé s'avouait donc coupable soit d'avoir une sensibilité à fleur de peau, soit d'être paresseux.

Chez Kroger, l'une des questions posées était bien plus simple : Quel adjectif vous décrit le mieux dans le cadre professionnel – « unique » ou « ordonné » ?

« Unique », expliqua Chamoro-Premuzic, révélait une « conception de soi, une ouverture et un narcissisme élevés », tandis qu'« ordonné » exprimait la conscienciosité et la maîtrise.

Notons qu'il n'est pas permis de répondre « les deux ». Les candidats à l'embauche doivent sélectionner l'un ou l'autre des choix proposés, sans la moindre idée de la manière dont le programme l'interprétera. Et certaines

analyses en tireront des conclusions peu flatteuses. Partout dans le pays, si vous vous rendez dans une école maternelle, vous entendrez souvent les maîtresses dire aux enfants qu'ils sont « uniques ». C'est exact, bien sûr, mais également une façon de renforcer leur estime de soi. Une douzaine d'années plus tard, cependant, quand cet élève postulera à un emploi au salaire minimum et sélectionnera « unique » sur son test de personnalité, le logiciel y verra peut-être bien un motif de refus : qui voudrait d'une main-d'œuvre peuplée de narcissiques ?

Les défenseurs des tests soulignent qu'ils comportent un grand nombre de questions, et qu'aucune réponse ne saurait à elle seule exclure un candidat. Certaines tendances générales dans les réponses fournies peuvent en revanche bel et bien le disqualifier. Mais on ignore quelles sont ces tendances. On ne nous dit pas ce que les tests recherchent. Le processus est totalement opaque.

Pire encore, une fois que des experts techniques ont calibré le modèle, celui-ci ne reçoit que très peu de remontées d'informations. L'univers sportif offre à nouveau ici un excellent contraste. La plupart des équipes professionnelles de basket emploient des passionnés de données, qui mettent en œuvre des modèles permettant d'analyser les joueurs en fonction d'une série d'indicateurs, parmi lesquels la vitesse de déplacement, la détente verticale, le pourcentage de lancers francs réussis et quantité d'autres variables. Lors des recrutements d'intersaison, les Lakers de Los Angeles dédaigneront peut-être un sensationnel meneur de l'université de Duke pour cause de mauvaises statistiques sur les passes décisives. Car les meneurs se doivent en effet d'être de bons passeurs. La saison suivante, à leur grand désarroi, le joueur rejeté remportera pourtant le titre de Rookie of the Year<sup>4</sup> pour les Jazz de l'Utah, et se classe en tête de la ligue sur le nombre de passes décisives. Les Lakers pourront dans un pareil cas revenir sur leur modèle, afin de découvrir sur quoi ils se sont trompés. L'ancienne équipe du joueur comptait peut-être sur lui pour marquer, ce qui aura mis à mal son nombre de passes. À moins qu'il ait appris chez les Jazz une meilleure technique de passe. Quoi qu'il en soit, les Lakers auront alors la possibilité de travailler à l'amélioration du modèle utilisé.

Imaginez maintenant que Kyle Behm, ayant été refusé chez Kroger, décroche ensuite un poste chez McDonald's. Il se révèle un employé remarquable. Il prend la direction de la cuisine au bout de quatre mois, puis celle du restaurant tout entier un an plus tard. Verra-t-on quiconque chez

Kroger réexaminer le test de personnalité, et chercher à comprendre comment ils ont pu à ce point se fourvoyer ?

Aucune chance, à mon avis. La différence tient au fait que les équipes de basket gèrent des individus, dont chacun vaut potentiellement plusieurs millions de dollars. Leurs moteurs d'analyse revêtent une importance cruciale en termes d'avantage concurrentiel, et ont un énorme besoin de données. Sans un retour d'informations permanent, leurs systèmes deviennent obsolètes et déficients. Les entreprises qui embauchent au salaire minimum gèrent à l'inverse de grandes masses de gens. Elles sabrent dans leurs dépenses en remplaçant les professionnels des ressources humaines par des machines, et ces machines filtrent de vastes populations en différents groupes plus maniables. Sauf si quelque chose tourne mal parmi son personnel – une épidémie de cleptomanie, par exemple, ou une baisse brutale de la productivité –, l'entreprise a peu de raisons de modifier son modèle de filtrage. Ce dernier remplit sa tâche – même s'il passe à côté de quelques talents d'exception.

L'entreprise peut se satisfaire du statu quo, mais les victimes de ses systèmes automatisés en pâtissent. Et comme vous pouvez vous en douter, je considère les tests de personnalité des services recruteurs comme des ADM. Ils cochent l'ensemble des cases. Ils sont en premier lieu très répandus et exercent un impact considérable. Le test de Kronos, avec tous ses défauts, a pratiquement gagné toute l'économie du recrutement. Avant cela, les employeurs avaient indubitablement des partis pris. Mais ces partis pris différaient d'une entreprise à l'autre, ce qui pouvait ici ou là ouvrir des portes à quelqu'un comme Kyle Behm. C'est de moins en moins vrai. Et d'une certaine manière, Kyle avait de la chance. Les demandeurs d'emploi, surtout pour un travail au salaire minimum, essuient sans cesse des refus et découvrent rarement pourquoi. C'est par pur hasard que l'ami de Kyle a eu vent du motif de son rejet, et qu'il a pu lui en parler. Et quand bien même, le procès contre les grands utilisateurs de Kronos ne mènerait probablement nulle part si le père de Kyle n'était pas avocat, avec assez de temps et d'argent pour monter un large recours juridique. Les candidats à des emplois peu qualifiés se trouvent rarement dans ce cas<sup>5</sup>.

Considérez pour finir la boucle de rétroaction qu'engendre le test de personnalité de Kronos. Le refus opposé aux gens souffrant de certains problèmes de santé mentale les empêche d'avoir un emploi et de mener une

vie ordinaire, ce qui les isole d'autant plus. C'est exactement ce que le Americans with Disabilities Act est supposé éviter.

\*\*\*

La majorité des postulants à l'embauche ne sont heureusement pas éjectés par des systèmes automatiques. Mais ils restent confrontés au défi de faire grimper leur candidature au sommet de la pile et d'obtenir un entretien. Ce qui a été longtemps un problème pour les minorités ethniques, ainsi que pour les femmes.

En 2001 et 2002, avant le développement des lecteurs automatiques de curriculum vitæ, des chercheurs de l'université de Chicago et du MIT ont envoyé 5 000 CV factices en réponse à des offres d'emploi publiées dans deux quotidiens, le *Boston Globe* et le *Chicago Tribune*. Les postes couvraient tout un éventail allant des tâches administratives au service client et à la vente. Chaque CV était conçu selon l'angle des origines ethniques. Une moitié d'entre eux affichaient des noms typiquement blancs comme Emily Walsh ou Brendan Baker, tandis que les autres, similaires en termes de qualifications, portaient des noms tels que Lakisha Washington ou Jamaal Jones, aux sonorités afro-américaines. Les chercheurs constatèrent que les entreprises rappelaient les noms blancs une fois et demi plus souvent que les noms noirs. Mais ils firent aussi une autre observation, peut-être encore plus saisissante. Parmi les candidats blancs, les excellents CV suscitaient beaucoup plus d'intérêt que les CV plus faibles, mais parmi les candidats noirs, l'excellence du CV d'une candidature à l'autre ne faisait guère de différence. Les préjugés gangrénaient manifestement toujours le marché du recrutement.

Le moyen idéal de déjouer ce genre de préjugés consiste à évaluer les candidats à l'aveugle. Les orchestres, longtemps dominés par les hommes, sont connus pour avoir commencé dès les années 1970 à auditionner leurs musiciens dissimulés derrière une tenture. Les relations et la réputation ne comptaient soudain plus du tout, ni les origines ethniques ou l'université que le musicien avait fréquentée. La musique jouée derrière le rideau se suffisait à elle-même. Depuis, le pourcentage de femmes dans les grands orchestres a été multiplié par cinq – même si elles ne représentent encore qu'un quart des effectifs.



Le problème est qu'il existe peu de professions en mesure d'élaborer une épreuve de recrutement aussi impartiale. Les musiciens cachés derrière un rideau peuvent réellement exécuter le travail auquel ils postulent, qu'il soit question d'un concerto pour violoncelle de Dvořák ou d'une bossa-nova à la guitare. Dans d'autres métiers, les employeurs doivent fouiller les curriculum vitæ à la recherche de qualités laissant présager d'une embauche réussie.

Pour vider les piles de CV qu'ils reçoivent, les départements des ressources humaines s'appuient, comme on peut s'en douter, sur des systèmes automatiques. Environ 72 % d'entre eux ne passent en fait jamais sous les yeux d'un être humain. Des programmes informatiques les parcourent et en extraient les compétences et les expériences recherchées par l'employeur. Ils attribuent ensuite un score à chaque CV, en fonction de son adéquation à l'emploi proposé. C'est aux agents de la DRH qu'il revient de fixer le seuil, mais plus ils parviennent à éliminer de candidats à l'aide de ce premier tri, moins ils devront consacrer d'heures à examiner les meilleurs CV ainsi présélectionnés.

Les candidats doivent donc rédiger leur CV en ayant en tête ce dispositif de lecture automatique. Il importe par exemple d'y parsemer généreusement les termes recherchés dans le cadre de l'offre spécifique à laquelle on répond. Il pourra notamment s'agir d'intitulés de postes (responsable des ventes, directeur financier, architecte logiciel), de langues étrangères ou de langages informatiques (le mandarin, le Java), ou de distinctions honorifiques (mentions universitaires, décorations scoutistes).

Les personnes les mieux renseignées apprennent ce qui plaît aux machines et ce qui les perturbe. Les images ne sont ainsi d'aucune utilité. La plupart des « scanners » de CV ne les traitent pas encore. Et les polices de caractère fantaisie n'ont d'autre effet que de troubler les machines, explique Mona Abdel-Halim, cofondatrice du site d'aide à la candidature Resunate.com. Les polices ordinaires comme Arial ou Courier restent le choix le plus sûr. Oubliez également les symboles, tels que les flèches. Ils ne font qu'embrouiller les choses, et empêchent les systèmes automatiques d'analyser correctement les informations.

Conséquence de ces programmes : de la même manière que vis-à-vis de l'admission à l'université, les individus qui possèdent l'argent et les ressources nécessaires pour bien préparer leur CV sortent gagnants. Les autres ne sauront peut-être jamais qu'ils envoient le leur dans un trou noir.

C'est un exemple de plus où les personnes riches et bien informées bénéficient là encore d'un avantage, tandis que les individus pauvres sont plus volontiers lésés.

Pour être tout à fait honnête, la production d'un CV a toujours été soumise à divers partis pris. Au sein des générations précédentes, ceux qui étaient au courant prenaient soin d'en organiser les rubriques de façon claire et cohérente, de le taper sur une machine à écrire de qualité, comme une IBM Selectric, et de l'imprimer sur un papier haut de gamme. Ce genre de CV avait plus de chances de passer l'obstacle des filtres humains. Les CV manuscrits, ou qui présentaient des bavures provoquées par une polycopieuse, finissaient plus souvent à la corbeille. Les inégalités d'accès aux perspectives de carrière n'ont en ce sens rien de nouveau. Elles ont simplement resurgi sous une nouvelle incarnation, pour guider cette fois les « gagnants » de la société au travers de filtres électroniques.

Le traitement inégal réservé par ces filtres s'étend bien au-delà de la question des CV. Nos moyens de subsistance dépendent de plus en plus de la capacité à défendre notre cause devant des machines. Google en offre l'exemple le plus évident. Pour une entreprise, qu'il s'agisse d'un gîte ou d'un garage automobile, la condition du succès tient à sa présence en première page des résultats de recherche. Mais les individus affrontent désormais des défis similaires, que ce soit pour glisser un pied dans la porte d'un employeur, gravir les échelons – ou même survivre aux vagues de licenciements. La clé consiste à apprendre ce que les machines recherchent. Or, dans un univers numérique vendu comme juste, scientifique et démocratique, les initiés trouvent ici aussi le moyen d'acquérir un avantage crucial.

\*\*\*

Dans les années 1970, le bureau des admissions de l'école de médecine de l'hôpital Saint-George, situé au sud de Londres, dans le quartier de Tooting, s'aperçut qu'il y avait une opportunité à saisir. Pour chacune des 150 places proposées tous les ans, l'école recevait plus de douze candidatures. Passer au peigne fin l'ensemble de ces dossiers demandait beaucoup de travail et mobilisait plusieurs agents. Comme ils avaient chacun des idées et des préférences différentes, le processus s'avérait en outre quelque peu aléatoire. Ne serait-il pas possible de programmer un

ordinateur, afin de trier les candidatures et d'en ramener le nombre à une quantité plus facile à gérer ?

De grandes organisations comme le Pentagone ou IBM utilisaient déjà des ordinateurs pour ce genre de tâches. Mais qu'une école de médecine mette au point dès la fin des années 1970 son propre programme d'évaluation automatisée, au moment où Apple sortait tout juste son premier micro-ordinateur personnel, voilà qui représentait une expérience audacieuse.

Celle-ci se solda néanmoins par un fiasco. Saint-George avait non seulement fait, semblait-il, un usage prématuré de la modélisation mathématique, mais il s'était aussi révélé un pionnier involontaire en matière d'ADM.

Le problème, comme avec tant d'autres ADM, était apparu dès l'origine, quand les administrateurs avaient énoncé les deux objectifs conjoints du modèle. Le premier visait à accroître l'efficacité, en laissant la machine effectuer l'essentiel du défrichage. Elle réduirait automatiquement les 2 000 candidatures à 500, après quoi des opérateurs humains prendraient le relais pour un long processus d'entretien. Le second objectif était celui de l'équité. L'ordinateur ne serait influencé ni par les humeurs ou les préjugés des administrateurs, ni par les urgentes sollicitations de quelque lord ou de quelque ministre. Lors de ce premier filtrage automatique, chaque candidat se trouverait jugé selon les mêmes critères.

Et quels critères allait-on utiliser ? Cet aspect-là semblait constituer le volet le plus facile. Saint-George disposait déjà de volumineuses archives résultant des sélections menées les années précédentes. Il s'agissait donc d'enseigner au système informatique à reproduire de façon identique les procédures suivies par les humains. Comme vous l'aurez sans doute deviné, ce sont précisément ces sources de données qui posèrent souci. L'ordinateur apprit des êtres humains comment discriminer les candidats, et s'acquitta de ce travail avec une stupéfiante efficacité.

Pour rendre justice aux administrateurs de Saint-George, les facteurs discriminatifs des données d'entraînement n'étaient pas tous ouvertement racistes. Bon nombre des candidatures portant un nom étranger, ou provenant d'un autre pays, émanaient d'individus qui ne maîtrisaient manifestement pas l'anglais. Au lieu d'envisager la possibilité que de grands médecins puissent apprendre la langue, ce qui est aujourd'hui une évidence, la tendance était simplement de les écarter (l'école devait après

tout éliminer les trois quarts des postulants, et il semblait assez simple de commencer par là).

Cependant, alors que le personnel humain de Saint-George rejetait depuis longtemps les candidatures truffées d'erreurs de grammaire ou de fautes d'orthographe, l'ordinateur – lui-même illettré – pouvait difficilement en faire de même. Il pouvait en revanche corrélérer les candidatures refusées au cours des années passées aux lieux de naissance et, dans une moindre mesure, aux noms de famille. Les individus originaires de certains endroits – d'Afrique, du Pakistan ou de quartiers à forte immigration au Royaume-Uni – reçurent de la sorte un score global plus faible et ne furent pas conviés à un entretien. Un énorme pourcentage d'entre eux étaient des gens de couleur. Les opérateurs humains avaient aussi rejeté les candidatures féminines, avec bien trop souvent la justification que leur carrière se verrait sans doute interrompue par les devoirs de la maternité. Naturellement, la machine les imita.

En 1988, la Commission pour l'égalité raciale du gouvernement britannique estima l'école de médecine coupable dans sa politique d'admission d'une discrimination raciale et sexuelle. Jusqu'à 60 des 2 000 candidats annuels pouvaient, selon la commission, s'être vu refuser un entretien uniquement en raison de leur appartenance ethnique ou de leur sexe.

La solution, pour les statisticiens de Saint-George – et pour ceux travaillant dans d'autres secteurs –, serait d'élaborer une version numérique de l'audition à l'aveugle, en éliminant les marqueurs indirects tels que l'origine géographique, le sexe, l'ethnie ou le nom pour se concentrer exclusivement sur des données touchant à la formation médicale. L'essentiel, c'est d'analyser les compétences que chaque candidat apporte à l'école, et non de le juger par comparaison à d'autres qui semblent partager certaines similitudes. À Saint-George, un minimum de réflexion et d'inventivité aurait pu en outre résoudre les défis auxquels les femmes et les étrangers se trouvaient confrontés. L'article du *British Medical Journal* accompagnant l'avis de la commission le disait bien. Si des questions liées à la maîtrise de la langue ou à la garde des enfants posaient problème vis-à-vis de candidats par ailleurs solides, la solution n'était pas de les refuser mais plutôt de leur apporter une aide – cours d'anglais ou service de garderie sur place – pour leur permettre de surmonter ces difficultés.

C'est un point sur lequel je reviendrai dans les prochains chapitres : nous avons vu à maintes reprises que les modèles mathématiques pouvaient passer au crible quantité de données afin de repérer les individus susceptibles de faire face à d'importants défis, liés aussi bien à la criminalité qu'à la pauvreté ou à l'éducation. Il appartient à la société, soit d'exploiter ces renseignements pour les rejeter et les punir, soit de leur tendre la main et de leur procurer les ressources dont ils ont besoin. La portée et l'efficience qui confèrent aux ADM un caractère si néfaste peuvent aussi servir à aider les gens. Tout dépend de l'objectif choisi.

\*\*\*

Dans ce chapitre, nous avons examiné jusqu'ici des modèles qui filtrent les candidats à l'embauche. Pour la plupart des entreprises, ces ADM sont conçues de manière à diminuer les coûts administratifs et à réduire le risque de mauvaises recrues (ou à éviter celles qui pourraient nécessiter un complément de formation). L'objectif de ces filtres est en résumé d'économiser de l'argent.

Les départements RH, bien entendu, aspirent aussi à réaliser des économies grâce aux choix de recrutement qu'ils effectuent. La rotation de la main-d'œuvre, ce qu'on appelle couramment le « *churn* », représente l'une des dépenses les plus lourdes pour une entreprise. Le remplacement d'un employé payé 50 000 dollars par an lui coûte, selon le Center for American Progress, environ 10 000 dollars, soit 20 % de son salaire annuel. Celui d'un employé de haut niveau peut coûter plusieurs fois cette somme – jusqu'à deux années de salaire.

Beaucoup de modèles de recrutement cherchent naturellement à calculer la probabilité que chaque candidat reste dans l'entreprise. Evolv, Inc., aujourd'hui filiale de Cornerstone OnDemand, a aidé Xerox à dénicher des recrues potentielles pour son centre d'appel, qui emploie plus de 40 000 personnes. La modélisation du *churn* prenait en compte, comme on peut s'y attendre, certains indicateurs tels que la durée de maintien moyenne dans les postes antérieurs. Mais d'intrigantes corrélations furent également mises au jour. Les gens que le système classait comme « créatifs » avaient tendance à rester plus longtemps où ils se trouvaient, tandis que ceux qui affichaient un score de « curiosité » élevé étaient davantage enclins à se tourner vers d'autres opportunités.

La corrélation la plus problématique concernait cependant la situation géographique. Les candidats les plus éloignés de leur poste avaient d'autant plus de chances de le quitter. C'était parfaitement logique : les longs trajets domicile-travail s'avèrent en effet pénibles. Mais les responsables de Xerox notèrent une autre corrélation : beaucoup de ceux qui subissaient ces longs trajets venaient de quartiers pauvres. L'entreprise, et c'est tout à son honneur, supprima donc de son modèle cette donnée discriminante, fortement corrélée au *churn*. Elle sacrifia un peu d'efficacité au profit de l'équité.

Alors que l'analyse du *churn* se focalise sur les candidats qui risquent le plus de faire défaut, la tâche la plus vitale pour les départements RH, d'un point de vue stratégique, consiste à repérer de futures « vedettes » – des personnes dont l'intelligence, l'inventivité et le dynamisme peuvent changer la destinée d'une entreprise tout entière. Dans les échelons supérieurs de l'économie, les entreprises sont en quête d'employés qui font preuve de créativité et possèdent une bonne disposition au travail d'équipe. Le défi pour un modélisateur vise donc à déceler, parmi le vaste univers du Big Data, les éléments d'information corrélés avec l'originalité et les aptitudes sociales.

Les CV seuls ne suffisent certainement pas. La plupart des points énumérés – l'université prestigieuse, les récompenses, et même les compétences – ne rendent que grossièrement compte de la capacité à fournir un travail d'excellente qualité. Bien qu'il existe sans nul doute quelque corrélation entre l'habileté technique et le fait d'être diplômé d'une grande école, elle est loin d'être parfaite. Beaucoup de programmeurs de talent viennent d'ailleurs – songez par exemple aux jeunes lycéens passionnés d'informatique. Les CV regorgent en outre d'exagérations, voire parfois de mensonges. À l'aide d'une rapide recherche sur LinkedIn ou Facebook, un système de recrutement peut aller voir plus loin, en identifiant certains amis et collègues du candidat. Mais il reste difficile, sur cette base-là, de prédire que tel ou tel ingénieur répondra idéalement aux besoins d'un cabinet de conseil d'une douzaine de collaborateurs situé à Palo Alto ou à Fort Worth<sup>6</sup>. Pour trouver la personne à même d'occuper une telle position, il faut balayer un éventail de données bien plus large et recourir à un modèle plus ambitieux.

Gild, une start-up basée à San Francisco, fait figure de pionnier dans ce domaine. Poussant l'étude d'une recrue potentielle bien au-delà de son

université d'origine ou de son curriculum vitæ, Gild parcourt des millions de sites d'emploi et analyse ce qu'elle appelle les « données sociales » de chaque personne. L'entreprise élabore des profils de candidats pour ses clients, principalement des sociétés technologiques, et les actualise à mesure que ces candidats acquièrent de nouvelles compétences. Gild prétend même pouvoir prédire qu'un employé vedette est susceptible de changer de poste, et alerter ses sociétés clientes quand vient le bon moment de lui faire une offre. Le modèle de Gild cherche toutefois non seulement à quantifier mais aussi à *qualifier* le « capital social » de chaque travailleur. Quelle est l'importance de cette personne dans la communauté des programmeurs ? Partage-t-elle du code et contribue-t-elle à des projets ? Imaginons qu'un codeur brésilien habitant à São Paulo – on l'appellera Pedro – passe toutes ses soirées de la fin du repas à une heure du matin en relation avec des collègues codeurs du monde entier, s'évertuant à résoudre des problèmes de *cloud* ou se creusant la tête pour des algorithmes de jeu sur des sites comme GitHub ou Stack Overflow. Le modèle pourrait tenter de jauger la passion de Pedro (qui obtiendrait probablement un score élevé) et son degré d'implication vis-à-vis des autres. Il évaluerait également la compétence et l'importance sociale de ses contacts. Ceux qui sont le plus suivis compteraient davantage. Si son principal contact en ligne se trouvait être Sergey Brin, cofondateur de Google, ou Palmer Luckey, fondateur de la firme de réalité virtuelle Oculus VR, le score social de Pedro battrait sans doute des records.

Mais les modèles comme celui de Gild tirent rarement de leurs données des signaux aussi explicites. Ils déploient donc un filet plus large, cherchant partout où il est possible d'en trouver des corrélations avec la notoriété professionnelle. Et forte d'une base de données de plus de 6 millions de programmeurs, Gild est capable de dégager toutes sortes de tendances. Vivienne Ming, sa responsable scientifique, a indiqué lors d'un entretien pour *Atlantic Monthly* qu'elle avait ainsi découvert une volée de talents qui fréquentaient un site particulier de manga japonais. Si Pedro passe du temps sur ce site de bandes dessinées, cela ne permettra évidemment pas de prédire qu'il est une super-vedette. Mais cet aspect fera bel et bien monter son score.

Dans le cas de Pedro, c'est tout à fait logique. Mais certains travailleurs peuvent avoir d'autres activités en dehors d'Internet, qu'un algorithme même le plus sophistiqué ne saurait déduire – du moins pas à l'heure

actuelle. Ils s'occupent peut-être de jeunes enfants, ou fréquentent un groupe de lecture. Le fait de ne pas passer six heures chaque soir à discuter de mangas ne devrait pas desservir d'éventuelles recrues. Et si ce site, comme l'essentiel de la sphère technologique, se trouve dominé par les hommes et peut présenter une tonalité sexuelle, bon nombre de femmes dans cette industrie éviteront en outre probablement de s'y rendre.

En dépit des problèmes posés, Gild n'est qu'un acteur parmi d'autres. Cette société n'a pas l'influence d'un géant mondial et n'est pas en position de définir une norme unique pour l'ensemble du secteur. Comparée à certaines des horreurs que nous avons vues – les publicités prédatrices, qui plongent des familles dans l'endettement, ou les tests de personnalité qui bloquent les perspectives d'embauche des individus –, Gild fait figure d'agneau. Son modèle prédictif appartient à une catégorie qui vise davantage à récompenser les gens qu'à les punir. L'analyse s'avère certainement inégale : des vedettes potentielles sont sans doute négligées. Mais je ne pense pas que les chasseurs de talents atteignent encore le niveau d'une ADM.

Il faut cependant noter que ces modèles de recrutement et d'« intégration » évoluent en permanence. L'univers des données ne cesse de s'étendre, à mesure que chacun d'entre nous génère un flux toujours croissant d'informations actualisées à propos de sa vie. Toutes ces données profiteront à nos employeurs potentiels, en leur fournissant le moyen de mieux nous connaître.

Les connaissances ainsi recueillies seront-elles vérifiées, ou serviront-elles simplement à justifier l'état actuel des choses et à renforcer les préjugés ? Quand je vois comment certaines entreprises exploitent les données, sans aucune rigueur et dans un but intéressé, cela me rappelle souvent la phrénologie, une pseudoscience qui fit brièvement fureur au XIX<sup>e</sup> siècle. Les phrénologues passaient leurs doigts sur le crâne du patient, à la recherche de bosses et de creux, liés, pensaient-ils, aux traits de personnalité prenant source dans vingt-sept zones du cerveau. Leur diagnostic concordait d'ordinaire avec les observations effectuées. Si un patient souffrait d'anxiété morbide ou d'alcoolisme, la palpation du crâne révélait généralement des irrégularités corrélées avec ces symptômes – ce qui en retour renforçait la foi dans l'aspect scientifique de la phrénologie.

La phrénologie était un modèle qui s'appuyait sur un non-sens pseudoscientifique pour émettre des jugements d'autorité. Pendant des



dizaines d'années, il n'a jamais été remis en cause. Le Big Data peut très bien tomber dans le même piège. Des modèles comme ceux qui ont rejeté la candidature de Kyle Behm et empêché les étudiants étrangers d'entrer à Saint-George possèdent le pouvoir d'exclure les gens, même quand la « science » qu'ils recèlent ne se résume guère qu'à un monceau d'hypothèses non vérifiées.

---

1. Vanderbilt est une université privée, classée comme étant la 14<sup>e</sup> meilleure université des États-Unis.

2. La plainte a depuis été intégrée dans une enquête plus large de la Commission américaine pour l'égalité des chances en matière d'emploi.

3. En France les tests de personnalité (comme le PAPI, le SOSIE ou le MBTI) sont également utilisés par un nombre croissant d'employeurs dans le cadre du recrutement.

4. Distinction remise au meilleur joueur disputant sa première année dans la ligue de basket-ball nord-américaine (*NdT*).

5. Beaucoup de lycéens qui prévoient d'entrer à l'université effectuent certes un ou deux jobs d'été, payés au salaire minimum. Mais lorsqu'ils vivent une expérience déplorable, ou se voient sous-estimés par une ADM arbitraire, cela les encourage à s'appliquer dans leurs études et tirer un trait sur des emplois aussi détestables (*note de l'auteur*).

6. Situées respectivement en Californie et au Texas, ces deux villes hébergent les sièges sociaux de plusieurs compagnies de très grande envergure, comme Facebook ou American Airlines, ainsi que des universités de renom, comme Stanford.

# L'ANGOISSE

## Sur le lieu de travail

# AU QUOTIDIEN

1001Ebooks

**L**es salariées de grandes sociétés américaines ont récemment inventé une nouvelle expression : le « clopening ». Elle désigne le fait pour un employé de rester tard le soir pour fermer la boutique ou le café (« *close* »), puis de revenir l'ouvrir au bout de quelques heures (« *open* »), avant le lever du soleil. Du point de vue de l'entreprise, il paraît souvent judicieux en termes de logistique qu'une seule et même personne se charge de la fermeture et de l'ouverture. Mais cette organisation entraîne un manque de sommeil et des horaires délirants.

L'extrême irrégularité des plannings devient de plus en plus courante, et touche tout particulièrement les travailleurs peu rémunérés qu'embauchent des entreprises comme Starbucks, McDonald's ou la chaîne de supermarchés Walmart. L'absence de notification préalable ne fait qu'aggraver le problème. Beaucoup d'employés découvrent à peine un ou deux jours à l'avance qu'ils vont devoir travailler un mercredi soir, ou assurer l'heure de pointe du vendredi. Leur vie s'en trouve complètement désorganisée, et la garde des enfants, bouleversée. On prend ses repas quand on peut, et de même pour dormir.

Ces horaires irréguliers sont un produit de l'économie des données. Dans le chapitre précédent, nous avons observé de quelle manière les ADM triaient les candidats à l'embauche, en écartaient certains et en ignoraient encore davantage. Nous avons vu que le logiciel utilisé incorporait souvent dans son code des préjugés néfastes, et apprenait à se comporter de façon injuste en se basant sur des données historiques. Nous poursuivons ici notre parcours jusqu'à l'emploi lui-même, où des ADM focalisées sur l'efficacité traitent les travailleurs comme les simples rouages d'une machine. Le

*clopening* n'est que l'un des fruits de cette tendance, qui a toutes les chances de se développer à mesure que la surveillance s'étendra sur le lieu de travail, donnant à l'économie des données d'autant plus de grain à moudre.

Pendant des décennies, avant que les entreprises ne croulent littéralement sous les données, l'élaboration des plannings n'a rien eu de scientifique. Imaginez une quincaillerie familiale faisant travailler ses vendeurs de 9 heures à 17 heures, six jours sur sept. Les parents ont une fille, qui part un an à l'université. Lorsqu'elle revient pour l'été, elle pose sur l'entreprise un regard nouveau. Elle s'aperçoit que pratiquement personne ne fréquente le magasin le mardi matin. Le vendeur surfe sur son téléphone, sans jamais être interrompu par la visite d'un client. C'est une perte de chiffre d'affaires. Le samedi, à l'inverse, les clients attendent longuement leur tour en grommelant.

Ces observations lui procurent des données précieuses, et elle aide alors ses parents à remodeler l'activité en conséquence. Ils commencent par fermer la boutique le mardi matin, et recrutent un temps partiel pour mieux gérer l'affluence du samedi. Ces changements ajoutent un peu d'intelligence au système établi, parfois absurde et inflexible.

Avec le Big Data, cette étudiante de première année se voit remplacée par des légions de titulaires d'un doctorat armés de puissants ordinateurs. Les entreprises peuvent désormais analyser le trafic client pour calculer le nombre exact d'employés nécessaires à chaque heure de la journée. Le but consiste bien entendu à dépenser le moins possible, ce qui suppose de maintenir l'effectif au strict minimum tout en s'assurant que des renforts restent disponibles en vue des périodes de pointe.

On pourrait penser que les mêmes schémas se répètent de semaine en semaine, et que les entreprises n'ont qu'à apporter quelques ajustements aux horaires établis, à l'instar des propriétaires de notre hypothétique quincaillerie. Mais de nouveaux logiciels de planning proposent des options bien plus sophistiquées. Ils analysent de nouveaux flux de données qui évoluent en permanence, allant des conditions météorologiques jusqu'aux habitudes piétonnières. Un après-midi pluvieux poussera ainsi sans doute les gens à délaisser parcs et jardins au profit des cafés. Ces derniers auront donc besoin d'un peu plus de personnel, au moins pour une heure ou deux. Un match de football américain interlycées pourra entraîner le vendredi soir un trafic piéton accru sur la rue principale – mais uniquement avant et après

la rencontre, pas pendant. Le volume des échanges sur Twitter suggère que 26 % d'acheteurs supplémentaires se précipiteront cette année sur les soldes du Black Friday, par rapport à l'an dernier. D'une heure à l'autre, les circonstances changent et la main-d'œuvre doit être déployée de manière à répondre aux fluctuations de la demande. Sinon, l'entreprise perd de l'argent.

Naturellement, les sommes ainsi économisées sortent tout droit des poches des employés. Sous le régime inefficace qui prévalait auparavant, les travailleurs bénéficiaient non seulement d'horaires prévisibles mais aussi de divers temps morts. On peut arguer qu'ils profitaient de l'inefficience du système : certains parvenaient à lire pendant le service, voire à étudier. Aujourd'hui, avec un logiciel chargé de chorégraphier le travail, chaque minute doit être remplie. Et ces minutes tomberont au moment où le programme l'exige, même si l'employé se voit alors imposer un *clopening* entre le vendredi et le samedi.

En 2014, le *New York Times* a publié un reportage sur une mère célibataire surmenée, Jannette Navarro, qui tentait de suivre des études en travaillant comme barista chez Starbucks, tout en s'occupant d'un enfant de 4 ans. Son planning toujours différent, incluant un *clopening* de temps à autre, lui rendait la vie presque impossible et l'empêchait d'accéder à une garde de jour régulière. Elle avait dû arrêter l'université. La seule chose qu'elle pouvait planifier, c'était le travail. Et son histoire n'avait rien d'exceptionnel. Selon les données du gouvernement américain, deux tiers des travailleurs dans le secteur de la restauration et plus de la moitié dans le commerce de détail ne sont informés des modifications de planning qu'une semaine au mieux au préalable – et souvent à peine un ou deux jours avant, ce qui les contraint à s'organiser hâtivement pour le transport ou la garde des enfants.

Dans les semaines qui suivirent la publication de l'article, les grandes entreprises mentionnées annoncèrent qu'elles allaient corriger leurs pratiques en matière de planning. Embarrassés par ce récit, les employeurs promirent d'intégrer à leur modèle une contrainte spécifique. Ils élimineraient le *clopening* et apprendraient à se contenter d'une optimisation un peu moins poussée. Starbucks, dont l'image de marque repose plus que d'autres sur le juste traitement des travailleurs, alla encore plus loin, indiquant que l'entreprise adapterait son logiciel pour éviter à ses

130 000 baristas de vivre un cauchemar au niveau de leurs horaires. Tous les plannings seraient désormais publiés au moins une semaine à l'avance.

Un an plus tard, d'après un second reportage du *Times*, Starbucks n'avait toutefois pas réussi à atteindre ces objectifs ni même à supprimer le *clopening*. Le problème venait de ce que le recours à un effectif minimum était ancré dans sa culture. Au sein de nombreuses entreprises, le salaire des managers est subordonné à l'efficacité de leur personnel, mesurée en fonction du chiffre d'affaires par heure ouvrée. Les logiciels de planning les aident à améliorer cet indicateur et donc à augmenter leur rémunération. Même quand l'encadrement leur dit de relâcher la pression, ils renâclent souvent. C'est contraire à tout ce qu'on leur a enseigné. Chez Starbucks, qui plus est, comme l'a expliqué un employé, si un manager dépasse son « budget de main-d'œuvre », un responsable régional est alors alerté. Et cela peut déboucher sur un rapport. Il s'avère en général plus facile de modifier simplement le planning de quelqu'un, même si l'engagement de l'entreprise à respecter un préavis d'une semaine s'en trouve bafoué.

Au bout du compte, les business models d'entreprises cotées en Bourse telles que Starbucks sont conçus dans l'optique d'augmenter le résultat net – ce que reflètent leur culture interne et leurs mesures d'incitation, mais aussi de plus en plus leurs logiciels de gestion opérationnelle. Si ces derniers autorisent des ajustements, comme le fait Starbucks, ils iront probablement dans le sens d'une augmentation des profits.

La technologie de la planification trouve pour l'essentiel son origine dans une branche puissante des mathématiques appliquées appelée « recherche opérationnelle », ou RO. Pendant des siècles, les mathématiciens ont exploité les rudiments de la RO pour aider les agriculteurs à planifier leurs semailles et les ingénieurs civils à tracer des routes, permettant une circulation efficace des biens et des personnes. Mais la discipline n'a vraiment décollé qu'à partir de la Seconde Guerre mondiale, quand les forces armées américaines et britanniques ont enrôlé des équipes de mathématiciens afin d'optimiser l'utilisation de leurs ressources. Les Alliés suivaient l'évolution d'un « ratio d'échange », qui comparait sous diverses formes les ressources dépensées de leur côté et celles détruites chez l'ennemi. Lors de l'opération « Starvation » (« Famine »), menée de mars à août 1945, le 21<sup>e</sup> groupe de bombardement se vit confier la mission de détruire les navires marchands japonais, afin d'empêcher les vivres et autres marchandises de gagner sans encombre les rives de l'archipel nippon.

Les équipes de RO s'efforcèrent de réduire au maximum, pour chaque bâtiment japonais coulé, le nombre d'aéronefs mouilleurs de mines abattus au cours de l'opération. Elles atteignirent un « ratio d'échange » de plus de 40 contre 1 – 15 appareils seulement perdus pour envoyer par le fond 606 bateaux japonais. Le rendement fut jugé remarquable, et on le devait en partie à leur travail.

Au lendemain de la guerre, les grandes entreprises (ainsi que le Pentagone) ont investi d'énormes ressources dans la RO. La science de la logistique a radicalement transformé notre manière de produire des biens et de les commercialiser.

Dans les années 1960, les constructeurs automobiles japonais ont accompli une autre avancée majeure, en concevant un système de fabrication baptisé « juste-à-temps ». Au lieu de stocker des montagnes de volants ou de boîtes de vitesses et d'aller les chercher dans de vastes entrepôts, l'idée était que l'usine d'assemblage passerait commande des pièces au fur et à mesure de ses besoins, plutôt que de payer pour qu'elles restent inutilisées. Toyota et Honda établirent des chaînes complexes d'approvisionnement, chaque fournisseur livrant en continu les pièces demandées. Comme si l'industrie tout entière formait un organisme unique, doté de ses propres systèmes de contrôle homéostatique.

La méthode était extrêmement efficace et s'est vite répandue à travers le monde. Dans beaucoup de régions, les entreprises peuvent mettre en place des chaînes d'approvisionnement « juste-à-temps » en un tournemain. Ces modèles mathématiques sous-tendent également l'activité de firmes telles qu'Amazon, FedEx et UPS.

Les logiciels de planification peuvent être considérés comme une extension de l'économie du juste-à-temps. Mais au lieu de commander des lames de tondeuse ou des écrans de téléphone portable pour qu'ils arrivent à point nommé, ce sont des gens que l'on gère – des gens qui manquent en général cruellement d'argent. Et parce qu'ils en ont tant besoin, les entreprises ont alors le pouvoir de les soumettre aux diktats d'un modèle mathématique.

Il convient d'ajouter que lesdites entreprises prennent néanmoins des mesures afin de ne pas leur rendre la vie *trop* dure. Elles savent toutes au centime près le coût de remplacement d'un travailleur à bout de forces qui finit par démissionner. Ces chiffres figurent eux aussi dans leurs données. Comme on l'a vu dans le chapitre précédent, elles disposent en sus d'autres

modèles permettant de réduire le *churn*, un remplacement d'employés qui sape les profits et l'efficacité.

Le problème du point de vue des employés tient à la surabondance d'une main-d'œuvre bon marché. Les gens ont besoin de travailler, ce qui explique qu'un si grand nombre d'entre eux s'accrochent à des emplois rémunérés à peine au minimum légal<sup>1</sup>. Cette surabondance, conjuguée à un manque de syndicats puissants, ne laisse aux travailleurs pratiquement aucune capacité de négociation. Les grandes enseignes de distribution et de restauration peuvent ainsi les astreindre à des plannings toujours plus absurdes, sans avoir à subir un *churn* excessif. Leurs gains augmentent, tandis que la vie de leurs salariés devient de plus en plus infernale. Et puisque ces programmes d'optimisation existent partout, les travailleurs savent fort bien qu'un changement d'emploi n'a guère de chances d'améliorer leur sort. Combinées les unes aux autres, ces dynamiques procurent aux entreprises ce qui n'est pas loin de s'apparenter à une main-d'œuvre captive.

Personne, j'en suis sûr, ne s'étonnera que les logiciels de planning constituent à mes yeux l'une des ADM les plus épouvantables. Elle possède comme on l'a vu un poids considérable, et profite d'individus qui ont déjà du mal à joindre les deux bouts. Elle s'avère de surcroît totalement opaque. Les travailleurs n'ont souvent aucune idée du moment où ils seront appelés à leur poste. C'est un programme arbitraire qui le leur ordonne.

Les logiciels de planning génèrent également une boucle de rétroaction néfaste. Prenons le cas de Jannette Navarro. Ses horaires aléatoires l'empêchent de reprendre ses études, réduisant de ce fait ses débouchés professionnels et l'enfermant dans un vivier surpeuplé de travailleurs mal payés. La longueur et l'irrégularité des horaires permettent en outre difficilement aux employés de s'organiser ou de réclamer de meilleures conditions. Ils sont au contraire confrontés à un état d'angoisse et à un manque de sommeil accrus, qui provoquent de violentes sautes d'humeur et que l'on estime responsables de 13 % des morts sur la route. Pire encore, comme le logiciel est prévu pour faire économiser de l'argent aux employeurs, il limite souvent les heures ouvrées à moins de trente par semaine, de sorte que les salariés ne soient pas éligibles à une couverture par l'assurance maladie de l'entreprise. Et avec leurs horaires chaotiques, la plupart sont dans l'impossibilité de se dégager du temps pour un deuxième

emploi. Le logiciel semble presque expressément conçu pour pénaliser les travailleurs à faibles revenus et les maintenir dans cette situation.

Il condamne également une importante proportion d'enfants à ne pas avoir d'habitudes régulières. Ceux-ci croisent leur mère au petit déjeuner, les yeux rougis de fatigue, ou la voient partir précipitamment sans dîner, ou l'entendent se disputer avec la sienne pour savoir qui pourra s'occuper d'eux le dimanche matin. Cette vie désordonnée les affecte profondément. D'après une étude réalisée par l'Economic Policy Institute, une association de défense, « les jeunes enfants et adolescents dont les parents travaillent selon des plannings imprévisibles, ou en dehors des heures de bureau standard, sont davantage exposés à des carences cognitives et comportementales ». Ces parents se reprocheront peut-être la mauvaise conduite ou l'échec scolaire de leur enfant, mais le vrai coupable, dans beaucoup de cas, c'est la pauvreté, qui amène les travailleurs à accepter des emplois aux horaires erratiques – et les modèles de planification, qui mettent encore plus de pression sur des familles déjà en difficulté.

La racine du problème, comme avec tant d'autres ADM, réside dans les objectifs choisis par les modélisateurs. Le modèle est optimisé dans une perspective d'efficacité et de rentabilité, et non pour l'équité ou le bien de l'« équipe ». C'est bien entendu la nature même du capitalisme. Le chiffre d'affaires est l'oxygène des entreprises. Il les maintient en vie. De leur point de vue, il serait profondément stupide, voire anormal, de renoncer à de potentielles économies. Voilà pourquoi la société doit pouvoir s'appuyer sur des forces contraires – comme une vigoureuse couverture médiatique, capable de mettre en lumière les abus de cette course à l'efficacité et de jeter l'opprobre sur les entreprises pour les pousser à bien faire. Et lorsqu'elles ne s'y appliquent pas assez, à l'instar de Starbucks, la société doit les dénoncer encore et encore. Elle a aussi besoin d'instances de contrôle, pour les garder dans le droit chemin, de syndicats puissants pour organiser les travailleurs et mieux faire entendre leurs besoins et leurs griefs, et de responsables politiques prêts à voter des lois pour réfréner les pires excès des entreprises. Suite à l'article du *New York Times* paru en 2014, les parlementaires démocrates du Congrès ont rapidement élaboré des textes visant à brider les logiciels de planification. Mais face à une majorité républicaine farouchement opposée à toute régulation étatique, les chances de les voir adoptés étaient nulles. La législation envisagée fut tuée dans l'œuf.



\*\*\*

En 2008, à l'heure où la grande crise économique s'approchait, une entreprise de San Francisco baptisée Cataphora lança un système informatique qui permettait d'évaluer les employés du secteur technologique en fonction d'un certain nombre d'indicateurs, parmi lesquels leur aptitude à générer des idées. La tâche n'avait rien d'aisé. Les logiciels ont en effet beaucoup de mal à distinguer une idée d'une simple succession de mots. Si l'on y réfléchit, la différence n'est souvent qu'affaire de contexte. Les idées d'hier – que la terre soit ronde, ou simplement que les gens puissent aimer partager leurs photos sur des réseaux sociaux – sont les faits d'aujourd'hui. Chaque être humain perçoit le moment où une idée devient un fait établi, et sait également (en dépit de fréquents désaccords) si elle a été démentie ou abandonnée. Cette distinction déconcerte en revanche y compris la plus sophistiquée des intelligences artificielles. Le système de Cataphora devait donc se tourner vers les humains eux-mêmes pour lui servir de guide.

Dans sa chasse aux idées, le logiciel fouillait parmi les courriers électroniques et la messagerie instantanée de l'entreprise. Son hypothèse directrice était que les meilleures d'entre elles auraient tendance à se répandre plus abondamment à travers le réseau. Si les gens « copiaient-collaient » certains groupes de mots et se les échangeaient, ces mots constituaient alors probablement des idées, et le logiciel pouvait les quantifier.

Des difficultés se posaient, toutefois. Les idées n'étaient pas les seuls groupes de mots amplement partagés sur les réseaux sociaux. Les blagues se montraient par exemple extrêmement virales, et non moins déroutantes pour les systèmes informatiques. Les ragots se propageaient eux aussi comme une traînée de poudre. Blagues et ragots suivaient cependant certains schémas particuliers, si bien qu'on pouvait apprendre au programme à en filtrer au moins une partie. Au fil du temps, le système parvenait à repérer les groupes de mots les plus susceptibles de représenter des idées. Il en suivait la trace à travers le réseau, comptabilisant le nombre de fois où ils étaient copiés, mesurant leur répartition et identifiant leur source.

Très vite, le rôle des employés semblait se préciser. Certains, selon les conclusions du système, étaient générateurs d'idées. Sur son organigramme

du personnel, Cataphora les indiquait sous la forme d'un cercle, d'autant plus gros et plus foncé qu'ils en produisaient davantage. D'autres faisaient office de connecteurs. À l'image des neurones dans un réseau distribué, ils transmettaient l'information. Les plus efficaces conféraient aux associations de mots un caractère viral. Le système les colorait également en foncé.

Qu'il fût ou non capable d'évaluer efficacement la circulation des idées, le concept n'avait en soi rien de maléfaisant. Ce type d'analyse peut s'avérer judicieux pour déterminer ce que les gens savent, et les associer avec leurs collègues et collaborateurs les plus prometteurs. IBM et Microsoft exploitent justement à cette fin des programmes développés en interne. L'algorithme est très similaire à celui d'un site de rencontres (et les résultats sans doute souvent tout aussi inégaux).

Le Big Data a aussi été utilisé pour étudier la productivité des opérateurs de centres d'appels. Il y a quelques années, des chercheurs du MIT ont analysé le comportement des téléconseillers de Bank of America, afin de découvrir pourquoi certaines équipes étaient plus productives que d'autres. Ils ont accroché au cou de chaque employé un dispositif qualifié de « badge sociométrique ». L'électronique embarquée suivait leurs déplacements et mesurait également, toutes les seize millisecondes, le ton de leur voix et les gestes effectués. Quand les gens se regardaient, elle le détectait et enregistrait le temps que chacun passait à parler, à écouter et à interrompre l'autre. Quatre équipes de téléconseillers – quatre-vingts personnes au total – ont porté le badge pendant six semaines.

Le travail de ces employés était strictement régenté. On les dissuadait de bavarder, car ils étaient censés passer le plus de minutes possible au téléphone pour résoudre les problèmes des clients. Les pauses café étaient programmées une par une.

À leur grande surprise, les chercheurs constatèrent que l'équipe la plus rapide et la plus efficace était aussi celle qui échangeait le plus. Ses membres faisaient fi des règles et bavardaient beaucoup plus que les autres. Et quand l'ensemble des opérateurs furent incités à communiquer davantage, la productivité du centre d'appel grimpa en flèche.

Mais les études de données pistant le comportement des employés peuvent aussi servir à réduire la main-d'œuvre. Alors que la récession de 2008 ravageait l'économie, les responsables RH du secteur des technologies commencèrent à examiner les organigrammes de Cataphora dans une optique différente. Ils virent que certains travailleurs étaient représentés par

de grands cercles foncés, tandis que d'autres étaient plus pâles et plus petits. S'il fallait procéder à des licenciements – et c'était le cas dans la plupart des entreprises – il semblait judicieux de renvoyer d'abord ces derniers.

Étaient-ils pour autant réellement inutiles ? Nous retrouvons à nouveau ici une sorte de phrénologie numérique. Si le système désigne tel travailleur comme un piètre générateur d'idée ou un médiocre connecteur, ce verdict devient sa propre vérité. C'est le score qu'on lui donne.

Peut-être peut-on apporter des arguments contraires. Le travailleur symbolisé par un cercle pâle génère peut-être de fabuleuses idées, mais ne les partage pas sur le réseau. Il prodigue peut-être d'incalculables conseils à l'heure du déjeuner, ou sait briser les tensions dans le bureau au moyen d'une plaisanterie. Il est peut-être apprécié de tous. Tout ceci revêt un intérêt considérable dans le cadre professionnel. Mais les systèmes informatiques peinent à trouver des équivalents numériques pour ce genre de compétences relationnelles. Les données qui s'y rapportent ne sont tout simplement pas collectées, et leur attribuer une valeur s'avère de toute façon difficile. Il est en général plus simple de les exclure du modèle.

Le système identifie donc ceux qui apparaissent comme médiocres. Et un bon nombre d'entre eux ont perdu leur emploi durant la récession. Cet aspect-là est à lui seul injuste. Mais ce qui est pire, c'est que les systèmes comme celui de Cataphora ne reçoivent qu'une faible remontée d'information. Un employé identifié comme médiocre, et par suite congédié, aura peut-être trouvé un nouveau poste et élaboré plusieurs brevets. Ces données ne sont d'ordinaire pas recueillies. Le système ne peut dès lors soupçonner qu'il a totalement méjugé cette personne, ni même un millier d'autres.

Cela pose un vrai problème, car les chercheurs ont besoin de ce retour d'information sur leur erreur – en l'espèce, la présence de faux négatifs – pour mener une analyse scientifique et déterminer ce qui n'a pas bien fonctionné, ce qui a été mal interprété, et les données qui ont été ignorées. C'est de cette façon que les systèmes apprennent et qu'ils gagnent en intelligence. Pourtant, comme on l'a vu, quantité d'ADM, des modèles de prédiction du récidivisme jusqu'à la notation des enseignants, génèrent allègrement leur propre réalité. Les managers supposent les scores obtenus suffisamment exacts pour être utiles, et l'algorithme rend les décisions délicates faciles à trancher. Ils peuvent ainsi renvoyer des employés et

réduire les coûts en rejetant la responsabilité de leurs choix sur un chiffre objectif, que celui-ci soit correct ou pas.

La société Cataphora est restée d'une taille modeste, et son modèle d'évaluation des employés ne constituait pour elle qu'une activité secondaire – l'essentiel de son travail portait sur l'identification de profils de fraude ou de délit d'initié à l'intérieur des entreprises. Elle a mis la clé sous la porte en 2012, et son logiciel a été vendu à une start-up du nom de Chenope. Mais les systèmes comme le sien possèdent la capacité de devenir de véritables ADM. Ils peuvent générer une fausse représentation des individus, et les sanctionner sans la moindre preuve que leur score soit corrélé à la qualité de leur travail.

Ce type de logiciel annonce l'essor des ADM dans un domaine nouveau. Durant quelques décennies, il a pu sembler que les travailleurs des secteurs de l'industrie et des services étaient les seuls susceptibles d'être modélisés et optimisés, tandis que ceux qui maniaient des idées, des juristes jusqu'aux ingénieurs chimistes, avaient la possibilité – au moins dans leur travail – d'échapper aux ADM. Le système de Cataphora fut un signe avant-coureur, indiquant que ce ne serait pas le cas. Et de fait, dans l'ensemble du secteur technologique, nombre d'entreprises s'efforcent aujourd'hui d'optimiser les « cols blancs » en étudiant leurs habitudes de communication. Les géants de la haute technologie, parmi lesquels Google, Facebook, Amazon, IBM et beaucoup d'autres, suivent énergiquement cette voie-là.

Pour l'instant, du moins, la diversité des systèmes se montre bienvenue. Elle laisse en tout cas espérer que les travailleurs rejetés par un modèle puissent être appréciés par un autre. À terme, un standard industriel finira néanmoins par émerger, et nous serons alors tous en danger.

\*\*\*

En 1983, l'administration Reagan lança une vibrante alerte sur l'état des écoles américaines. Dans un rapport intitulé *A Nation at Risk* (« Une nation en péril »), une commission présidentielle avertit qu'une « marée montante de médiocrité » dans les établissements scolaires menaçait « l'avenir même des États-Unis en tant que nation et en tant que peuple ». Le rapport ajoutait que « si une puissance étrangère hostile » avait tenté d'imposer à l'Amérique des écoles aussi mauvaises, « elle aurait fort bien pu y voir un acte de guerre ».

L'indice le plus remarquable de cet échec résidait dans l'apparent effondrement des scores à l'examen SAT. Entre 1963 et 1980, le score aux épreuves de lecture et de rédaction avait baissé de 50 points, et celui en mathématiques, de 40 points. L'aptitude des Américains à rivaliser au sein de l'économie mondiale dépendait de leurs compétences, et celles-ci semblaient se détériorer.

À qui fallait-il imputer ce triste état de fait ? Le rapport ne laissait planer aucun doute là-dessus : les enseignants en étaient responsables. Il appelait à passer à l'action, c'est-à-dire à tester les élèves – et à utiliser les résultats obtenus pour identifier les enseignants peu performants. Comme on l'a vu dans l'introduction, cette pratique peut leur coûter leur poste. Sarah Wysocki, l'institutrice de Washington renvoyée après que sa classe eut affiché des scores étonnamment bas, fut la victime de ce genre de test. Mon but en relatant cette histoire était de montrer une ADM à l'œuvre, et à quel point elle pouvait être arbitraire, injuste et sourde à tout recours.

Mais outre le fait de s'occuper des enfants et de les éduquer, les enseignants sont aussi évidemment des employés, et je voudrais ici approfondir un peu l'étude des modèles qui évaluent leurs performances, car ceux-ci pourraient bien se propager à d'autres pans de la population active. Prenons le cas de Tim Clifford, professeur d'anglais exerçant en collège dans la ville de New York, avec vingt-six ans d'expérience. Il y a de cela quelques années, Clifford apprit qu'il avait subi un échec terrible lors d'une évaluation basée sur le modèle dit de la valeur ajoutée, similaire à l'évaluation qui avait conduit au renvoi de Sarah Wysocki. Il avait obtenu le score abyssal de 6 sur 100.

Clifford fut anéanti. « Je ne voyais pas comment il était possible, en ayant travaillé si dur, de récolter de si mauvais résultats », m'a-t-il raconté plus tard. « Pour être honnête, quand j'ai pris connaissance de mon score lamentable, j'ai d'abord eu honte et n'en ai parlé à personne pendant un ou deux jours. J'ai découvert cependant que deux autres enseignants de mon école avaient obtenu un score encore plus faible. J'y ai puisé le courage de partager mes résultats, car je voulais qu'ils sachent qu'ils n'étaient pas seuls. »

Si Clifford n'avait pas été titulaire, il aurait pu être remercié dès cette année-là. « Même en tant que titulaire, m'expliqua-t-il, un mauvais score plusieurs années de suite fait forcément de vous une cible, dans une certaine mesure. » Lorsque des enseignants titulaires enregistrent de mauvais scores,

cela vient de surcroît conforter les chantres de la réforme scolaire, qui font valoir que la sécurité de l'emploi protège les éducateurs incompetents. Clifford avait abordé la rentrée suivante avec appréhension.

Le modèle de la valeur ajoutée lui avait attribué une note éliminatoire, mais sans lui fournir aucun conseil pour l'améliorer. Clifford continua donc d'enseigner comme il l'avait toujours fait, en espérant que tout irait pour le mieux. L'année suivante, il décrocha un score de 96.

« On pourrait croire que j'en étais ravi, mais ce n'était pas le cas. Je savais que mon mauvais score était infondé, et ne pouvais donc guère me réjouir d'obtenir un score élevé découlant de la même formule fautive. Les 90 points d'écart entre les deux me firent simplement réaliser combien l'utilisation du modèle de la valeur ajoutée en matière d'éducation était en tout point ridicule. »

Infondé : voilà bien le terme qui convient. L'interprétation erronée des statistiques traverse en réalité toute l'histoire de l'évaluation des enseignants. Le problème a pour origine une bourde majeure dans l'analyse initiale du rapport *Nation at Risk*. Il est apparu en effet que le jugement de ces chercheurs, qui dénonçaient une catastrophe nationale, reposait lui-même sur une erreur fondamentale qu'un étudiant de première année aurait dû détecter. S'ils souhaitaient illustrer les défaillances de l'enseignement aux États-Unis, leur propre méprise dans la lecture des statistiques aurait en fait pu servir de preuve numéro un.

Sept ans après la publication en grande pompe du rapport, des chercheurs des Sandia National Laboratories (qui comptent parmi les principaux laboratoires nationaux du département américain de l'Énergie) examinèrent à nouveau les données rassemblées pour le rédiger. En matière de statistiques, ces gens-là étaient tout sauf des amateurs – ils conçoivent et entretiennent des armes nucléaires –, et ils eurent tôt fait de découvrir l'erreur. En moyenne, les scores des examens SAT avaient certes bel et bien chuté. Néanmoins, au cours de ces dix-sept années, le nombre d'élèves candidats au test avait explosé. Les universités ouvraient de plus en plus leurs portes aux étudiants pauvres et issus des minorités. Les perspectives s'élargissaient, signe de réussite sociale. Mais cet afflux de nouveaux venus faisait naturellement baisser la moyenne des scores. Pourtant, quand les statisticiens répartissaient la population par tranches de revenus, les scores individuels de chaque groupe, des plus pauvres aux plus riches, se révélaient en hausse.

En statistique, ce phénomène porte le nom de « paradoxe de Simpson » : quand un corpus de données affiche une tendance d'ensemble, mais que, une fois divisé en sous-groupes, c'est la tendance inverse qui apparaît pour chacun d'entre eux. La conclusion accablante du rapport *Nation at Risk*, celle qui avait impulsé tout le mouvement d'évaluation des enseignants, procédait donc d'une interprétation gravement inexacte des données.

Les scores divergents de Tim Clifford résultaient d'un autre usage bâclé des statistiques, beaucoup trop courant, celui-ci. Les scores des enseignants, calculés sur la base de tests, ne mesuraient *rien*. Une telle affirmation pourrait sembler exagérée. Le fait est que les enfants passaient des tests, et que leurs scores étaient corrélés à ceux de Clifford. Mais ces derniers, aussi bien l'humiliant 6 que le gratifiant 96, reposaient presque entièrement sur des approximations si fragiles qu'ils en devenaient au fond aléatoire.

Le problème venait de ce que les administrateurs, dans leur quête d'équité, avaient perdu de vue l'exactitude. Ils comprenaient bien qu'il n'était pas juste, lorsque des enfants de médecins et d'avocats se dirigeaient vers des universités d'élite, d'accorder un trop grand mérite aux enseignants des écoles riches. Ceux des districts pauvres ne devaient pas non plus être tenus aux mêmes normes de réussite. On ne pouvait attendre qu'ils accomplissent des miracles.

Plutôt que d'évaluer les enseignants sur une échelle absolue, ils essayèrent donc d'ajuster le modèle en tenant compte des inégalités sociales. Au lieu de comparer les élèves de Tim Clifford à ceux d'autres quartiers, ils compareraient ces élèves à des modélisations prévisionnelles d'eux-mêmes. Chacun d'entre eux se voyait prédire un score. S'il le dépassait, sa réussite était mise au crédit de l'enseignant. S'il ne l'atteignait pas, c'est à ce dernier qu'incombait la faute. La méthode vous semble primitive ? C'est parce qu'elle l'est.

D'un point de vue statistique, dans leurs tentatives d'affranchir les tests des notions de classe sociale et de couleur, les administrateurs sont passés d'un modèle primaire à un modèle secondaire. Au lieu de baser les scores sur une évaluation directe des élèves, ils se sont appuyés sur ce qu'on appelle le « terme d'erreur » – l'écart entre les résultats et les attentes. D'un point de vue mathématique, il s'agit là d'une proposition beaucoup plus imprécise. Puisque les attentes dérivent elles-mêmes des statistiques, cela revient à ajouter des suppositions aux suppositions. Il en ressort un modèle

produisant quantité de résultats aléatoires, ce que les statisticiens qualifient de « bruit ».

On serait tenté de croire que l'étude d'une large population permettrait de dégager des scores plus nets. La ville de New York, avec 1,1 million d'élèves dans l'enseignement public, devrait après tout fournir un corpus de données suffisamment vaste pour générer des prédictions pertinentes. Si 80 000 collégiens de dernière année passent le test, ne serait-il pas possible dans ce cas d'établir des moyennes fiables pour les établissements en difficulté, ceux qui excellent et ceux qui sont entre les deux ?

Certes. Et si Tim Clifford enseignait à un grand échantillon d'élèves, disons 8314, on pourrait alors avec raison évaluer cette cohorte par rapport à la moyenne de l'année précédente, et en tirer certaines conclusions. Les grands nombres compensent les exceptions et les aberrations statistiques. Des tendances devraient théoriquement apparaître. Mais il est presque impossible que l'évolution d'une classe de vingt-cinq ou trente élèves corresponde à celle d'une population plus étendue. Si une classe comporte certains types d'élèves, elle aura tendance à progresser plus vite que la moyenne, tandis que d'autres progresseront plus lentement. Clifford n'avait reçu à peu près aucune information sur cette ADM opaque qui lui attribuait des scores aussi divergents, mais il présumait que la variation dans la composition de ses classes n'y était pas étrangère. L'année où il avait obtenu un mauvais score, il avait eu « à la fois beaucoup d'élèves en enseignement adapté et beaucoup de très bons éléments ». Or, d'après lui, « enseigner aux plus faibles ou aux meilleurs – ou aux deux en même temps – génère des problèmes. Les scores des premiers sont difficiles à améliorer, en raison de leurs troubles d'apprentissage, et ceux des seconds, parce qu'ils sont déjà élevés et qu'il reste peu de marge de progrès ».

L'année suivante, il avait hérité d'une combinaison d'élèves différente, situés pour une plus grande part entre ces deux pôles. Et les résultats donnèrent l'impression qu'un piètre enseignant était devenu spectaculairement bon. Ces situations s'avéraient beaucoup trop fréquentes. Une analyse réalisée par Gary Rubinstein, blogueur et formateur, montra qu'un enseignant sur quatre pratiquant la même matière plusieurs années de suite enregistrerait un écart de 40 points. Ce constat laissait à penser que les données d'évaluation étaient quasi aléatoires. Ce n'était pas les performances des enseignants qui fluctuaient ainsi de manière aussi brutale, mais la note générée par une ADM défailante.



Bien que les scores qu'elle produit ne signifient rien, la modélisation de la valeur ajoutée exerce un impact envahissant et délétère. « J'ai vu de grands enseignants se convaincre sur la base de ces chiffres qu'ils étaient au mieux médiocres, m'expliqua Clifford, ce qui les a poussés à délaisser les excellents cours qu'ils avaient coutume de donner, et à mettre davantage l'accent sur la préparation aux tests. Pour un jeune enseignant, une mauvaise évaluation de sa valeur ajoutée est vécue comme une punition, tandis qu'un bon score peut lui inspirer le faux sentiment d'une réussite qui n'est pas méritée. »

À l'instar de tant d'ADM, la modélisation de la valeur ajoutée découle de bonnes intentions. L'administration Obama s'est aperçue très tôt que les réformes instaurées en 2001 par la loi « No Child Left Behind » (« Aucun enfant laissé pour compte »), imposant des tests normalisés à très fort enjeu, tendaient à sanctionner les districts scolaires pauvres et désavantagés. Ceux qui parvenaient à démontrer l'efficacité de leurs enseignants se virent donc accorder des dérogations, garantissant aux écoles concernées de ne pas être punies même si leurs élèves étaient à la traîne<sup>2</sup>.

L'utilisation de modèles basés sur la valeur ajoutée résulte dans une grande mesure de ce changement réglementaire. Mais fin 2015, l'engouement pour l'évaluation des enseignants a connu peut-être un tournant encore plus décisif. Tout d'abord, le Congrès et la Maison-Blanche sont convenus d'abroger « No Child Left Behind » et d'y substituer une loi offrant aux États plus de latitude pour développer leurs propres approches, destinées à remettre sur pied les districts peu performants. Elle leur permet également de prendre en compte un éventail de critères plus large, parmi lesquels l'investissement des élèves et des enseignants, le climat scolaire, la sécurité, et l'accès à des cours avancés, un dispositif permettant aux lycéens de valider à l'avance des modules universitaires. Les responsables éducatifs peuvent en d'autres termes étudier de manière individuelle ce qui se passe dans chaque établissement, et accorder moins d'attention à des ADM telle que le modèle de la valeur ajoutée – ou mieux encore, s'en débarrasser complètement.

Vers la même époque, le groupe de travail sur l'éducation mis en place par le gouverneur de l'État de New York, Andrew Cuomo, a réclamé un moratoire de quatre ans sur l'utilisation d'examens pour évaluer les enseignants. Ce changement, quoique bienvenu, ne signale pourtant pas un rejet clair et net des ADM correspondantes, et encore moins une

reconnaissance de leur caractère injuste. L'élan est en réalité venu des parents, lesquels s'étaient plaints que ce régime de tests épuisait leurs enfants et prenait trop de temps sur l'année scolaire. Au printemps 2015, un mouvement de boycott avait conduit 20 % des élèves du primaire au collège à ne pas passer les tests, et il ne cessait de s'amplifier. En cédant à la pression des parents, l'administration Cuomo porta un coup à la modélisation de la valeur ajoutée. Sans une évaluation exhaustive de l'ensemble des élèves, l'État manquerait en effet de données pour l'alimenter.

Clifford s'était réjoui de cette nouvelle mais restait prudent. « Le boycott a forcé la main de Cuomo, me confia-t-il par e-mail. Il a eu peur de perdre l'appui des électeurs aisés situés dans les meilleurs districts scolaires, ceux-là mêmes qui étaient ses plus fervents soutiens. Pour prendre les devants, il a introduit ce moratoire sur l'usage des scores d'évaluation. » Clifford craint cependant que les tests ne soient rétablis.

C'est bien possible. Et puisque la modélisation de la valeur ajoutée est un outil qui a désormais fait ses preuves à l'encontre des syndicats enseignants, je ne m'attends pas à ce qu'elle disparaisse de sitôt. Elle est aujourd'hui bien ancrée, utilisée ou en cours de développement, sous une forme ou sous une autre, dans quarante États ainsi que dans la capitale fédérale (le district de Columbia). Raison de plus pour faire connaître, au même titre que les autres, ces ADM-là. Une fois qu'ils sauront les identifier et qu'ils auront compris leurs travers statistiques, les gens exigeront des évaluations plus juste vis-à-vis à la fois des élèves et des enseignants. Si les tests ont en revanche comme objectif de trouver un bouc émissaire et d'intimider les travailleurs, une ADM débitant des scores dénués de sens obtient alors comme on l'a vu la note maximale.

---

1. Aux États Unis, le salaire minimum est fixé par le gouvernement puis varie à la hausse en fonction des États. En septembre 2018, il est de 7,25 dollars de l'heure, soit environ 6,23 euros.

2. Parmi les sanctions prévues figure la possibilité pour les élèves des écoles faibles de fréquenter un autre établissement d'un meilleur niveau. Dans les cas extrêmes, la loi exige la fermeture des écoles défaillantes et leur remplacement par des écoles sous contrat (privées mais à financement public) (NdA).

# DOMMAGES

## Obtenir un crédit COLLATÉRAUX

1001  books

**A**utrefois, les banquiers locaux étaient d'éminents personnages au sein de leur ville. C'étaient eux qui contrôlaient l'argent. Si vous vouliez acheter une nouvelle voiture ou contracter un emprunt, vous mettiez vos habits du dimanche et rendiez visite à l'un d'entre eux. En tant que membre de votre communauté, ce banquier connaissait probablement certains détails de votre vie. Il savait si vous alliez à l'église ou pas. Il connaissait toute l'histoire des démêlés judiciaires de votre frère aîné. Il savait ce que votre patron (et son partenaire de golf) disaient de vous sur le plan professionnel. Il savait naturellement de quelle origine vous étiez, et il jetait aussi un coup d'œil aux chiffres inscrits sur votre demande de prêt.

Les quatre premiers facteurs influençaient souvent, consciemment ou non, le jugement du banquier. Et il était probable qu'il fasse davantage confiance aux gens issus de son propre milieu. C'est dans la nature humaine. Mais cela signifie que pour des millions d'Américains le système en place avant l'ère numérique s'avérait tout aussi cruel que certaines des ADM que j'ai pu décrire. Les exclus, notamment les femmes et les membres de certaines minorités, se faisaient régulièrement recalier. Il leur fallait réunir un dossier financier impressionnant – et se mettre en quête de banquiers à l'esprit ouvert.

Ce n'était tout simplement pas juste. Apparut alors un algorithme, et la situation s'améliora. Un mathématicien du nom d'Earl Isaac imagina avec un ami ingénieur, Bill Fair, un modèle qu'ils baptisèrent FICO, destiné à évaluer le risque de défaut de paiement d'un emprunteur. Ce score FICO découlait d'une formule qui tenait uniquement compte des finances de

l'individu – principalement de son niveau d'endettement et du bon acquittement de ses factures. FICO n'opérait aucune discrimination selon la couleur de peau. Et il se révéla d'un apport fantastique pour l'industrie bancaire, car il prédisait le risque avec une précision bien plus grande, tout en ouvrant la porte à des millions de nouveaux clients. Bien entendu, les scores FICO existent toujours. Ils sont utilisés par les agences d'évaluation de crédit telles que Experian, Transunion ou Equifax, qui alimentent chacune le modèle via différentes sources pour établir leurs propres scores. Ceux-ci possèdent beaucoup d'aspects louables qui les distinguent des ADM. Ils intègrent tout d'abord une boucle de rétroaction parfaitement claire. Les sociétés de crédit peuvent détecter quels sont les emprunteurs qui font défaut, et rapprocher ces chiffres de leurs scores. Si les emprunteurs dotés d'un score élevé semblent faire plus souvent défaut que le modèle ne le prévoit, FICO et les agences d'évaluation peuvent ajuster les modèles pour en améliorer la précision. Ce qui constitue un usage judicieux des données statistiques.

Les scores de crédit sont aussi relativement transparents. Le site web de FICO fournit par exemple des recommandations simples sur la façon d'améliorer son score (réduire sa dette, régler ses factures à temps, et cesser de commander de nouvelles cartes de crédit). Il faut également noter que les organismes évaluateurs sont soumis à réglementation. Si vous vous posez des questions sur votre score, vous êtes en droit de réclamer votre rapport de crédit, qui renferme toutes les informations entrant dans son calcul, dont l'historique de vos remboursements d'emprunts et de vos factures d'eau et d'énergie, votre dette totale, et le pourcentage de crédit disponible que vous utilisez. Bien que le processus puisse se montrer d'une lenteur abominable, si vous découvrez des erreurs, vous pouvez les faire corriger.

Depuis l'époque pionnière de Fair et Isaac, l'utilisation des scores a bien sûr largement proliféré. Statisticiens et mathématiciens nous mesurent aujourd'hui de toutes les manières imaginables en agrégeant un vaste fatras de données, de nos codes postaux à nos derniers achats en passant par nos habitudes de navigation sur Internet. Beaucoup de leurs modèles pseudoscientifiques s'efforcent de prédire notre solvabilité, en attribuant à chacun ce qu'on appelle des e-scores. Ces chiffres, qui nous sont rarement connus, ouvrent les portes à certains et les claquent au nez des autres. À la différence des scores FICO, auxquels ils ressemblent, les e-scores sont

arbitraires, obscurs, non réglementés, et souvent injustes – en bref, ce sont des ADM.

La société Neustar, basée en Virginie, en offre un parfait exemple. Neustar fournit des services de ciblage marketing aux entreprises, dont un en particulier qui les aide à gérer le trafic dans les centres d'appel. En un éclair, cette technologie parcourt les données disponibles sur chaque appelant et opère un classement hiérarchique. Ceux qui se classent au sommet sont considérés comme des prospects plus rentables et sont rapidement orientés vers un opérateur humain. Ceux en bas de l'échelle patientent bien plus longtemps ou sont renvoyés sur un centre de délestage externalisé, où leur appel sera traité en grande partie par des machines.

Les émetteurs de cartes de crédit<sup>1</sup> comme Capital One réalisent le même genre de calculs ultrarapides dès qu'un visiteur se rend sur leur site web. Ils ont souvent accès aux données de navigation et d'achat en ligne, qui leur procurent quantité d'indications sur le client potentiel. Il y a ainsi fort à parier que l'internaute qui recherche une Jaguar neuve soit plus riche que celui qui vérifie l'état d'une Ford Taurus de 2003 sur Carfax.com. La plupart des systèmes d'évaluation récupèrent également la localisation de l'ordinateur du visiteur. Par croisement avec des données immobilières, ils peuvent en tirer des conclusions quant à son niveau de vie. Une personne connectée depuis le quartier huppé de Balboa Terrace à San Francisco fera un bien meilleur prospect que celle située de l'autre côté de la baie à East Oakland, un quartier nettement moins bien coté.

L'existence de ces e-scores ne devrait surprendre personne. On a vu des modèles qui se nourrissent de données similaires pour nous proposer des prêts dits « rapaces » aux conditions abusives ou estimer la probabilité que nous avons de dérober un véhicule. Pour le meilleur ou pour le pire, ils nous ont menés à des études (ou en prison) et orientés vers un emploi, puis ils ont optimisé notre place au sein de l'entreprise. Le moment venu d'acquérir peut-être un logement ou une voiture, rien de plus normal que des modèles financiers exploitent le même gisement de données pour nous jauger.

Mais songez donc au cercle vicieux engendré par les e-scores. Il y a de très fortes chances que le système d'évaluation attribue un faible score à l'emprunteur qui habite le secteur mal famé de East Oakland. Dans cette partie-là de la ville, beaucoup de gens font défaut sur leurs remboursements. L'offre de carte de crédit qui s'affiche sur son écran ciblera par conséquent une tranche de population plus risquée. D'où un crédit disponible plus

faible et des taux d'intérêt plus élevés pour ceux qui ont déjà du mal à joindre les deux bouts.

La plupart des publicités prédatrices que nous avons évoquées, notamment pour des prêts sur salaire ou des universités privées à but lucratif, sont générées par le biais de ce genre d'e-scores. Ils se substituent aux scores de crédit : puisque la loi interdit aux entreprises de les utiliser à des fins marketing, elles s'accommodent de ces ersatz expéditifs.

Cette interdiction répond à une certaine logique. Notre historique de crédit renferme après tout des données extrêmement personnelles, et il semble donc légitime de conserver la maîtrise de sa divulgation. Mais la conséquence en est que les entreprises finissent par puiser dans des viviers de données qui ne sont guère réglementés, comme les parcours de navigation en ligne ou les informations de géolocalisation, pour aboutir à la création d'un marché parallèle de la donnée. Elles s'octroient ce faisant la possibilité d'échapper en grande partie à la surveillance des autorités. Elles mesurent alors leur réussite à l'aune des gains réalisés en termes d'efficience, de flux de trésorerie et de profits. À de rares exceptions près, des concepts comme ceux de justice et de transparence n'ont pas leur place dans ces algorithmes.

Comparons cela un instant au banquier des années 1950. Consciemment ou non, ce dernier soupesait diverses données qui n'avaient que peu ou pas de rapport avec la capacité du potentiel emprunteur à assumer un prêt. Il notait les origines ethniques du client assis de l'autre côté du bureau, et en tirait des conclusions. Le casier judiciaire de son père était susceptible de lui porter préjudice, tandis que son assiduité à l'église pouvait jouer en sa faveur.

Tous ces éléments étaient des données indirectes. Pour évaluer la responsabilité financière du client, le banquier aurait pu procéder à une étude impartiale des chiffres (comme le faisaient sans nul doute certains confrères exemplaires). Au lieu de quoi il traçait des corrélations avec l'ethnie, la religion et les liens familiaux. Il s'abstenait de la sorte de sonder l'emprunteur en tant qu'individu et l'inscrivait au contraire comme membre d'un groupe – ce que les statisticiens appelleraient aujourd'hui un « *bucket* » (littéralement, un « seau »). Les « gens comme lui », décrétait-il, étaient ou n'étaient pas dignes de confiance.

La grande avancée de Fair et Isaac fut d'abandonner les données de remplacement au profit de données financières pertinentes, telles que

l'historique d'acquittement des factures. Leur analyse se concentrait sur l'individu dont il était question – et non sur d'autres personnes aux caractéristiques similaires. Les e-scores, à l'inverse, nous renvoient dans le passé. Ils analysent l'individu par le biais d'un véritable maelstrom de données indirectes. En l'espace de quelques millisecondes, ils effectuent des milliers de calculs pour le comparer à « des gens comme lui ». Et si un nombre suffisant de ses « semblables » se révèlent être de mauvais payeurs, ou pire des criminels, cet individu-là se verra traité comme tel.

On me demande de temps en temps comment enseigner l'éthique à une classe d'experts en données. Je commence habituellement par une discussion sur la manière de bâtir un modèle d'e-score, et leur demande s'il paraît judicieux d'utiliser l'origine ethnique comme donnée d'entrée. Ils répondent invariablement que ce serait injuste et sans doute illégal. La question suivante porte sur l'utilisation ou non du code postal. À première vue, cela semble de bonne guerre. Mais il ne faut pas longtemps pour que les étudiants s'aperçoivent qu'ils codifient ainsi dans leur modèle les injustices du passé. Quand ils prennent en compte un attribut tel que le code postal, ils décident que l'histoire des comportements humains dans cette zone foncière devrait déterminer, au moins en partie, le type de prêt qu'un de ses habitants peut obtenir.

En d'autres termes, les modélisateurs d'e-scores doivent se contenter de répondre à la question « Comment les gens comme vous se sont-ils conduits par le passé ? », alors qu'il leur faudrait demander dans l'idéal « Comment vous êtes-vous comporté, *vous*, en particulier ? ».

La différence entre ces deux questions est colossale. Imaginez qu'un immigré aux moyens modestes, extrêmement motivé et responsable, cherche à lancer son affaire et soit obligé de passer par un tel système pour commencer à investir. Qui prendrait le risque de financer quelqu'un comme lui ? Sans doute pas un modèle reposant sur des données démographiques et comportementales de cet ordre.

Il convient de remarquer que, dans l'univers statistique qu'elles habitent, les données indirectes fonctionnent souvent. Dans une majorité de cas, le proverbe « qui se ressemble s'assemble » se vérifie bel et bien. Les riches peuvent s'acheter des croisières et des BMW. Les pauvres, eux, auront souvent besoin d'un prêt sur salaire. Et puisque ces modèles statistiques semblent fonctionner la plupart du temps, l'efficacité augmente et les profits s'envolent. Les investisseurs doublent leur mise, convaincus par ces

systèmes scientifiques capables apparemment de ranger des milliers de personnes dans les bons *buckets*. C'est le triomphe du Big Data.

Mais qu'advient-il d'un individu mal appréhendé qui se retrouve placé dans la mauvaise catégorie ? Cela arrive, et il n'y a aucune remontée d'information permettant de corriger le système. Un moteur chargé de traiter de grandes masses de statistiques n'a aucun moyen de savoir qu'il a orienté un précieux client potentiel vers les méandres infernaux d'un centre d'appel. Pire, dans l'univers dérégulé des e-scores, les perdants n'ont guère de recours pour se plaindre, et encore moins pour rectifier l'erreur. Dans le royaume des ADM, ils constituent des dommages collatéraux. Comme cet obscur système opère tout entier depuis de lointaines fermes de serveurs, ils découvrent en outre rarement de quoi il retourne. La plupart en concluent probablement, avec raison, que la vie est tout simplement injuste.

\*\*\*

Au sein du monde que j'ai décrit jusqu'à présent, les e-scores alimentés par des millions de données indirectes œuvrent dans l'ombre, tandis que nos rapports de crédit, remplis de données pertinentes et appropriées, sont régis par la loi. Mais ce n'est hélas pas si simple. Trop souvent, les rapports de crédit servent eux aussi de données indirectes.

Cela n'a rien d'étonnant : dans notre société, beaucoup d'institutions – des grandes entreprises jusqu'aux administrations publiques – recherchent des personnes fiables et dignes de confiance. Dans le chapitre sur l'obtention d'un emploi, on les a vues trier des CV et écarter les candidats dont les tests psychologiques révélaient des caractéristiques personnelles indésirables. Une autre approche trop répandue consiste à prendre en compte le score de crédit du postulant. Si les gens paient leurs factures à temps et évitent de s'endetter, ne serait-ce pas là, s'interrogent les employeurs, une marque de sérieux et de fiabilité ? Ce n'est pas *exactement* la même chose, ils le savent. Mais n'y aurait-il pas entre ces deux aspects un recoupement significatif ?

C'est ainsi que l'emprise des rapports de crédit s'est étendue bien au-delà de son périmètre originel. Par excès de facilité, la solvabilité a remplacé d'autres mérites. Inversement, une mauvaise notation de crédit semble à présent signaler un ensemble de fautes et de défauts qui n'ont rien à voir avec le fait de régler ses factures. Comme nous le verrons par la suite,



toutes sortes d'entreprises tirent de ces rapports leur propre version des scores de crédit et les utilisent comme données indirectes. Cette pratique est aussi nuisible que généralisée.

Dans le cadre d'applications spécifiques, l'usage de telles données pourrait paraître inoffensif. Certains services de rencontres en ligne appariant par exemple les individus en fonction de leurs scores de crédit. L'un d'entre eux, CreditScoreDating, proclame d'ailleurs : « Les bons scores de crédit sont sexy. » On peut certes se demander s'il est sage de lier le comportement financier au sentiment amoureux. En tout cas, les clients de CreditScoreDating savent au moins où ils mettent les pieds et pourquoi. À eux de décider.

Mais si vous cherchez un emploi, il y a de fortes chances pour qu'un remboursement de carte de crédit manqué ou des pénalités de retard sur votre prêt étudiant puissent vous desservir. Selon une enquête de la Society for Human Resource Management, association de professionnels en gestion des ressources humaines, près de la moitié des employeurs américains filtrent les recrues potentielles en examinant leur rapport de crédit. Certains contrôlent également le niveau de solvabilité de leurs employés actuels, surtout s'ils sont en lice pour une promotion.

Avant de mener ces vérifications, les entreprises doivent d'abord en demander l'autorisation. Mais ce n'est en général qu'une formalité ; chez beaucoup d'employeurs, les personnes qui refusent de divulguer leurs données de crédit ne seront même pas envisagées pour un poste. Et si leurs antécédents s'avèrent mauvais, d'autres candidats risquent fort de leur passer devant. En 2012, une enquête sur l'endettement par carte de crédit dans les familles à revenus faibles et moyens l'a très clairement montré. Un sujet sur dix s'est entendu dire par des employeurs que les irrégularités de son historique de crédit avaient ruiné ses chances – et qui sait combien ont été éconduits pour ce motif sans en être jamais informés ? Car si la loi stipule que les employeurs doivent alerter les demandeurs d'emploi quand des problèmes de solvabilité les mettent hors course, on imagine sans peine qu'un certain nombre se contentent de leur annoncer qu'ils ne conviennent pas, ou que d'autres candidats étaient plus qualifiés.

L'utilisation des scores de crédit en matière d'embauche et de promotion crée un cercle vicieux dangereux en termes de pauvreté. En effet, si votre rapport de crédit vous empêche de décrocher un poste, ledit rapport va alors probablement se dégrader, rendant plus difficile encore l'obtention d'un

travail. Ce n'est pas sans rappeler la difficulté que rencontrent les jeunes lorsqu'ils cherchent leur premier emploi – et se voient disqualifiés par manque d'expérience. Ou la situation compliquée des chômeurs de longue durée, qui trouvent peu d'employeurs disposés à les recruter car ils sont restés trop longtemps inactifs. C'est une boucle de rétroaction qui entraîne les malchanceux pris au piège dans une spirale d'échec.

Naturellement, les employeurs n'éprouvent que peu d'indulgence vis-à-vis de cet argument. Une bonne qualité de crédit, soulignent-ils, constitue l'un des attributs d'une personne responsable – le type de personne qu'ils souhaitent recruter. Mais concevoir l'endettement comme un problème moral n'en est pas moins une erreur. Quantité d'individus vaillants et dignes de confiance perdent chaque jour leur emploi à la suite de faillites d'entreprises, de réductions de budgets ou de délocalisations. Les chiffres augmentent durant les périodes de récession. Et beaucoup de ces nouveaux chômeurs se retrouvent sans assurance maladie. À ce stade, il leur suffit d'avoir un accident ou de tomber malades pour ne pas acquitter une mensualité d'emprunt. Même après la promulgation du Affordable Care Act (loi sur les soins abordables, surnommée « Obamacare ») qui a réduit le nombre de personnes sans assurance, les dépenses médicales demeurent aux États-Unis la première cause de faillite personnelle.

Les gens qui ont épargné peuvent bien entendu conserver leur crédit intact pendant les moments difficiles. Ceux qui n'ont que leur salaire pour vivre d'un mois sur l'autre sont bien plus vulnérables. Un excellent score de crédit ne fait donc pas simplement office de donnée indirecte pour évaluer le degré de responsabilité et le caractère avisé des décisions prises. C'est également un indicateur de richesse. Et la richesse s'avère fortement corrélée avec les origines ethniques.

Jugez plutôt. En 2015, les ménages blancs détenaient en moyenne dix fois plus d'argent et de biens que les ménages noirs et hispaniques. Et tandis que 15 % seulement des Blancs possédaient un patrimoine nul ou négatif, plus d'un tiers des foyers noirs et hispaniques n'avaient aucun matelas de sécurité. Cet écart de richesse augmente avec l'âge. Arrivés à la soixantaine, les Blancs sont onze fois plus riches que les Afro-Américains. Compte tenu de ces chiffres, il n'est pas difficile d'affirmer que l'engrenage de la pauvreté, résultat du contrôle de la qualité de crédit par les employeurs, affecte la société de manière inégale et selon des lignes de fracture raciales. À l'heure où j'écris ceci, dix États ont adopté une loi pour

proscrire l'utilisation des scores de crédit dans le cadre d'un recrutement. Lorsqu'elle les a interdites, l'administration de la ville de New York a déclaré que les vérifications de solvabilité « impactaient de façon disproportionnée les candidats de couleur et à faibles revenus ». Cette pratique reste cependant légale dans les quarante autres États.

Il ne s'agit pas de dire que les services du personnel élaborent intentionnellement partout en Amérique un piège à pauvres, et encore moins un piège raciste. Ces services pensent sans doute que les rapports de crédit renferment des faits pertinents qui les aident à trancher des choix importants. « Plus il y a de données, mieux c'est » : voilà bien après tout le principe directeur qui régit l'ère de l'information. Pour autant, au nom de l'équité, certaines de ces données ne devraient pas être exploitées.

\*\*\*

Imaginez un instant : récemment diplômé de la faculté de droit de l'université de Stanford, vous passez un entretien d'embauche dans un prestigieux cabinet juridique de San Francisco. L'associé principal consulte son fichier informatique et éclate de rire : « Je lis ici que vous avez été arrêté pour possession d'un laboratoire de méthamphétamine à Rhode Island ! » Il secoue la tête. Votre nom de famille est très courant, et les ordinateurs commettent assurément de stupides erreurs. L'entretien se poursuit.

Dans les strates supérieures de l'économie, ce sont généralement des êtres humains qui prennent les décisions sensibles, se servant des ordinateurs comme d'un outil pratique. Mais pour la majorité de la population, surtout aux échelons les plus bas, l'essentiel du travail, comme nous l'avons vu, est automatisé. Si des erreurs apparaissent dans un dossier – et c'est souvent le cas –, les algorithmes même les mieux conçus prendront la mauvaise décision. Comme le disent de longue date les professionnels du traitement de données : « *garbage in, garbage out* » – si ce qui entre est faux, ce qui sort sera faux.

L'individu qui fait les frais de ce processus automatisé peut en endurer les conséquences pendant des années. Les listes d'interdiction de vol sur des compagnies aériennes pour terrorisme, générées par ordinateur, sont ainsi connues pour regorger d'erreurs. Une personne innocente dont le nom ressemble à celui d'un terroriste présumé affrontera une épreuve infernale

chaque fois qu'elle devra monter dans un avion (les passagers fortunés, en revanche, ont souvent la possibilité d'acquiescer contre paiement le statut de « voyageur de confiance », qui leur permet de franchir tranquillement les dispositifs de sécurité. Ils versent de fait de l'argent pour se prémunir d'une ADM).

Les anomalies de ce genre se manifestent partout. En 2013, la Federal Trade Commission a indiqué que 5 % des consommateurs – soit environ 10 millions de personnes – avaient pâti d'une erreur dans leurs rapports de crédit qui avait engendré des frais d'emprunt plus élevés. C'est ennuyeux, mais les rapports de crédit font au moins partie du versant réglementé de l'économie des données. Les consommateurs peuvent (et devraient) exiger d'en prendre connaissance une fois par an afin de corriger ces bêtises potentiellement coûteuses<sup>2</sup>.

Le versant non réglementé présente toutefois un danger encore plus grand. Une multitude d'entreprises, qu'il s'agisse de géants comme Acxiom Corp. ou de petites firmes éphémères, achètent des informations auprès des commerçants, des publicitaires, des éditeurs d'applications mobiles, des organisateurs de loteries et des opérateurs de réseaux sociaux afin de réunir sur chaque consommateur du pays un vaste ensemble de faits. Elles pourront noter par exemple qu'un consommateur souffre de diabète, qu'il habite avec une personne qui fume, qu'il conduit une voiture de loisirs du type SUV, ou qu'il possède deux chiens de race Colley (lesquels resteront peut-être en vie dans son dossier longtemps après leur disparition). Ces entreprises recueillent également toutes sortes de données administratives en libre accès, notamment les registres de vote, les relevés d'arrestations et les ventes immobilières. Tous ces renseignements sont incorporés dans le profil du consommateur, qu'elles revendent ensuite.

Certains courtiers en données sont sans nul doute plus fiables que d'autres. Mais toute société cherchant à profiler des centaines de millions de personnes à partir de milliers de sources différentes collectera à coup sûr beaucoup de faits erronés. Prenons le cas de Helen Stokes, une habitante de Philadelphie. Helen souhaitait entrer dans une maison de retraite des environs, mais ne cessait d'être recalée car son dossier faisait état d'arrestations. Elle avait certes été interpellée à deux reprises lors d'altercations avec son ancien mari. Mais aucune condamnation ne s'était ensuivie, et elle avait obtenu que les faits soient effacés des bases de données de l'administration. Ces relevés d'arrestations restaient néanmoins

consignés dans les fichiers assemblés par une entreprise du nom de RealPage, Inc., qui vérifiait les antécédents des locataires.

Pour RealPage et d'autres entreprises de ce genre, la création et la vente de rapports sont une source de revenus. Les gens comme Helen Stokes ne sont pas des clients, mais des produits. Répondre à leurs réclamations demande du temps et occasionne des frais. Stokes aura beau affirmer en effet que les arrestations ont été effacées, vérifier ses dires se révèle chronophage et dispendieux. Un opérateur humain fort coûteux devra peut-être se rendre quelques minutes sur Internet, voire – grands dieux ! – passer un ou deux coups de téléphone. Que Stokes ait été obligée de porter plainte pour faire nettoyer son dossier n'a donc rien de très étonnant. Et même après que RealPage lui a répondu, combien d'autres courtiers sont encore susceptibles de vendre des fichiers renfermant les mêmes informations, mensongères et nuisibles ? Nul ne le sait.

Bien entendu, certains permettent aux consommateurs d'accéder à leurs données. Mais ces rapports sont fortement expurgés. Ils mentionnent les faits mais n'indiquent pas toujours les conclusions que les algorithmes en ont tirées. L'individu qui prendra la peine de consulter son fichier chez l'un de ces nombreux courtiers verra par exemple apparaître son emprunt immobilier, une facture de téléphonie mobile, et une note de 459 dollars pour la réparation de la porte du garage. Mais il ne verra pas qu'il est rangé dans le *bucket* des « ruraux ayant du mal à joindre les deux bouts » ou des « préretraités sans épargne ». Heureusement pour les courtiers, peu de gens ont l'occasion de découvrir ces détails. Si c'était le cas – et la Federal Trade Commission pousse à ce qu'ils rendent davantage de comptes –, ils se trouveraient probablement assaillis de millions de réclamations. Leur business model aurait de fortes chances d'en être bouleversé. Pour l'heure, ce n'est qu'à la faveur d'une divulgation involontaire, souvent par pur hasard, que les consommateurs apprennent l'existence de fichiers erronés.

Il y a quelques années, par exemple, une résidente de l'Arkansas du nom de Catherine Taylor n'a pas obtenu un poste à l'antenne locale de la Croix-Rouge. Cela peut arriver. Mais la lettre de refus qui lui est parvenue était accompagnée d'une précieuse information. Son rapport d'antécédents contenait une accusation pénale pour intention de produire et de vendre de la méthamphétamine. Ce n'était pas le genre de candidat que la Croix-Rouge cherchait à recruter.

Taylor se renseigna et apprit que les charges concernaient une autre Catherine Taylor, née le même jour. Elle découvrit plus tard qu'au moins dix autres entreprises l'accablaient de rapports inexacts – dont un relatif à sa demande d'allocation logement, qui avait été rejetée. Était-ce la conséquence d'une erreur d'identité ?

Dans le cadre d'un processus complètement automatisé, ç'aurait été le cas. Mais un être humain était intervenu. Lors du dépôt de leur demande, Taylor et son époux avaient rencontré une employée de l'office du logement afin de vérifier leurs antécédents. L'employée en question, Wanda Taylor – aucun lien de parenté –, se servait des informations fournies par le courtier Tenant Tracker. Celles-ci étaient remplies d'erreurs et mélangeaient plusieurs identités. Elles désignaient par exemple Taylor comme possible pseudonyme de Chantel Taylor, une criminelle née elle aussi le même jour. Elles la rattachaient également à l'autre Catherine Taylor, condamnée dans l'Illinois pour vol, falsification de documents et possession d'une substance prohibée.

Son dossier n'était en résumé qu'un amalgame nocif. Mais Wanda Taylor, qui avait l'expérience de ce genre de choses, entreprit de tout passer en revue. Elle raya immédiatement le supposé pseudonyme, Chantel, qui lui paraissait improbable. Elle lut dans le fichier que la voleuse de l'Illinois portait le nom de « Troy » tatoué sur la cheville. Après avoir examiné celle de Catherine, elle barra également le nom de cette délinquante. Au terme du rendez-vous, un être humain consciencieux avait dissipé la confusion générée par des logiciels collecteurs de données sur Internet. L'office du logement savait maintenant à quelle Catherine Taylor il avait affaire.

Reste la question suivante : combien y a-t-il de Wanda Taylor occupées à éliminer les fausses identités et toutes les autres erreurs contenues dans nos données ? Pas assez, tant s'en faut. Dans l'économie des données, les humains font figure d'intrus et de vestiges. Les systèmes sont conçus pour fonctionner autant que possible sur un mode automatique. C'est la manière efficace de procéder ; c'est là que résident les profits. Les erreurs, comme dans tout programme statistique, sont inévitables, mais le moyen le plus rapide de les réduire consiste à affiner les algorithmes qui pilotent les machines. Les humains présents sur le terrain ne font que gripper le mécanisme.

Cette tendance vers l'automatisation progresse à grands pas du fait que les ordinateurs décryptent de mieux en mieux ce que nous écrivons, se

montrant capables dans certains cas de traiter en une seconde des milliers de documents. Mais ils comprennent encore de travers toutes sortes de choses. Aussi brillant fût-il, le superordinateur Watson d'IBM entraîné pour concourir au jeu télévisé *Jeopardy !* s'est trouvé désarçonné environ 10 % du temps sur des aspects liés à la langue ou au contexte. On l'a entendu répondre que le régime alimentaire d'un papillon était « casher », et il a confondu *Oliver Twist*, le personnage de Charles Dickens, avec le groupe techno-pop des années 1980 Pet Shop Boys.

De telles méprises ne manqueront pas de s'accumuler dans nos profils de consommateur, perturbant et égarant les algorithmes qui gèrent une part de plus en plus grande de nos vies. Ces erreurs, résultat de la collecte automatisée de données, infectent les modèles prédictifs et alimentent les ADM. Et cette collecte n'est appelée qu'à croître. Les ordinateurs s'emploient déjà à étendre leur champ d'action au-delà de l'écrit. Ils recueillent paroles et images, dont l'exploitation leur permet d'obtenir davantage d'informations sur tout ce qui compose l'univers environnant – y compris sur nous-mêmes. Ces nouvelles technologies fouilleront de nouveaux gisements pour étoffer nos profils, tout en augmentant le risque d'erreurs.

Google a récemment analysé les photos d'un trio de jeunes Afro-Américains à la mine réjouie, et son service de balisage automatique les a identifiés comme des « gorilles ». L'entreprise s'est répandue en excuses, mais dans les systèmes comme celui de Google, les erreurs sont inévitables. C'est sans aucun doute une défaillance du système d'apprentissage de la machine (et probablement pas l'acte d'un raciste évoluant librement au sein du Googleplex), qui avait conduit l'ordinateur à confondre *Homo sapiens* avec notre proche cousin. Le logiciel avait parcouru seul des milliards d'images de primates et avait établi ses propres distinctions. Il s'était intéressé à chaque détail, des différentes nuances de couleur jusqu'à l'écartement des yeux en passant par la forme des oreilles. Pour autant, il n'avait manifestement pas été assez testé avant sa mise en ligne.

Les erreurs de ce genre constituent des occasions d'apprentissage – à condition cependant que le système reçoive un feed-back, une remontée d'informations. Ce fut le cas ici. Mais l'injustice persiste. Quand des systèmes automatiques explorent nos données en vue de nous attribuer un e-score, ils projettent naturellement le passé sur l'avenir. Comme nous l'avons vu avec les modèles de récidivisme des sentences pénales et les

algorithmes des sociétés d'emprunt aux pratiques prédatrices, les pauvres sont censés le rester à jamais et se voient traités en conséquence – on leur refuse des opportunités, ils sont plus souvent jetés en prison, et se font escroquer en matière de prêts et de services. C'est un processus inexorable, souvent caché, sans recours, et injuste.

Impossible toutefois d'attendre des systèmes automatiques qu'ils résolvent le problème. En dépit de leur stupéfiante puissance, les machines ne savent pas encore effectuer les ajustements nécessaires pour garantir l'équité, du moins pas de leur propre chef. Parcourir les données de façon à valoriser ce qui est juste leur reste totalement étranger, et cela s'avère pour elles extrêmement compliqué. Les êtres humains sont seuls en mesure d'imposer cette contrainte.

Il y a ici un paradoxe. Si l'on en revient une dernière fois au banquier des années 1950, on constate que son esprit était habité de distorsions typiquement humaines – désirs, préjugés, méfiance envers les personnes venues de l'extérieur. Pour remplir sa tâche de façon plus juste et plus efficiente, il a confié le travail – comme l'ensemble de l'industrie bancaire – à un algorithme.

Soixante ans plus tard, le monde est dominé par des systèmes automatiques qui se nourrissent à la chaîne de dossiers truffés de défauts. Le besoin de contexte, de bon sens et d'équité, éléments que seuls les humains peuvent apporter, se fait urgemment sentir. Or, si l'on abandonne cette question au profit du marché – qui valorise l'efficacité, la croissance et les flux de trésorerie (tout en tolérant un certain niveau d'erreur) –, les humains recevront l'ordre de ne pas s'immiscer dans la mécanique.

\*\*\*

Il s'agira là d'un important défi, car alors même que deviennent évidents les problèmes associés à nos vieux modèles d'évaluation de la qualité de crédit, de nouveaux et puissants acteurs font aujourd'hui irruption. Facebook, par exemple, a breveté un score de crédit d'un nouveau genre, basé sur nos réseaux sociaux. L'objectif paraît de prime abord légitime. Imaginez qu'un jeune diplômé parte cinq ans en mission religieuse pour apporter l'eau potable dans des villages démunis en Afrique. De retour chez lui, il ne possède aucun score de crédit et a du mal à obtenir un prêt. Mais ses camarades de classe sur Facebook sont désormais banquiers d'affaires,



titulaires d'un doctorat, ou concepteurs de logiciels. Une analyse fondée sur les affinités indiquera alors qu'on peut miser sur lui. Malheureusement, cette même analyse desservira sans doute une vaillante femme de ménage habitant la ville pauvre d'East St. Louis, qui risque de compter parmi ses amis beaucoup de chômeurs et quelques personnes incarcérés.

Pendant ce temps, le secteur bancaire officiel ratisse frénétiquement les données personnelles pour essayer de stimuler l'activité. Mais les banques agréées sont assujetties à une réglementation nationale et à des exigences de communication, ce qui signifie que le profilage des clients leur fait courir un risque en termes de réputation et de légalité. La société de cartes de paiement American Express l'a appris à ses dépens en 2009, juste au moment où la grande crise économique s'enracinait. Cherchant sans doute à réduire le profil de risque de son propre bilan, Amex a abaissé le plafond de dépenses de certains clients. Contrairement aux acteurs informels de l'économie du e-score, le géant de la carte de crédit était toutefois tenu de leur envoyer un courrier pour en expliquer la raison.

Et c'est là qu'Amex a porté un coup bas. Les détenteurs de carte qui fréquentaient certains commerces, écrivit l'entreprise, avaient davantage de chances de connaître des retards de paiement. Simple affaire de statistiques : il existait une corrélation flagrante entre les habitudes d'achat et les taux de défaut de paiement. Les clients mécontents d'Amex n'avaient plus qu'à deviner quel établissement avait altéré leur qualité de crédit. Était-ce la visite hebdomadaire à l'hypermarché populaire Walmart, ou peut-être l'entretien des freins de la voiture chez Grease Monkey qui les avait placés dans la catégorie des mauvais payeurs potentiels ?

Quelle qu'en soit la cause, voilà qui les laissait aborder une terrible récession avec un crédit encore plus faible. Pire, l'abaissement de leur plafond de dépenses apparaîtrait d'ici quelques jours à peine sur leurs rapports. Il y figurait d'ailleurs sans doute déjà avant même que les courriers n'arrivent. Cela aurait pour conséquence de diminuer leurs scores et d'accroître le coût de leurs emprunts. Beaucoup de ces détenteurs de carte, peut-on dire sans trop s'avancer, fréquentaient des « magasins associés à un mauvais taux de remboursement » pour la bonne raison qu'ils ne roulaient pas sur l'or. Et comme par hasard, un algorithme s'en était aperçu et les rendait encore plus pauvres.

Leur colère attira l'attention de la presse grand public, y compris du *New York Times*, et Amex s'empessa d'annoncer qu'elle ne ferait pas de

corrélation entre les commerces fréquentés et le niveau de risque (elle a affirmé plus tard que les termes utilisés dans son message avaient été mal choisis, et qu'elle s'était contentée d'analyser des schémas de consommation globaux, et non pas des enseignes spécifiques).

Ce fut un casse-tête et un embarras pour American Express. Si elle avait effectivement découvert une forte corrélation entre la fréquentation d'un magasin donné et le risque de crédit, il n'était certainement plus possible à présent de s'en servir. Comparée à la majorité des acteurs de l'économie d'Internet, Amex est encadrée, régulée, en un sens handicapée (sans qu'elle ait pour autant matière à se plaindre. Au fil des décennies, les lobbyistes au service des entreprises dominantes ont élaboré nombre des réglementations en vigueur dans l'optique de défendre les puissances établies – et de bloquer l'arrivée d'importuns compétiteurs).

Faut-il donc s'étonner que les nouveaux arrivants de l'industrie financière choisissent une voie plus libre, exempte de régulation ? La liberté d'expérimenter conditionne après tout la capacité d'innovation. Or, avec des pétaoctets de données comportementales à portée de main et une absence quasi totale de contrôle, les opportunités pour la création de business models inédits se révèlent immenses.

Diverses entreprises cherchent par exemple à prendre la place des prêteurs sur salaire. Ces banques de dernier recours s'adressent aux travailleurs pauvres ; elles les maintiennent à flot d'une paie à la suivante en leur appliquant des taux d'intérêt exorbitants. Au bout de vingt-deux semaines, un prêt de 500 dollars peut finir par en coûter 1 500. Si un nouvel acteur efficace découvrait de nouveaux moyens d'évaluer le risque, puis d'extraire de cette masse de gens désespérés les candidats solvables, il pourrait alors leur facturer des intérêts légèrement inférieurs tout en continuant d'engranger d'énormes profits.

Telle était l'idée de Douglas Merrill. Ancien directeur de l'exploitation chez Google, Merrill estimait possible d'exploiter le Big Data pour mesurer le risque et proposer de la sorte des prêts sur salaire à taux réduit. En 2009, il a fondé la start-up ZestFinance. Sur la page web de l'entreprise, Merrill proclame que « toute donnée peut servir à évaluer la qualité de crédit ». Autrement dit, tout est permis.

ZestFinance achète des données qui indiquent si les candidats ont bien réglé leurs factures de téléphonie mobile à temps, ainsi que nombre d'autres données en accès libre ou disponibles à la vente. Comme l'a promis Merrill,

elle applique des taux d'intérêt plus bas que ceux de la plupart des prêteurs sur salaire. Un prêt de 500 dollars chez ZestFinance revient habituellement au bout de vingt-deux semaines à 900 dollars – 60 % de moins que la norme du secteur.

C'est un progrès, mais le système est-il juste ? Les algorithmes de l'entreprise analysent jusqu'à 10 000 données par candidat, parmi lesquelles certaines observations insolites : ont-ils une bonne orthographe et font-ils un usage correct des majuscules dans leur formulaire de demande ? Combien de temps mettent-ils à le lire ? Se donnent-ils la peine de consulter les modalités du contrat ? « Ceux qui respectent les règles », soutient ZestFinance, présentent un moindre risque de crédit.

Peut-être bien. Mais les fautes de ponctuation et d'orthographe sont aussi le signe d'un faible niveau d'instruction, qui se trouve fortement corrélé avec la classe sociale et les origines ethniques. Lorsque des personnes pauvres ou immigrées s'avèrent éligibles à un prêt, leurs compétences linguistiques parfois médiocres sont donc susceptibles de faire grimper les frais. Si elles ont ensuite du mal à payer, ces difficultés pourront confirmer qu'elles présentaient dès le départ un risque élevé, et affaiblir encore leurs scores de crédit. On observe là une boucle de rétroaction perverse, et le fait de régler ses factures dans les délais ne joue à cet égard qu'un rôle minime.

Quand de nouvelles initiatives prennent appui sur des ADM, les problèmes s'ensuivent inmanquablement, même si les acteurs concernés affichent les meilleures intentions du monde. Prenons le cas de l'industrie du prêt « de pair à pair » (P2P). Elle est née au cours de la décennie passée, avec l'idée que les prêteurs et les emprunteurs noueraient contact sur des plateformes de mise en relation. Ce serait la démocratisation de l'activité bancaire. Davantage de personnes accèderaient à un prêt, tandis que des millions de gens ordinaires se changeraient en autant de petits banquiers et réaliseraient un joli retour sur investissement. On court-circuiterait des deux côtés les grandes banques avides et cupides.

L'une des premières plateformes de prêt entre particuliers, Lending Club, a été lancée sous forme d'une application Facebook en 2006, et a bénéficié un an plus tard d'une levée de fonds lui permettant de devenir un nouveau type de banque. Pour calculer le risque emprunteur, Lending Club mêlait au traditionnel rapport de crédit des données recueillies aux quatre coins du Web. Son algorithme générait en un mot des e-scores, dont elle affirmait qu'ils étaient plus précis que les scores de crédit.

Lending Club et son principal rival Prosper sont encore de minuscules acteurs. Ils ont émis pour moins de 10 milliards de dollars de prêts, une goutte d'eau dans un marché du prêt à la consommation qui s'élève à 3 000 milliards. Ils suscitent cependant beaucoup d'attention. Des dirigeants de Citigroup et de Morgan Stanley siègent au conseil d'administration des organismes P2P, et le fonds d'investissement de Wells Fargo est le plus gros financeur de Lending Club. Son introduction en Bourse au mois de décembre 2014 fut la plus importante de l'année dans le secteur technologique : la firme a levé 870 millions de dollars et atteint une capitalisation de 9 milliards, ce qui en a fait en termes de valeur boursière la quinzième banque des États-Unis.

Tout cet engouement a peu à voir avec la démocratisation du capital ou l'élimination des intermédiaires. D'après une enquête publiée dans *Forbes*, les capitaux institutionnels représentent aujourd'hui plus de 80 % de toute l'activité sur les plateformes P2P. Pour les grandes banques, ces nouvelles plateformes offrent une alternative pratique vis-à-vis d'une économie bancaire étroitement régulée. En passant par des systèmes de pair à pair, le prêteur peut analyser à peu près toutes les données qu'il souhaite et élaborer ses propres e-scores. Il peut établir des corrélations de risque en fonction des quartiers, des codes postaux et des magasins que les clients fréquentent – sans jamais être tenu de s'en expliquer par d'embarrassants courriers.

Qu'est-ce que cela signifie pour nous ? Avec le développement inexorable des e-scores, nous sommes regroupés et catégorisés sur la base de formules secrètes, reposant pour certaines sur des portefeuilles de données gorgés d'erreurs. On nous envisage non comme des individus mais comme les membres de différentes tribus, et cette classification nous colle à la peau. Plus la sphère financière se trouve polluée par les e-scores, et plus les personnes démunies voient leurs opportunités s'amenuiser. À dire vrai, comparé à cette foule d'ADM opérant de manière incontrôlée, l'agent de prêt de jadis, pétri de préjugés, ne paraît plus si terrible. Un emprunteur pouvait au moins lire dans son regard et en appeler à son humanité.

---

1. Il existe deux principaux types de cartes bancaires aux États-Unis. La *credit card* permet de faire des achats avec de l'argent directement prêté par la banque. La somme dépensée est due à la fin de chaque mois, mais les banques encouragent souvent à payer cette dette de manière échelonnée, avec toutefois de forts taux d'intérêt. Le plafond autorisé et le taux d'intérêt dépendent du score de crédit du détenteur de la carte. La carte de débit, *debit card* est l'équivalent de la carte bancaire française, qui sert à retirer l'argent détenu par le consommateur sur son compte, voire davantage (un découvert) à condition de payer un intérêt à la banque.

2. Précisons malgré tout que leur rectification peut virer au cauchemar. Patricia Armour, résidente du Mississippi, a vainement demandé pendant deux ans que Experian supprime de son fichier une dette de 40 000 dollars qui n'était plus due. Il lui a fallu contacter le procureur général du Mississippi, a-t-elle expliqué au *New York Times*, pour que Experian finisse par corriger son dossier (*NdA*).

## Souscrire à une assurance

# ZONE À RISQUE

1001Ebooks

À la fin du XIX<sup>e</sup> siècle, un célèbre statisticien allemand du nom de Frederick Hoffman créa une puissante ADM. Employé par la compagnie d'assurances Prudential Life, Hoffman ne pensait probablement pas à mal. Plus tard, les recherches effectuées au cours de sa vie contribueraient en effet grandement à la santé publique. Il réaliserait de précieux travaux sur la malaria et serait parmi les premiers à établir un lien entre cancer et tabac. Malheureusement, un beau jour du printemps 1896, Hoffman publia un rapport de 330 pages qui eut pour effet d'entraver la cause de l'égalité raciale aux États-Unis et de renforcer le statut de citoyens de seconde zone auquel des millions de gens se trouvaient relégués. S'appuyant sur des statistiques exhaustives, son rapport défendait l'argument selon lequel les Afro-Américains vivaient une existence si précaire que leur groupe ethnique tout entier était impossible à assurer.

L'analyse d'Hoffman, comme beaucoup des ADM dont nous avons parlé, était à cet égard défailante. Il confondait causalité et corrélation, de sorte que les volumineuses données réunies ne servaient qu'à confirmer sa thèse, à savoir que l'origine ethnique constituait un puissant prédicteur de l'espérance de vie. Le racisme était à ce point enraciné dans sa pensée qu'il ne s'était apparemment jamais demandé si la pauvreté et l'injustice n'avaient pas quelque chose à voir avec le taux de mortalité des Afro-Américains ; si le manque d'écoles décentes, la vétusté des installations sanitaires, la dangerosité des environnements de travail et les difficultés d'accès aux soins médicaux n'étaient pas responsables de leurs morts précoces.

Hoffman avait en outre commis une erreur fondamentale sur le plan statistique. À l'instar de la commission présidentielle, auteure du rapport *A Nation at Risk* de 1983, il avait omis de stratifier ses résultats.

Il n'envisageait les Noirs que sous l'aspect d'un vaste groupe homogène et s'était donc abstenu de les séparer en différents segments géographiques, sociaux ou économiques. Entre un maître d'école noir menant une vie tranquille à Boston ou à New York, et un métayer s'échinant pieds nus douze heures par jour dans le delta du Mississippi, il ne faisait aucune distinction. Hoffman était aveuglé par la notion de race.

C'était d'ailleurs également le cas de l'industrie qui l'employait. Au fil du temps, bien entendu, l'opinion des assureurs a quelque peu évolué, et ils ont commencé à vendre des contrats aux familles afro-américaines. Après tout, il y avait là aussi de l'argent à gagner. Pendant des décennies, ils se sont toutefois cramponnés à l'idée d'Hoffman, selon laquelle des groupes entiers d'individus étaient plus risqués à assurer que d'autres – et certains trop risqués. Les compagnies d'assurances, au même titre que les banquiers, délimitaient en rouge les quartiers où elles refusaient d'investir. Cette pratique cruelle, connue sous le nom de « *redlining* », a été rendue illégale par divers textes législatifs et notamment par le Fair Housing Act (« loi sur le logement équitable »), promulgué en 1968.

Un demi-siècle plus tard, le *redlining* se perpétue, quoique sous des formes bien plus subtiles. Il s'inscrit dans le code des ADM les plus récentes. Comme Hoffman, les créateurs de ces nouveaux modèles confondent corrélation et causalité. Ils punissent les pauvres, et en particulier les minorités ethniques. Ils s'appuient également sur des tonnes de statistiques, qui confèrent à leurs analyses l'allure bien étudiée d'une science impartiale.

Sur le trajet d'une existence régie par les algorithmes, nous nous sommes frayé un chemin à travers les études et avons décroché un emploi (même si ce dernier nous soumet à un planning chaotique). Nous avons contracté des prêts et observé que notre solvabilité se substituait à l'appréciation de nos vices et vertus. Il est temps à présent de protéger nos biens les plus précieux – notre logement, notre véhicule et la santé de notre famille – et de prendre des dispositions pour les proches que nous laisserons un jour derrière nous.

L'assurance s'est développée à partir de la science actuarielle, discipline dont les racines remontent au XVII<sup>e</sup> siècle. Une bourgeoisie européenne en plein essor connaissait à cette époque une prospérité grandissante – offrant à beaucoup, pour la première fois, le luxe de penser aux générations futures.

À mesure que les progrès des mathématiques fournissaient les outils nécessaires pour effectuer des prédictions, une première génération de

collecteurs de données se mit en quête d'évaluer de nouveaux paramètres. L'un d'eux était un marchand de tissus londonien du nom de John Graunt. Il explora les registres de naissances et de décès, et conduisit en 1682 la première étude des taux de mortalité d'une collectivité entière. Il calcula par exemple que les enfants de Londres présentaient, au cours de chacune de leurs six premières années d'existence, un risque de décès de 6 % (les statistiques lui permirent également de réfuter le mythe selon lequel une épidémie de peste se répandait systématiquement l'année où un nouveau monarque montait sur le trône). Pour la première fois, les mathématiciens pouvaient calculer la trajectoire de vie la plus probable d'une personne. Bien entendu, ces chiffres ne fonctionnaient pas à l'échelle individuelle. Mais avec un échantillon suffisamment large, moyenne et écart devenaient prévisibles.

Les mathématiciens ne prétendaient pas anticiper le destin de chaque individu, qui restait de l'ordre de l'inconnu. Mais ils pouvaient prédire au sein de grands groupes humains la prévalence des accidents, des incendies et des décès. Dans les trois siècles qui ont suivi, une vaste industrie de l'assurance s'est développée autour de ces prédictions. Cette nouvelle industrie donna la chance aux gens, pour la première fois, de mutualiser leur risque collectif et de protéger ainsi les individus en cas de malheur.

À l'heure actuelle, avec l'évolution de la science des données et des réseaux informatiques, l'assurance affronte une transformation fondamentale. Disposant d'informations toujours plus nombreuses – entre autres sur notre génome, nos habitudes en matière de sommeil, d'exercice physique et de régime alimentaire, et nos compétences de conducteur –, les assureurs vont opérer de plus en plus un calcul de risque en fonction de l'individu, et se libérer des généralités communes à une population plus large. Beaucoup de personnes accueilleront ce changement d'un œil favorable. Données à l'appui, une fervente adepte d'une vie saine peut démontrer dès aujourd'hui qu'elle dort huit heures par nuit, qu'elle parcourt chaque jour quinze kilomètres à pied, et qu'elle suit un régime frugal à base de légumes verts, d'amandes et d'huile de poisson. Pourquoi ne bénéficierait-elle pas d'une réduction sur son assurance maladie ?

Le mouvement vers l'individualisation, comme nous le verrons, est encore embryonnaire. Mais les assureurs exploitent d'ores et déjà les données pour nous répartir en « tribus » de plus petite taille, et nous proposer divers produits et services à des tarifs variables. Certains pourront



parler d'une prestation sur mesure. Le problème, c'est qu'elle n'est pas complètement individualisée. Les modèles nous classent sans qu'on le voie dans différents groupes, dont les comportements ressemblent au nôtre à première vue. Quelle que soit la qualité de l'analyse, son opacité ouvre la porte à une véritable escroquerie.

Prenons l'assurance automobile. En 2015, des enquêteurs de *Consumer Reports* (l'équivalent américain de *Que choisir*) ont mené aux États-Unis une grande étude nationale sur les disparités tarifaires. Ils ont analysé plus de 2 milliards de devis provenant de tous les principaux assureurs, pour d'hypothétiques clients situés dans chacune des 33 419 circonscriptions postales du pays. Ils ont constaté de la sorte de profondes injustices, rattachées – comme on l'a vu au chapitre précédent – aux scores de crédit.

Les assureurs obtiennent ces scores à partir des rapports de solvabilité et génèrent ensuite, à l'aide d'un algorithme propriétaire, leurs propres évaluations (ou e-scores) qui servent de mesure indirecte pour attester d'une conduite responsable. Mais *Consumer Reports* a découvert que les e-scores, qui intègrent toutes sortes de données démographiques, comptent souvent davantage que l'historique du conducteur. En d'autres termes, la manière dont on gère son argent peut avoir plus d'impact que la manière dont on conduit. Dans l'État de New York, par exemple, passer d'une cote de crédit « excellente » à simplement « bonne » pouvait faire grimper la prime d'assurance annuelle de 255 dollars, tandis qu'en Floride les adultes qui présentaient en tant que conducteur un historique vierge mais affichaient un mauvais score de crédit payaient en moyenne 1 552 dollars de plus que les mêmes personnes gratifiées d'une parfaite solvabilité mais condamnées pour conduite en état d'ivresse.

Nous avons déjà exposé que le recours croissant aux scores de crédit dans l'ensemble de l'économie jouait en défaveur des pauvres. On a donc en l'espèce une nouvelle illustration de cette tendance, d'autant plus flagrante que l'assurance du véhicule est obligatoire pour tout conducteur. Ce qui diffère ici, c'est l'accent mis sur un élément accessoire alors même que des données bien plus pertinentes sont disponibles. Pour des assureurs automobiles, je ne saurais imaginer d'information plus probante qu'une mention de conduite en état d'ivresse. C'est la preuve d'un risque dans un domaine qu'ils s'efforcent précisément de prédire. Elle vaut bien mieux que d'autres indicateurs indirects pris en considération, comme la moyenne des notes obtenues au lycée. Malgré cela, il arrivera qu'elle pèse beaucoup

moins dans leur formule qu'un score fondé sur des données financières, réunies dans un rapport de solvabilité (dont nous avons vu en outre qu'il est parfois erroné).

Dès lors, pourquoi leurs modèles se focalisent-ils sur les scores de crédit ? À l'instar d'autres ADM, les systèmes automatiques sont capables d'éplucher ces scores avec une efficacité impressionnante, et sur une échelle considérable. Mais la principale raison touche selon moi à la question des profits. Si un assureur possède un système qui lui permet d'obtenir d'un conducteur à l'historique vierge 1 552 dollars de plus par an, pourquoi s'en priver ? Les victimes de son ADM, comme on l'a vu ailleurs, ont de grandes chances d'être pauvres, de manquer d'instruction et d'être issues pour bon nombre de l'immigration. Il est peu probable qu'elles se sachent escroquées. Et dans des quartiers qui comptent davantage de bureaux de prêt sur salaire que de courtiers en assurances, difficile en outre de chercher de meilleurs tarifs. En bref, bien qu'un e-score ne soit pas forcément corrélé avec une conduite prudente, il produit en revanche un vivier lucratif de conducteurs vulnérables. Beaucoup d'entre eux sont prêts à tout pour pouvoir conduire – leur emploi en dépend. Les faire payer plus cher s'avère bénéfique, au moins sur le plan des résultats financiers.

Pour l'assureur automobile, en revanche, c'est du gagnant-gagnant. Un bon conducteur affecté d'un mauvais score de crédit fait courir un risque faible et se montre extrêmement rémunérateur. La compagnie peut de surcroît utiliser une partie des bénéfices générés par ce contrat pour pallier les inefficiences du modèle – notamment lorsque des conducteurs dotés d'un rapport de solvabilité impeccable, et versant de ce fait des primes peu élevées, fracassent leur véhicule en état d'ébriété.

Cela pourrait sembler un brin cynique. Mais voyez plutôt l'algorithme d'optimisation tarifaire mis en œuvre par Allstate, cet assureur qui affirme dans son slogan : « Vous êtes entre de bonnes mains. » Selon la Consumer Federation of America (CFA), une association de défense des consommateurs, Allstate analyse les données démographiques et les comportements de consommation pour déterminer la probabilité que les clients fassent jouer la concurrence. Si c'est peu probable, il est alors judicieux de leur appliquer des tarifs supérieurs. Et c'est exactement ce que fait Allstate.

Mais ce n'est pas tout. Dans un dossier déposé auprès du département des assurances du Wisconsin, la CFA a dénombré chez Allstate 100 000

microsegments de tarification, définis sur la base du montant maximum que chaque groupe pourrait accepter de payer. Certains obtiennent en conséquence jusqu'à 90 % de rabais par rapport à la cotisation moyenne, tandis que d'autres subissent un surcoût de 800 %. « La tarification d'Allstate a perdu tout lien avec les règles de fixation des primes fondées sur l'évaluation du risque, et ne respecte plus la règle de droit », a expliqué J. Robert Hunter, responsable du secteur assurances pour la CFA et ancien commissaire aux assurances du Texas. Allstate a répondu que les accusations de la CFA étaient inexactes, mais la compagnie a néanmoins reconnu que « les considérations de marché, conformes aux pratiques de l'industrie, s'étaient révélées opportunes dans l'élaboration des tarifs d'assurances ». Ses modèles étudient en d'autres termes une foule de données indirectes pour calculer le tarif à appliquer aux clients. Et les autres acteurs du secteur font de même.

La tarification qui en découle est injuste. De tels abus ne pourraient survenir si les prix des assurances étaient transparents, et si les clients pouvaient facilement les comparer. Mais comme d'autres ADM, celle-ci est opaque. Chaque individu vit une expérience différente, et les modèles sont optimisés afin d'extorquer aux ignorants et aux personnes acculées le plus d'argent possible. Le résultat – une nouvelle boucle de rétroaction –, c'est que les conducteurs pauvres, qui sont les moins à même de payer des primes exorbitantes, se voient dépossédés jusqu'au dernier centime. Le modèle est finement ajusté de manière à tirer de ce sous-groupe le maximum d'argent. Certains décrochent inévitablement et se retrouvent dans l'incapacité d'honorer leur prêt automobile, de rembourser le solde de leur carte de crédit ou de payer leur loyer. Leur score de solvabilité en est derechef pénalisé, ce qui les rejette sans nul doute au sein d'un microsegment encore plus misérable.

\*\*\*

En même temps que son rapport accablant sur les assureurs automobile, *Consumer Reports* a également lancé une campagne visant l'Association nationale des commissaires aux assurances (National Association of Insurance Commissioners, ou NAIC), déclinée y compris sur Twitter : « @NAIC\_News à l'attention des commissaires aux assurances : « Fixez le

prix selon ma façon de conduire, pas selon l'idée que vous vous faites de moi ! #ReformerAssuranceAuto »

L'idée sous-jacente était que les conducteurs devaient être jugés en fonction de leurs antécédents – nombre de PV pour excès de vitesse, ou implication éventuelle dans un accident – et non pas de leurs habitudes de consommation, ni de celles de leurs amis ou de leurs voisins. À l'ère du Big Data, cependant, l'exhortation à nous juger sur notre façon de conduire revêt un sens tout à fait nouveau.

Les compagnies d'assurances possèdent aujourd'hui de multiples moyens d'étudier dans les moindres détails le comportement des conducteurs. Pour s'en faire une idée, il suffit de regarder l'industrie du transport routier.

Beaucoup de camions embarquent de nos jours un dispositif d'enregistrement électronique qui consigne chaque virage, chaque accélération, chaque pression sur les freins. En 2015, Swift Transportation, le plus important transporteur routier des États-Unis, a aussi commencé à installer des caméras pointées dans deux directions : l'une vers la route, l'autre vers le visage du conducteur.

Cette surveillance a pour objectif déclaré de réduire les accidents. Environ 700 camionneurs meurent chaque année sur les routes américaines, et leurs accidents ôtent également la vie à de nombreuses personnes circulant dans d'autres véhicules. Outre la tragédie humaine, cela coûte beaucoup d'argent. D'après la Federal Motor Carrier Safety Administration (Agence fédérale pour la sécurité du transport routier), le coût moyen d'un crash mortel s'élève à 3,5 millions de dollars.

Mais avec un laboratoire d'analyse aussi immense à leur disposition, les compagnies de transport ne se limitent pas à la sécurité. Si l'on combine la géolocalisation, les technologies d'enregistrement à bord et les caméras, les conducteurs de poids lourds livrent un flux permanent et très étoffé de données comportementales. Les transporteurs peuvent désormais analyser différents itinéraires, évaluer la gestion du carburant, et comparer les résultats à différentes heures du jour et de la nuit. Ils peuvent même calculer les vitesses idéales pour divers revêtements de chaussée. Et ils utilisent ces données pour déterminer les configurations qui procurent le plus fort revenu au moindre coût.

Ils peuvent aussi comparer individuellement les conducteurs. Des tableaux de suivi analytique établissent pour chacun une fiche de performance. En un ou deux clics, un manager peut ainsi identifier les pires

et les meilleurs d'entre eux sur un large éventail de critères. Naturellement, ces données de surveillance permettent aussi d'estimer le niveau de risque de chaque conducteur.

Cette promesse n'a pas échappé à l'industrie de l'assurance. Les plus grandes compagnies, telles que Progressive, State Farm ou Travelers, proposent déjà aux conducteurs un rabais sur leurs primes s'ils acceptent de partager leurs données de conduite. Une petite unité télémétrique installée dans la voiture, version simplifiée de la boîte noire des avions, enregistre la vitesse du véhicule et la manière dont le conducteur freine et accélère. Un moniteur GPS suit ses déplacements.

Voilà qui répond en théorie à l'idéal défendu par la campagne de *Consumer Reports*. Le conducteur individuel passe au premier plan. Prenons le cas des jeunes de 18 ans. Ils versent par tradition des primes astronomiques, au motif que leur tranche d'âge commet statistiquement plus que sa part d'imprudences. Or, aujourd'hui, un lycéen de terminale qui évite les démarrages en trombe, qui respecte les limitations de vitesse et qui freine lentement à l'approche d'un feu rouge pourrait bénéficier d'un tarif réduit. Les compagnies d'assurances accordent depuis longtemps un avantage aux jeunes automobilistes qui ont suivi des cours d'instruction routière au lycée, ou qui obtiennent une mention. Ces éléments tiennent lieu d'indicateurs indirects quant à leur responsabilité au volant. Mais les véritables données sont celles qui touchent à la conduite elle-même. Voilà qui est mieux, n'est-ce pas ?

Mais il reste néanmoins quelques problèmes. Premièrement, si le système associe le risque à la localisation géographique, les conducteurs pauvres se trouvent lésés. Ils ont en effet de plus grandes chances de circuler dans des quartiers que les assureurs considèrent comme risqués. Beaucoup effectuent aussi de longs trajets domicile-travail à des horaires irréguliers, ce qui se traduit par un risque accru.

Eh bien soit ! pourrait-on se dire. Si les quartiers pauvres sont plus risqués, en particulier pour les vols de véhicules, pourquoi les compagnies d'assurances devraient-elles ignorer cette information ? Et si des trajets professionnels plus longs augmentent la probabilité d'accident, les assureurs sont aussi en droit d'en tenir compte. Leur appréciation demeure basée sur le comportement de l'automobiliste, et non sur des détails sans rapport, comme sa cote de crédit ou les historiques de conduite des

personnes de son âge. Beaucoup de gens considéreront qu'il s'agit là d'une amélioration.

Et dans une certaine mesure, c'est effectivement le cas. Mais prenons une hypothétique conductrice habitant un secteur mal famé de Newark, dans le New Jersey, qui doit parcourir vingt kilomètres jusqu'au Starbucks où elle travaille comme barista, dans la banlieue aisée de Montclair. Son planning est chaotique et comporte un *clopping* de temps à autre. Elle ferme donc l'établissement à 23 heures, rentre à Newark, puis revient peu avant 5 heures du matin. Pour gagner dix minutes et économiser 1,50 dollar dans chaque sens au péage du Garden State Parkway, elle emprunte un raccourci qui la fait passer par une route où se succèdent bars et clubs de strip-tease.

Un assureur rompu au traitement de données notera que les voitures qui circulent sur cette route au petit matin présentent un risque supérieur d'accident. On y croise plus d'un conducteur ivre. Et pour être honnête, notre barista rajoute elle-même un facteur de risque en empruntant ce raccourci et en partageant la chaussée avec tous ces gens qui sortent des bars. L'un d'entre eux pourrait la percuter. Pour le géolocalisateur de la compagnie d'assurances, cependant, non contente de frayer avec des ivrognes, rien ne dit qu'elle n'en soit pas également une.

Ainsi donc, les modèles qui suivent notre comportement individuel tirent malgré tout quantité de leurs observations, et évaluent le risque, en nous comparant à d'autres. Au lieu de regrouper les individus qui parlent arabe ou ourdou, qui habitent la même localité ou qui ont un salaire proche, ils réunissent cette fois ceux qui agissent de manière semblable. La prédiction est que des personnes au comportement analogue prendront des niveaux de risque similaires. Si vous ne l'aviez pas encore remarqué, c'est à nouveau le principe du « qui se ressemble s'assemble », empreint en grande partie des mêmes injustices.

Chez la plupart des gens, l'opposition aux boîtes noires dans les voitures, quand j'en discute avec eux, ne tient pas tant à l'analyse des données qu'à la surveillance elle-même. Ils m'affirment avec insistance qu'ils n'accepteront pas d'être épiés. Ils refusent d'être pistés et de voir leurs informations vendues à des publicitaires ou transmises à la National Security Agency. Certains parviendront peut-être à résister, mais le respect de la vie privée exigera de plus en plus d'en payer le prix.

N'en étant qu'à leurs débuts, les systèmes de suivi des assureurs automobile fonctionnent pour l'instant sur la base d'un consentement

préalable et explicite (l'« *opt-in* »). Ne mettent en marche leur boîte noire que les conducteurs disposés à être pistés<sup>1</sup>. Ils se voient récompensés d'une remise allant de 5 à 50 %, avec la promesse d'obtenir plus tard des réductions plus importantes (et ce sont les autres conducteurs qui subventionnent ces rabais, par le versement de primes plus élevées). Mais plus les assureurs recueilleront d'informations, et plus ils pourront améliorer la fiabilité de leurs prédictions. C'est la nature même de l'économie des données. Ceux qui sauront en extraire le plus de connaissances, et les convertir en profits, sortiront gagnants. Ils prédiront avec une précision accrue les niveaux de risque associés à différents groupes (même si les comportements individuels ne cesseront jamais de les déconcerter). Et plus ils tireront bénéfice des données, plus ils déploieront d'efforts pour en collecter davantage.

À partir d'un certain moment, les dispositifs de suivi deviendront probablement la norme. Et les consommateurs désireux de gérer leur assurance à l'ancienne, en ne dévoilant que les renseignements essentiels, devront payer une surprime – d'un montant sans doute exorbitant. Dans l'univers des ADM, le respect de la vie privée représente de plus en plus un luxe que seuls les riches peuvent se permettre.

En parallèle, la surveillance modifiera la nature profonde de l'assurance. C'est une industrie qui, au sein d'une collectivité, met traditionnellement à contribution la majorité pour répondre aux besoins d'une minorité de malchanceux. Dans les villages que nous habitions autrefois, il y a de cela plusieurs siècles, les familles, les communautés religieuses et les voisins se prêtaient main-forte pour prendre soin les uns des autres en cas d'incendie, d'accident ou de maladie. Dans l'économie de marché, nous sous-traitons ce soin auprès des compagnies d'assurances, qui conservent pour elles une part de l'argent à titre de profit.

Plus ils en apprendront sur nous, et plus les assureurs seront en mesure d'identifier les clients qui semblent présenter le plus de risque. Ils pourront alors soit leur appliquer des tarifs stratosphériques, soit, quand la loi le permet, refuser de les couvrir. On s'éloigne là grandement de l'objectif initial de l'assurance, qui est d'aider la société à équilibrer son risque. Dans un monde fondé sur le ciblage, le tarif que nous payons n'est plus calculé en fonction d'une moyenne. Il est au contraire lesté de coûts fixés par anticipation. Au lieu d'adoucir les aléas de l'existence, les compagnies d'assurances nous en réclameront d'avance le paiement. Voilà qui sape

l'objet même de l'assurance... et l'impact sera particulièrement rude pour ceux qui sont le moins en mesure de l'encaisser.

À mesure qu'elles décortiqueront nos habitudes de vies et nos profils physiques, les compagnies d'assurances vont nous répartir dans de nouveaux types de tribus. Ces dernières ne reposeront cependant pas sur des critères traditionnels tels que l'âge, le sexe, le patrimoine net ou le code postal, mais plutôt sur des critères comportementaux, et seront générées presque entièrement par des machines.

Pour observer de quelle manière ce genre de tri est appelé à se généraliser, prenons l'exemple de l'entreprise new-yorkaise Sense Networks, spécialisée dans les données. Il y a une dizaine d'années, ses chercheurs ont commencé à analyser des données de téléphonie mobile qui révélaient où les gens s'étaient rendus. Fournies par les opérateurs européens et américains, celles-ci restaient anonymes : elles ne montraient que de simples points se déplaçant sur des cartes. (Bien entendu, il n'aurait pas fallu chercher beaucoup pour associer l'un de ces points avec l'adresse où il retournait tous les soirs de la semaine. Mais Sense ne s'intéressait pas aux individus ; elle s'intéressait aux tribus.)

L'équipe injecta ces données sur les utilisateurs de mobiles new-yorkais dans son système d'apprentissage automatique (*machine learning*), sans toutefois fournir guère plus de directives à l'ordinateur. Ce dernier ne reçut aucune instruction pour isoler les banlieusards ou « milléniaux » (*millenials*), la génération née entre le début des années 1980 et le début des années 2000, ou pour créer différentes catégories d'acheteurs. Le logiciel trouverait des similitudes par lui-même. Beaucoup seraient idiots – les gens qui passaient plus de la moitié de leur journée dans des rues commençant par la lettre « J », ou ceux qui prenaient la plupart de leurs pauses déjeuner à l'extérieur. Mais si le système explorait des millions de données de ce type, des schémas finiraient par apparaître. Des corrélations se feraient jour, dont un grand nombre que des êtres humains n'auraient vraisemblablement jamais pris en compte.

Au fil des jours, tandis que l'ordinateur de Sense digérait cette masse énorme de données, les points se mirent à revêtir différentes couleurs. Certains virèrent au rouge, d'autres au jaune, au bleu ou au vert. Les tribus étaient en train d'émerger.

Que représentaient-elles ? La machine seule le savait, et elle n'était pas douée de parole. « Nous ne discernions pas forcément ce que ces gens



avaient en commun, a expliqué Greg Skibiski, cofondateur et ancien PDG de Sense. Ils n'entraient pas dans les catégories traditionnelles que nous pouvions définir. » L'équipe de Sense put alors suivre les déplacements des tribus, parées de leurs couleurs, à travers New York. En journée, certains quartiers affichaient une dominance de bleu puis devenaient rouges le soir, parsemés de jaune. L'une des tribus, s'est remémoré Skibiski, semblait fréquenter un lieu précis tard dans la nuit. Était-ce une discothèque ? Une fumerie de crack ? Quand l'équipe de Sense a regardé l'adresse, elle s'est aperçue qu'il s'agissait d'un hôpital. Les membres de cette tribu semblaient se blesser ou tomber plus souvent malades. Ou peut-être étaient-ils médecins, infirmiers ou urgentistes.

Sense a été revendue en 2014 à YP, société de publicité sur mobile née d'une scission avec l'opérateur téléphonique AT&T. Sa démarche de tri va donc servir pour l'heure à diffuser des publicités ciblées auprès de telles ou telles tribus. Mais on peut imaginer que des systèmes d'apprentissage automatique alimentés par divers flux de données comportementales nous rangeront bientôt non pas dans une seule tribu, mais dans plusieurs centaines, voire plusieurs milliers de tribus différentes. Leurs membres réagiront par exemple au même genre de publicités. D'autres présenteront peut-être des affinités sur le plan politique, ou iront plus fréquemment en prison. D'autres encore adoreront les fast-foods.

Ce que je veux souligner, c'est que des océans de données comportementales iront directement nourrir dans les années à venir les systèmes d'intelligence artificielle. Et ces systèmes demeureront aux yeux de la population de vraies boîtes noires. Tout au long de ce processus, nous apprendrons rarement à quelles tribus nous « appartenons » et pourquoi nous en faisons partie. À l'ère de l'intelligence artificielle, la plupart des variables resteront un mystère. Beaucoup de ces tribus muteront en outre d'heure en heure, si ce n'est de minute en minute, à mesure que les systèmes renverront les individus d'un groupe à l'autre. Car la même personne agit en effet très différemment à 8 heures et à 20 heures.

Ces programmes automatiques détermineront de plus en plus la manière dont les autres machines nous traitent – celles qui choisissent les publicités que nous voyons, qui fixent les prix auxquels nous sommes soumis, qui nous attribuent un rendez-vous chez le dermatologue, ou qui organisent nos trajets. Elles feront preuve d'une extrême efficience, d'un caractère

apparemment arbitraire, et n'auront à répondre de rien. Personne ne comprendra leur logique ni ne saura l'expliquer.

Si nous ne reconquérons pas un peu de maîtrise, ces futures ADM nous sembleront de plus en plus mystérieuses et toutes-puissantes. Elles nous imposeront leur volonté, presque sans que l'on s'en aperçoive.

\*\*\*

En 1943, au plus fort de la Seconde Guerre mondiale, alors que les armées et les industries américaines avaient besoin de chaque soldat et de chaque travailleur disponible, l'administration fiscale des États-Unis (Internal Revenue Service) modifia le Code des impôts afin d'accorder un statut d'exonération aux dispositifs d'assurance santé des employeurs. La nouvelle semblait anodine, et ne faisait en tout cas pas le poids face aux gros titres sur la capitulation des Allemands à Stalingrad ou les débarquements alliés en Sicile. À l'époque, environ 9 % seulement des travailleurs américains bénéficiaient à titre d'avantage social d'une couverture santé privée. Mais avec ce nouveau statut d'exonération fiscale, les entreprises firent en sorte d'attirer les travailleurs, denrée rare, en leur garantissant ce genre de prestation. En l'espace de dix ans, 65 % des Américains allaient se soumettre aux systèmes mis en place par leurs employeurs. Les entreprises possédaient déjà une mainmise énorme sur leurs finances. Au cours de cette décennie-là, elles ont aussi acquis – qu'elles l'aient ou non souhaité – une certaine emprise sur leur corps.

Soixante-quinze ans plus tard, les coûts médicaux ont métastasé et engloutissent aujourd'hui 3 000 milliards de dollars par an. Près d'un dollar sur cinq que gagne un Américain alimente la vaste industrie de la santé <sup>2</sup>.

Les employeurs, qui ont longtemps ponctionné les travailleurs pour réduire leurs coûts, appliquent à présent une nouvelle tactique pour en combattre l'augmentation : le « *wellness* », c'est-à-dire l'hygiène de vie. Elle implique une surveillance accrue, incluant quantité de données issues des objets connectés – les bracelets Fitbit, les montres Apple Watch et autres capteurs qui transmettent des informations actualisées sur le fonctionnement de notre organisme.

L'idée, comme on l'a vu tant de fois, procède de bonnes intentions. Elle est en réalité encouragée par le gouvernement. Le Affordable Care Act, ou « Obamacare », invite les entreprises à enrôler les travailleurs dans des

programmes de bien-être, et même à instaurer en matière de santé des incitations financières. Les employeurs peuvent désormais, grâce à la loi, offrir des récompenses et infliger des pénalités atteignant jusqu'à 50 % du coût de la couverture santé. Selon une étude de la Rand Corporation, plus de la moitié des entreprises d'au moins cinquante employés ont d'ores et déjà un programme de *wellness* en place, et d'autres se joignent chaque semaine à cette tendance.

Ces programmes de *wellness* ne manquent pas de justifications. S'ils fonctionnent – et c'est un « si » majuscule, comme nous le verrons –, le travailleur et sa famille en sont les premiers bénéficiaires. Mais s'ils évitent aux salariés de développer des pathologies cardiaques ou diabétiques, les employeurs y gagnent aussi. Moins les employés d'une entreprise font de séjours aux urgences, et moins son personnel tout entier paraît présenter de risque aux yeux de la compagnie d'assurances, ce qui par suite fait baisser les primes. Si l'on parvient à dépasser son caractère intrusif, le *wellness*, ou bien-être par l'entreprise, peut donc apparaître comme une solution gagnant-gagnant.

Le problème, c'est que son caractère intrusif lui est indissociable. Impossible de faire comme s'il n'existait pas. Idem pour l'aspect coercitif. Prenons le cas d'Aaron Abrams, professeur de mathématiques à la Washington and Lee University, en Virginie. Il est couvert par Anthem Insurance, qui met en œuvre un programme de *wellness*. Pour respecter le programme, Abrams doit accumuler 3 250 « points santé ». Il en gagne un pour chaque connexion quotidienne, et 1 000 pour une visite annuelle chez le médecin ou un bilan santé effectué sur le campus. Il obtient également des points en répondant à une « enquête santé », où il se fixe lui-même des objectifs mensuels, et davantage encore s'il les atteint. S'il choisit de ne pas participer au programme, il doit payer chaque mois 50 dollars de supplément sur sa prime d'assurance.

Abrams avait été engagé pour enseigner les mathématiques. Or, à l'instar de millions d'autres Américains, son travail consiste maintenant en partie à suivre une série de diktats médicaux et à partager ces données non seulement avec son employeur, mais aussi avec l'entreprise tierce qui gère le programme. Il le supporte mal, et entrevoit le jour où l'université sera en mesure d'étendre sa surveillance. « Imaginer que “l'autosuiivi” de ma marche à pied permette à quelqu'un de reconstituer mes déplacements quotidiens, je trouve cela plus qu'effrayant », déclare-t-il.

Mes propres craintes vont un peu plus loin. Une fois que les entreprises auront amassé des mines de données sur l'état de forme des employés, qu'est-ce qui les empêchera d'élaborer des scores de santé et de s'en servir pour trier les candidats à l'embauche ? La plupart des indicateurs indirects collectés, qu'il s'agisse du nombre de pas parcourus ou des habitudes de sommeil, ne sont soumis à aucune protection juridique, si bien qu'il serait en théorie parfaitement légal de les utiliser à d'autres fins. Voilà qui serait en outre dans l'ordre des choses. Les entreprises, comme on l'a vu, écartent régulièrement des postulants sur la base des scores de crédit et des tests de personnalité. Les scores de santé représentent à cet égard une prochaine étape, aussi naturelle qu'inquiétante.

Certaines entreprises définissent dès à présent pour leurs employés d'ambitieuses normes en matière de santé, et les pénalisent s'ils sont en dessous. Le manufacturier de pneus Michelin leur fixe des objectifs relatifs à différentes mesures, allant de la tension artérielle à la glycémie, en passant par le cholestérol, les triglycérides et le tour de taille. Ceux qui n'atteignent pas la cible dans trois catégories se voient contraints de payer un supplément annuel de 1 000 dollars pour leur assurance santé. En 2013, la chaîne de drugstores américaine CVS a annoncé qu'elle imposerait à ses employés de déclarer leur niveau de masse adipeuse, leur glycémie, leur tension artérielle et leur taux de cholestérol – faute de quoi ils auraient à verser 600 dollars par an.

La démarche de CVS provoqua cette réaction houleuse de la part d'Alissa Fleck, chroniqueuse chez Bitch Media : « Votre attention s'il vous plaît ! Si vous vous battez depuis des années pour retrouver la forme, qu'importe ce que cela signifie pour vous, vous pouvez immédiatement abandonner ce que vous faites car CVS a tout compris. Quels que soient les ridicules efforts que vous tentiez, vous n'aviez simplement pas l'incitation adéquate. Sauf que le régime dont on parle existe en réalité déjà, c'est celui de l'humiliation et de la grossophobie. Laissez quelqu'un vous dire que vous êtes en surpoids, ou préparez-vous à payer une forte amende. »

Au cœur de la question du poids s'inscrit une statistique aujourd'hui discréditée : l'indice de masse corporelle (IMC), basé sur une formule conçue il y a deux siècles par un mathématicien belge, Lambert Adolphe Jacques Quetelet, qui ne connaissait pratiquement rien à la santé ni au corps humain. Il cherchait simplement une équation facile à utiliser pour mesurer

l'obésité au sein d'une vaste population. Il s'est appuyé à cette fin sur ce qu'il a appelé « l'homme moyen ».

« C'est un concept utile, écrit le mathématicien et auteur scientifique Keith Devlin. Mais si on essaie de l'appliquer à une seule personne, on en arrive par exemple à lui attribuer de manière absurde 2,4 enfants. Les moyennes mesurent des populations entières et ne sont souvent pas applicables aux individus. » Devlin ajoute que l'IMC, avec son score chiffré, donne à la « charlatanerie mathématique » une apparence d'autorité scientifique.

L'IMC se calcule en divisant le poids en kilogrammes du sujet par le carré de sa taille en mètres. C'est un indicateur indirect qui rend grossièrement compte de sa condition physique. Il aboutit volontiers à la conclusion que les femmes sont en surpoids (ce ne sont pas en effet des « hommes moyens »). Comme la graisse pèse moins que le muscle, les athlètes affutés affichent de surcroît fréquemment de très forts IMC. Dans cet univers parallèle, le joueur de basket LeBron James, avec ses 2,03 mètres et ses 113 kilos, appartient à la catégorie des individus en surcharge pondérale. Quand un mécanisme de récompenses et de sanctions économiques est rattaché à l'IMC, d'importants groupes de travailleurs se retrouvent pénalisés en raison de leur morphotype. Cela frappe tout particulièrement les femmes noires, qui possèdent souvent un indice élevé.

Mais n'est-il pas bénéfique, demanderont les défenseurs du *wellness*, d'aider les gens à s'occuper de leur poids et de leurs autres problèmes de santé ? La question essentielle ici est de savoir si cette aide revêt l'aspect d'une main tendue ou d'une injonction. Si les entreprises mettent en place des programmes de *wellness* gratuits et fondés sur le volontariat, peu de personnes auront de raisons de s'y opposer (et les travailleurs qui s'inscrivent à de tels programmes enregistrent de fait des progrès, même s'ils auraient fort bien pu les réaliser sans eux). Lier la rémunération à une statistique erronée comme l'IMC et contraindre les employés à modeler leur corps selon l'idéal de l'entreprise constitue en revanche une atteinte à la liberté. Cela donne aux employeurs une bonne excuse pour punir les gens dont ils n'apprécient pas l'apparence – et soustraire en même temps un peu d'argent de leurs poches.

Tout ceci est fait au nom de la santé. Parallèlement, l'industrie du *wellness*, qui représente un marché de 6 milliards de dollars, clairotte haut et fort ses succès – sans apporter souvent la moindre preuve. « Voici les

faits, indique Joshua Love, président de Kinema Fitness. Les gens en forme travaillent plus dur, sont plus heureux, aident les autres et sont plus efficaces. Ceux qui ne prennent pas soin d'eux sont en général apathiques, épuisés et sans entrain, l'attitude au travail étant souvent un reflet de leur mode de vie. »

Love n'a fourni naturellement aucune référence à l'appui de ces grandes assertions. Et même si elles étaient vraies, il n'existe guère de démonstration que les programmes de *wellness* obligatoires améliorent réellement la santé des travailleurs. Un rapport de recherche du California Health Benefits Review Program, un organisme chargé d'évaluer l'impact des avantages sociaux liés à la santé, conclut qu'ils ne réussissent à faire baisser ni la tension artérielle, ni la glycémie, ni le taux de cholestérol moyen des participants. Quand les gens parviennent néanmoins à perdre du poids dans le cadre d'un de ces programmes, ils ont tendance ensuite à reprendre les kilos perdus. (Le seul point laissant effectivement apparaître des résultats positifs concerne l'arrêt du tabac.)

Il s'avère également que les programmes de *wellness*, en dépit de réussites individuelles très médiatisées, n'entraînent souvent aucune baisse des dépenses en matière de santé. Une étude dirigée en 2013 par Jill Horwitz, professeur de droit à l'UCLA, met en pièces leurs soutènements économiques. Selon le rapport, des études randomisées « amènent à douter » que les employés fumeurs et obèses génèrent des frais médicaux plus élevés que les autres. S'il est vrai qu'ils sont davantage susceptibles de connaître des problèmes de santé, ceux-ci tendent à survenir plus tard au cours de leur vie, quand ils ne sont plus couverts par l'entreprise mais par le système Medicare. Les plus grosses économies engendrées par les programmes de *wellness* viennent en fait des pénalités infligées aux travailleurs. À l'instar des algorithmes de planning, ils procurent en réalité aux entreprises un outil supplémentaire pour piller la fiche de paie de leurs salariés.

Malgré les griefs qu'ils m'inspirent, ces programmes ne comptent pas (encore) comme des ADM à part entière. Ils sont assurément très répandus, s'immiscent dans la vie de millions d'employés et les frappent au portefeuille. Mais ils ne sont pas opaques et, à l'exception du bancal indice de masse corporelle, ne reposent pas sur des algorithmes mathématiques. Ils constituent simplement un cas évident et généralisé de vol sur salaire, enveloppé d'une rhétorique sanitaire fleurie.

Les employeurs font déjà une consommation excessive de nos données. Ils s'en servent comme nous l'avons vu pour nous évaluer, en tant que recrues potentielles et en tant que travailleurs. Ils s'efforcent de cartographier nos pensées, nos relations amicales, et de prédire notre productivité. Puisqu'ils sont déjà fortement impliqués dans l'assurance – la couverture médicale du personnel représentant un poste de dépense majeur –, il est tout naturel qu'ils étendent leur surveillance pour contrôler à grande échelle la santé des employés. Et si les entreprises concoctaient leurs propres modèles reliant l'état de forme à la productivité, cela pourrait déboucher sur la création d'une ADM en bonne et due forme.

---

1. En France, on appelle ce système la télémédecine embarquée, un procédé qui se démocratise de manière croissante dans les pratiques des assureurs automobile, sur la base du volontariat de la clientèle assurée.

2. Le système de santé aux États-Unis est largement privatisé et représente un budget colossal pour les individus comme les États. La population de plus de 65 ans est couverte par un système fédéral appelé Medicare, les plus démunis, par Medicaid, mais la majorité de la population doit recourir à des assurances privées parfois hors de prix, sans compter un coût des soins très élevé. La majorité est donc assurée par leur employeur, avec une prise en charge plus ou moins bonne. Cependant, certains hôpitaux publics, services d'urgence ou centres de santé offrent des soins gratuits.

# LE CITOYEN

## POUR CIBLE

La vie civique

1001Ebooks

**C**omme vous le savez à présent, je suis révoltée par toutes sortes d'ADM. Imaginons donc que je décide de lancer une campagne réclamant une régulation plus stricte, et que je poste une pétition sur ma page Facebook. Quels sont parmi mes amis ceux qui la verront apparaître sur leur fil d'actualités ?

Je n'en ai aucune idée. Dès que j'aurai cliqué sur « envoyer », cette pétition appartiendra à Facebook, et l'algorithme du réseau social jugera seul du meilleur usage à en faire. Il calculera les probabilités qu'elle a d'intéresser chacune de mes connaissances. Certaines d'entre elles, il le sait, signent souvent des pétitions et les partagent éventuellement avec leurs propres relations. D'autres ont plutôt tendance à les survoler sans s'y arrêter. Un certain nombre de mes amis m'accordent toutefois davantage d'attention et cliquent volontiers sur les articles que je publie. L'algorithme de Facebook prend tout cela en compte pour décider des personnes qui verront ma pétition. Chez beaucoup de mes contacts, elle sera si profondément enfouie dans leur fil d'actualités qu'ils ne l'apercevront jamais.

Voilà ce qui se passe quand le réseau immensément puissant que nous partageons avec 1,5 milliard d'utilisateurs se trouve être également une société cotée en Bourse. Bien que Facebook puisse donner l'impression d'une place de village moderne, c'est l'entreprise qui détermine, en fonction de ses propres intérêts, ce que les gens voient et apprennent sur les pages de son réseau social. Aux États-Unis, à l'heure où j'écris ces lignes, environ deux tiers des adultes possèdent un compte Facebook<sup>1</sup>. Ils passent trente-neuf minutes par jour sur le site, quatre de moins seulement qu'ils



n'en consacrent aux échanges dans le monde réel. Près de la moitié d'entre eux, selon un rapport du Pew Research Center, comptent sur Facebook pour fournir au moins une partie de leurs informations, ce qui nous amène à la question suivante : en ajustant son algorithme et en modelant les nouvelles que nous recevons, Facebook peut-il manipuler le système politique ?

Les propres chercheurs de l'entreprise se sont penchés sur le sujet. Lors des élections américaines de 2010 (législatives) et de 2012 (présidentielle), Facebook a mené des expériences pour mettre au point un outil qu'il a baptisé le « mégaphone électoral ». L'idée était d'inciter les gens à faire savoir qu'ils avaient voté. Cela paraissait plutôt raisonnable. En parsemant de statuts « j'ai voté » les fils d'actualités des utilisateurs, Facebook encourageait plus de 61 millions de citoyens américains à accomplir leur devoir civique et à faire entendre leur voix. En révélant leur comportement électoral, le site attisait de surcroît entre proches la pression d'aller voter. Des études ont en effet montré que la satisfaction muette de s'acquitter d'un devoir civique a moins de chances de faire bouger les gens que le jugement éventuel de leurs amis et de leurs voisins.

En parallèle, les chercheurs de Facebook ont étudié l'influence de différents types d'actualisations de statut sur la pratique électorale des individus. Aucun scientifique n'avait jamais travaillé au sein d'un laboratoire humain de cette envergure. En l'espace de quelques heures, Facebook pouvait recueillir des renseignements sur des dizaines de millions de personnes, sinon plus, et mesurer l'impact réciproque de leurs propos et des liens partagés. Il pouvait alors exploiter ce savoir pour influencer sur leurs actions – en l'occurrence, celle de voter.

C'est un pouvoir considérable. Et Facebook n'est pas la seule entreprise à l'exercer. D'autres sociétés cotées, parmi lesquelles Google, Apple, Microsoft, Amazon et des opérateurs de téléphonie mobile tels que Verizon et AT&T disposent d'une vaste somme d'informations sur une majeure partie de l'humanité – et par conséquent les moyens de nous diriger là où elles le souhaitent.

Leur objectif premier consiste généralement, comme on l'a vu, à gagner de l'argent. Mais leurs profits dépendent néanmoins étroitement des politiques publiques. L'administration leur applique des règles, ou choisit de ne pas le faire, approuve ou bloque leurs fusions et acquisitions, et fixe leur régime d'imposition (en fermant souvent les yeux sur les milliards placés à l'étranger dans des paradis fiscaux). Voilà pourquoi les sociétés

technologiques, comme toutes les autres entreprises américaines, inondent Washington de lobbyistes et injectent discrètement sous forme de contributions des centaines de millions de dollars dans le système politique. Elles se dotent aujourd'hui de la capacité d'orienter notre comportement électoral – et par là même de façonner le gouvernement –, en ajustant simplement leurs algorithmes.

La campagne menée par Facebook partait d'un objectif constructif et en apparence innocent : inciter les gens à voter. Et elle avait réussi. Après comparaison des registres de vote, les chercheurs ont estimé qu'elle avait conduit 340 000 personnes de plus aux urnes. Une foule suffisante pour faire basculer des États entiers, si ce n'est même les scrutins nationaux. En 2000, après tout, George W. Bush ne l'avait emporté qu'avec 537 voix d'avance en Floride. De toute évidence, un unique algorithme mis en œuvre par Facebook un jour d'élection pouvait non seulement modifier l'équilibre du Congrès mais décider aussi du nom du président.

La puissance de Facebook ne résulte pas seulement de sa portée mais aussi de son aptitude à utiliser ses propres clients pour influencer leurs amis. La grande majorité des 61 millions de personnes participant à l'expérience avaient reçu un message sur leur fil d'actualités qui les encourageait à voter. Ce message comportait une série de six photos d'amis Facebook de l'utilisateur, pris au hasard parmi ceux qui avaient cliqué sur le bouton « j'ai voté ». Les chercheurs observèrent aussi deux groupes témoins d'environ 600 000 membres chacun, l'un exposé à la campagne « j'ai voté » mais sans les photos d'amis, et l'autre qui n'avait rien reçu du tout.

En disséminant ses messages à travers le réseau, Facebook a pu étudier l'impact du comportement de nos amis sur le nôtre. Les gens allaient-ils inciter leurs connaissances à voter, et leur comportement en serait-il affecté ? D'après les calculs des chercheurs, voir ses amis participer au vote faisait toute la différence. Les gens accordaient beaucoup plus d'attention aux statuts « j'ai voté » lorsqu'ils provenaient de leurs relations, et se montraient davantage enclins à les partager. Environ 20 % des personnes qui voyaient que leurs amis avaient voté cliquaient elles aussi sur le bouton. Parmi celles qui n'avaient pas reçu le message de la part de leurs amis, seules 18 % le faisaient. Impossible de savoir à coup sûr si tous les gens qui avaient cliqué étaient effectivement allés voter, ni d'ailleurs si les autres étaient restés chez eux. Reste qu'avec 61 millions d'électeurs potentiels sur

le réseau, un éventuel écart de deux points pouvait prendre une ampleur énorme.

Deux ans plus tard, Facebook fit un pas supplémentaire. Pendant les trois mois précédant l'élection présidentielle, entre d'un côté le candidat sortant Barack Obama et de l'autre Mitt Romney, l'un de ses chercheurs, Solomon Messing, modifia l'algorithme du fil d'actualités d'environ 2 millions d'utilisateurs, tous politiquement engagés. Ces derniers reçurent une plus grande proportion de nouvelles sérieuses, par opposition aux vidéos de chats, aux remises de diplômes et aux photos de Disney World qu'ils avaient l'habitude de voir. Si leurs amis partageaient un article d'information, celui-ci s'affichait en tête de leur fil d'actualités.

Messing voulait vérifier si le fait de recevoir davantage d'informations de la part d'amis modifiait le comportement politique des individus. Au lendemain de l'élection, il envoya des questionnaires de sondage. Les réponses indiquèrent que la participation parmi ce groupe avait légèrement progressé, passant de 64 à 67 %. « Quand ce sont vos amis qui vous apportent le journal, a expliqué Lada Adamic, chercheuse en sciences sociales informatiques chez Facebook, cela génère des conséquences intéressantes. » Bien entendu, ce n'était pas réellement vos amis qui le livraient mais Facebook lui-même. On pourrait faire valoir que la presse exerce depuis des lustres un pouvoir similaire. Les rédacteurs en chef sélectionnent la nouvelle qui fera la une et décident de l'angle selon lequel elle sera présentée. Ils choisissent soit de montrer des Palestiniens bombardés, soit des Israéliens en deuil ; un policier en train de secourir un bébé, ou de frapper un manifestant. Ces choix peuvent sans nul doute influencer à la fois l'opinion publique et les scrutins électoraux. Même chose pour les informations télévisées. Mais quand le *New York Times* ou CNN couvrent une actualité, la différence est que tout le monde le voit. Leur décision éditoriale est parfaitement claire et assumée. Elle n'a rien d'opaque. Et les gens débattent ensuite (souvent sur Facebook) de son bien-fondé.

Facebook ressemble davantage au monde du *Magicien d'Oz* : l'intervention des êtres humains est invisible. Lorsqu'on visite le site, on fait défiler les statuts postés par nos amis. La machine semble n'être qu'un intermédiaire neutre. Beaucoup de gens le croient encore. En 2013, quand une chercheuse de l'université de l'Illinois, Karrie Karahalios, a réalisé un sondage sur l'algorithme de Facebook, elle a découvert que 62 % des gens

n'avaient pas du tout conscience que l'entreprise intervenait dans le fil d'actualités. Ils étaient persuadés que le système partageait instantanément tout ce qu'ils postaient avec l'ensemble de leurs amis.

L'emprise potentielle de Facebook sur notre système politique s'étend bien au-delà de la présentation des nouvelles et des campagnes d'incitation au vote. En 2012, des chercheurs ont mené une expérience sur 680 000 utilisateurs pour observer si les mises à jour de leur fil d'actualités étaient susceptibles d'affecter leur humeur. Des tests en laboratoire avaient déjà clairement montré que les humeurs sont contagieuses. Si vous côtoyez un individu grognon, vous avez toutes les chances de le devenir aussi, même brièvement. Mais ce genre de contagion pouvait-elle se propager en ligne ?

À l'aide d'un logiciel d'analyse linguistique, Facebook a trié les mises à jour positives (« c'est super ! ») et négatives (« je suis dégoûté ! »). Les chercheurs ont alors réduit le volume de publications pessimistes sur la moitié des fils d'actualités, et le quotient de publications optimistes sur l'autre moitié. L'étude des publications ultérieures a révélé que l'altération des fils d'actualités avait effectivement modifié l'humeur des utilisateurs. Ceux qui avaient vu moins de mises à jour joyeuses produisaient des messages plus négatifs. Une tendance similaire apparaissait du côté positif.

Conclusion des chercheurs : « Les états émotionnels peuvent se transmettre d'une personne à l'autre [...] ce qui conduit les gens à éprouver à leur insu les mêmes émotions. » Les algorithmes de Facebook sont capables autrement dit d'agir sur le moral de millions de gens sans que ceux-ci ne s'en aperçoivent. Que se passerait-il s'ils jouaient avec leurs émotions un jour de scrutin ?

Je n'ai aucune raison de croire que les experts en sciences sociales de Facebook manipulent activement le système politique. Ce sont pour la plupart des universitaires sérieux, effectuant des recherches sur une plateforme dont ils auraient à peine pu rêver il y a une vingtaine d'années. Mais ce qu'ils ont démontré, c'est l'immense pouvoir que possède Facebook pour peser sur ce que l'on apprend, sur ce que l'on ressent et sur le fait que l'on aille ou non voter. C'est une plateforme gigantesque, puissante et opaque. Ses algorithmes nous sont cachés, et nous ne voyons que les résultats des expériences que les chercheurs choisissent de publier.

Il en va à peu près de même pour Google. Son algorithme de recherche semble être axé sur l'accroissement du chiffre d'affaires. Mais les résultats

des requêtes, si Google en décidait ainsi, pourraient avoir un effet considérable sur ce que les gens apprennent et la manière dont ils votent. Deux chercheurs, Robert Epstein et Ronald E. Robertson, ont récemment demandé à des électeurs indécis, aux États-Unis ainsi qu'en Inde, de recourir à un moteur de recherche afin de se renseigner sur les élections à venir. Les moteurs utilisés étaient programmés pour fausser les résultats affichés et privilégier un parti plutôt que l'autre. Ces résultats, ont-ils indiqué, avaient modifié les préférences de vote de 20 %.

Si l'effet était si fort, c'est en partie parce que les gens font largement confiance aux moteurs de recherche. Selon un rapport du Pew Research Center, quelque 73 % des Américains pensent que les réponses qu'ils donnent sont à la fois exactes et impartiales. Si des entreprises comme Google trafiquaient les résultats de recherche pour favoriser telle conclusion politique au lieu de telle autre, elles mettraient donc en péril leur propre réputation et s'exposeraient à des sanctions réglementaires.

Cela dit, qui le saurait ? Ce que l'on découvre à propos de ces géants d'Internet provient essentiellement de l'infime part de leurs travaux d'étude qu'ils acceptent de partager. Leurs algorithmes, d'une importance vitale, constituent de véritables secrets industriels. Ils mènent leur activité à l'abri des regards.

Je ne qualifierais pas pour l'instant les algorithmes de Facebook ou de Google d'ADM politiques, car je n'ai aucune preuve que ces entreprises utilisent leurs réseaux dans l'optique de nuire<sup>2</sup>. Les risques d'abus n'en restent pas moins immenses. C'est au cœur du code informatique, et derrière d'imposants pare-feu, que les choses se jouent. Et comme nous allons le voir, ces technologies possèdent aussi la capacité de nous ranger chacun dans une petite niche politique bien confortable.

\*\*\*

Vers la fin du printemps 2012, l'ex-gouverneur du Massachusetts, Mitt Romney, s'était assuré l'investiture républicaine. L'étape suivante consistait à former un trésor de guerre en vue de la confrontation électorale avec le président Obama. Le 17 mai, il se rendit donc à Boca Raton, en Floride, pour une collecte de fonds organisée au luxueux domicile de Marc Leder, investisseur en capitaux privés. Leder avait déjà versé 225 000 dollars au comité de soutien pro-Romney « Restore Our Future » (« Reconstruire

notre avenir ») et donné 63 330 dollars de plus au comité « Romney Victory » (« Victoire pour Romney »). Il avait réuni tout un tas d'amis fortunés, évoluant en majorité dans le domaine de la finance et de l'immobilier, afin de rencontrer le candidat. Naturellement, un traiteur avait été engagé pour l'occasion.

Romney pouvait raisonnablement supposer qu'il entrait là dans un cercle fermé, dont les membres partageaient en grande partie les opinions de Marc Leder. S'il s'était agi d'un discours télévisé, il aurait bien pris soin de ne pas froisser de potentiels électeurs républicains, qui comptaient dans leurs rangs aussi bien des chrétiens évangéliques que des financiers de Wall Street, des exilés cubains ou des mères au foyer. L'une des raisons pour lesquelles la plupart des discours politiques se révèlent ennuyeux (et ceux de Romney, comme ses partisans s'en plaignaient eux-mêmes, l'étaient tout particulièrement) tient au fait qu'ils cherchent à satisfaire tout le monde. Dans le cadre d'une réunion intime chez Marc Leder, un petit groupe d'individus influents pourrait en revanche approcher le vrai Mitt Romney et entendre les convictions réelles du candidat, sans aucun filtre. Ils lui avaient déjà fait des dons conséquents. Une discussion franche et sincère, c'était le minimum attendu en échange de leur investissement.

Se délectant de cette assistance qu'il croyait favorable et entièrement acquise à ses idées, Romney se permit d'observer que 47 % de la population se composait de « profiteurs », vivant de la générosité du gouvernement fédéral. Ces personnes ne voteraient jamais pour lui, expliqua-t-il – d'où l'importance cruciale de convaincre les 53 autres pour cent. Mais il s'avéra que Romney avait mal visé. Les serveurs qui circulaient parmi les convives, proposant boissons et petits fours, n'étaient pas des leurs. Et comme presque tout le monde dans les pays développés, ils possédaient un téléphone équipé d'une caméra. Les remarques dédaigneuses de Romney, filmées par un barman, firent le tour d'Internet. Cet impair lui ôta très probablement toute chance de conquérir la Maison-Blanche.

Le succès de Romney lors du raout de Boca Raton exigeait à la fois un ciblage précis et une totale confidentialité. Il voulait faire figure, pour Marc Leder et ses amis, de candidat idéal. Et au domicile de Leder, il était persuadé d'être en lieu sûr pour l'incarner. Dans un monde utopique, une infinité de ces lieux, parfaitement sûrs et ciblés, permettraient aux politiciens d'adapter leur discours à chaque sous-groupe – sans rien laisser

voir aux autres. Un seul candidat pourrait ainsi en dissimuler beaucoup, et chaque frange de l'électorat verrait uniquement les aspects qu'elle apprécie.

Cette duplicité, ou cette « multiplicité », n'a rien de nouveau en politique. Les politiciens s'efforcent depuis longtemps de présenter un visage différent en fonction des publics, qu'il s'agisse de déguster des *kielbasa* à Milwaukee<sup>3</sup>, de citer la Torah à Brooklyn<sup>4</sup>, ou de s'engager en faveur de l'éthanol de maïs dans l'Iowa<sup>5</sup>. Mais comme Romney l'a découvert, les caméras peuvent désormais les démasquer s'ils exagèrent ces contorsions.

Le marketing client moderne leur ouvre néanmoins de nouvelles voies pour s'adresser à tel ou tel type d'électeurs, et leur dire ce qu'ils veulent entendre. Ces derniers ont alors tendance à prendre l'information pour argent comptant, car elle vient confirmer leurs convictions antérieures – phénomène que les psychologues appellent un « biais de confirmation ». C'est notamment pour cette raison qu'aucun des donateurs invités à la réception de Romney n'avait contesté son assertion selon laquelle près de la moitié des électeurs engloutissaient les aides publiques. Elle ne faisait qu'étayer ce qu'ils croyaient déjà.

Cette alliance de la politique et du marketing client s'est développée au cours des cinquante dernières années, à mesure que les rituels tribaux de la politique américaine, à base de responsables locaux et de longues listes de numéros téléphoniques, laissaient place à la science du marketing. Dans *The Selling of the President*, qui relatait la campagne de Richard Nixon en 1968, le journaliste Joe McGinnis exposa aux lecteurs le travail des membres de l'état-major chargés de « vendre » le candidat à la présidence comme un bien de consommation. En organisant des groupes de discussion, ils avaient pu adapter son discours en fonction des différentes régions et des différentes tranches de population visées.

Mais au fil du temps, les politiciens réclamèrent une démarche plus détaillée, qui permettrait dans l'idéal de s'adresser à chaque électeur de manière personnalisée. Ce souhait donna naissance aux campagnes de publipostage. Reprenant les tactiques de l'industrie de la carte de crédit, les états-majors politiques bâtirent d'énormes bases de données de clients – d'électeurs, en l'occurrence – et les répartirent en divers sous-groupes reflétant leurs valeurs et leurs caractéristiques démographiques. Pour la première fois, des voisins immédiats pouvaient recevoir de la part d'une même personnalité politique des courriers ou des prospectus différents, l'un

faisant serment de protéger l'environnement et l'autre mettant l'accent sur l'ordre public.

Le publipostage constituait les prémices du micro-ciblage. La convergence du Big Data et du marketing client fournit désormais aux politiciens des outils beaucoup plus puissants. Ils peuvent à présent cibler des microgroupes de citoyens, pour solliciter aussi bien leur vote qu'un soutien financier, et appâter chacun à l'aide d'un message méticuleusement ajusté que personne d'autre n'est appelé à voir. Qu'il prenne la forme d'une bannière sur Facebook ou d'un e-mail pour une collecte de fonds, chaque message permet au candidat de vendre discrètement plusieurs versions de lui-même – sans que nul ne sache celle qui prendra les rênes après l'investiture.

\*\*\*

En juillet 2011, plus d'un an avant que le président Obama ne se présente à sa réélection, un expert en données du nom de Rayid Ghani publia une mise à jour sur LinkedIn :

« Nous recrutons des experts en analyse qui veulent changer les choses. La campagne pour la réélection d'Obama renforce son équipe d'analyse pour travailler sur des problèmes d'extraction de données à grande échelle qui produiront un impact majeur. Plusieurs postes sont disponibles, pour tous niveaux d'expérience. Nous cherchons des experts en statistiques, en *machine learning*, en extraction de données, en analyse textuelle et en analyse prédictive, pour traiter de grandes quantités de données et contribuer à la stratégie électorale. »

Ghani, informaticien formé à l'université Carnegie Mellon, allait diriger l'équipe chargée du traitement de données pour la campagne d'Obama. Dans son emploi antérieur, chez Accenture Labs à Chicago, Ghani avait développé des applications afin d'exploiter le Big Data dans le secteur de la consommation, et il espérait bien pouvoir mettre à profit ses compétences dans le domaine politique. Le but pour la campagne d'Obama était de former des tribus d'électeurs partageant les mêmes idées, avec des valeurs et des priorités aussi uniformes que les invités de la réception chez Marc Leder – mais sans les serveurs indiscrets. On pourrait alors leur adresser



des messages ciblés, ceux qui auraient le plus de chances de les orienter vers des objectifs spécifiques – voter, participer à l’organisation ou collecter des fonds, notamment.

L’un des projets de Ghani chez Accenture consistait à modéliser les clients de supermarchés. Une grande chaîne alimentaire avait confié à son équipe une vaste base de données d’achats anonymisés. L’idée était d’explorer ces données afin d’étudier les habitudes de chaque consommateur, et de les répartir ensuite dans plusieurs centaines de « *buckets* » différents. Il y avait les impulsifs, qui prenaient des sucreries aux présentoirs des caisses, et les obsédés de l’alimentation saine prêts à payer trois fois plus cher pour du chou kalé bio. Ces catégories-là étaient évidentes. Mais d’autres étaient plus surprenantes. Ghani et son équipe repéraient par exemple les personnes qui restaient fidèles à une marque, et celles qu’une infime remise suffisait à faire changer. Ces clients « influençables » se voyaient eux aussi rangés dans des *buckets* distincts. L’objectif final était d’établir un plan différent pour chaque client et de le guider à travers le magasin, en le dirigeant vers tous les produits qu’il avait le plus de chances de vouloir acheter.

Hélas pour les entreprises clientes d’Accenture, cette visée ultime dépendait de l’avènement des caddies informatisés, qui ne se sont encore guère imposés et ne s’imposeront peut-être jamais. En dépit de résultats décevants dans les supermarchés, la démarche scientifique de Ghani se révélait néanmoins parfaitement transposable au monde politique. Ces acheteurs volages qui changeaient de marque pour économiser quelques centimes présentaient par exemple un comportement très proche de celui des électeurs indécis. Dans un supermarché, on pouvait estimer ce que coûterait d’inciter chaque client à délaisser une marque de ketchup ou de café au profit d’une autre plus rentable. L’enseigne pouvait alors sélectionner, disons, les 15 % de clients les plus susceptibles d’en changer et leur offrir des coupons de réduction. Un ciblage intelligent s’avérait essentiel. Il n’était certainement pas souhaitable de distribuer des coupons aux acheteurs disposés à payer le prix fort. Autant jeter l’argent par les fenêtres<sup>6</sup>.

Ce type de calcul pouvait-il fonctionner pour les électeurs indécis ? Munis d’une masse énorme de données sur les consommateurs, les caractéristiques démographiques et les pratiques de vote, Ghani et son équipe se mirent en quête de le découvrir. Ils se heurtèrent néanmoins à une

différence cruciale. Dans le cadre du projet élaboré pour les supermarchés, l'ensemble des données disponibles se rattachait précisément aux achats en magasin, permettant d'étudier les tendances de consommation afin de prédire (et d'influencer) ce que les gens allaient acheter. En politique, en revanche, on n'avait accès qu'à très peu de données pertinentes. Pour chacune des deux campagnes, les équipes d'analyse avaient besoin de données supplétives, qui nécessitaient des recherches.

Elles interrogèrent tout d'abord de manière très approfondie plusieurs milliers de personnes. Ces gens se classaient dans différents groupes. Certains se souciaient du système éducatif ou des droits des homosexuels, tandis que d'autres s'inquiétaient des questions de sécurité sociale ou de l'impact de la fracturation hydraulique sur les nappes phréatiques. Certains apportaient un soutien inconditionnel au président. D'autres hésitaient. Un bon nombre d'entre eux l'appréciaient mais ne prenaient généralement pas la peine d'aller voter. Certains – et c'était là un élément vital – se montraient prêts à contribuer financièrement à la campagne d'Obama.

Une fois que l'équipe de Ghani eut bien appréhendé ce petit groupe d'électeurs, leurs désirs, leurs craintes, et ce qu'il fallait mettre en œuvre pour modifier leur comportement, le défi suivant fut de trouver des millions d'autres électeurs (et de donateurs) qui leur ressemblaient. Cette tâche supposait d'éplucher les données de consommation et les caractéristiques démographiques des votants interrogés, et de dresser leurs profils mathématiques. Il suffisait ensuite de parcourir les bases de données nationales, de trouver des individus aux profils similaires, et de les placer dans les mêmes *buckets*.

L'équipe de campagne put alors cibler chaque groupe en lui adressant des publicités, sur Facebook notamment ou les sites de médias qu'il fréquentait, afin de vérifier s'il réagissait comme prévu. Elle effectua le même type de test A/B qu'utilise Google pour déterminer la nuance de bleu qui génère le plus de clics sur un bouton. En essayant différentes approches, elle découvrit par exemple que les e-mails avec uniquement « Hey ! » comme objet ennuyaient les gens, mais mobilisaient aussi davantage et suscitaient parfois plus de dons. Via des milliers de tests et d'ajustements, l'équipe parvint au bout du compte à prendre la mesure de son public – dont un contingent ô combien important de 15 millions d'électeurs indécis.

Grâce à ce processus, chaque campagne a élaboré divers profils d'électeurs américains. Chaque profil contenait de nombreux scores, qui

évaluaient non seulement leur valeur en tant qu'électeur, bénévole et donateur potentiel, mais reflétaient aussi leurs positions sur différents sujets. Tel électeur pouvait afficher un score élevé sur les questions d'environnement et un score faible sur la sécurité nationale ou le commerce mondial. Ces profils politiques ressemblent beaucoup à ceux que les entreprises d'Internet, comme Amazon et Netflix, utilisent pour gérer leurs dizaines de millions de clients. Leurs moteurs d'évaluation produisent des analyses coûts-bénéfices de façon quasi continue, afin de maximiser le chiffre d'affaires par client.

Quatre ans plus tard, la campagne d'Hillary Clinton s'est appuyée sur la méthodologie mise en place par l'équipe d'Obama. Elle a engagé une start-up spécialisée dans le microciblage – The Groundwork –, financée par le président du conseil d'administration de Google, Eric Schmidt, et dirigée par Michael Slaby, directeur technologique de la campagne d'Obama en 2012. L'objectif, d'après un article publié dans *Quartz*, consistait à bâtir une version politique des systèmes de traitement de données développés par des entreprises comme Salesforce pour gérer leurs millions de clients.

La collecte de données nouvelles et pertinentes éveille comme on peut l'imaginer un intense appétit. Et certaines des méthodes employées s'avèrent détestables, sans parler même de leur caractère intrusif. Fin 2015, le *Guardian* a rapporté qu'une société d'analyse de données politiques, Cambridge Analytica, avait rémunéré des universitaires au Royaume-Uni pour récolter les profils Facebook d'électeurs américains, comprenant les caractéristiques démographiques et l'enregistrement des *likes* de chaque utilisateur. Ces informations leur permirent d'établir l'analyse psychographique de plus de 40 millions d'électeurs, en les classant selon le fameux modèle des cinq grands traits de personnalité, les Big Five, que l'on retrouve ici : l'ouverture d'esprit, la conscienciosité, l'extraversion, l'agréabilité et le névrosisme. Divers groupes collaborant à la campagne présidentielle de Ted Cruz utilisèrent alors ces études pour concevoir des publicités télévisées ciblées en fonction des différents types d'électeurs, et les diffusèrent au milieu des émissions qu'ils étaient le plus susceptibles de regarder. Quand la Republican Jewish Coalition (« Coalition juive républicaine ») s'était réunie en mai 2015 au Venetian, un hôtel-casino de Las Vegas, les équipes de Cruz avaient lancé par exemple une série de publicités en ligne visibles uniquement à l'intérieur du complexe hôtelier,

qui mettaient l'accent sur l'attachement du candidat à l'égard d'Israël et de sa sécurité.

Il convient ici de noter que ces opérations de ciblage n'ont pas toutes fait la preuve de leur efficacité. Certaines frisent sans nul doute la supercherie. Après tout, les professionnels du microciblage se promeuvent eux-mêmes auprès des équipes de campagne et des comités de soutien, qui possèdent des millions de dollars à dépenser. Ils leur font miroiter la qualité inestimable de leurs bases de données et la précision de leur ciblage, autant de promesses dont beaucoup sont forcément exagérées. Les politiciens ne se contentent donc pas d'émettre des promesses discutables ; ils s'en abreuvent aussi (pour un coût exorbitant). Comme l'équipe d'Obama l'a démontré, certaines de ces méthodes se révèlent pourtant fructueuses. Experts en données dignes de ce nom ou simples bonimenteurs, toute l'industrie s'intéresse par conséquent de très près au corps électoral.

Dans le domaine politique, les praticiens du microciblage font toutefois face à des contraintes spécifiques, qui rendent leur travail beaucoup plus complexe. La valeur de chaque électeur augmente ou diminue par exemple selon la probabilité que son État sera ou non disputé. Un électeur indécis habitant un « État charnière<sup>7</sup> » comme la Floride, l'Ohio ou le Nevada s'avère extrêmement précieux. Mais si les sondages montrent que cet État penche résolument en faveur des démocrates ou des républicains, la valeur de l'électeur en question s'effondre, et l'on réoriente alors très vite le budget marketing vers d'autres votants dont la valeur est au contraire en hausse<sup>8</sup>.

De ce point de vue-là, l'électorat peut pratiquement s'envisager de la même manière que les marchés financiers. Les informations qui circulent font monter ou baisser la valeur des électeurs, comme celle des investissements. Dans ces nouveaux marchés politiques, nous représentons chacun une action cotée, dont le cours ne cesse de fluctuer. Et chaque équipe de campagne doit décider si elle souhaite investir sur nous, et comment. Si nous en valons la peine, elle doit ensuite non seulement choisir les informations qui nous seront fournies, mais aussi dans quelles quantités et par quels moyens elle entend les délivrer.

Les responsables de campagne effectuent depuis des décennies des calculs similaires, à l'échelle macroscopique, pour planifier leurs dépenses télévisuelles. À mesure que les sondages évoluent, ils peuvent ainsi réduire la diffusion de publicités à Pittsburgh et reporter ces ressources budgétaires

sur Tampa ou Las Vegas. Avec le microciblage, cependant, on ne se concentre plus sur une région entière, mais sur l'individu. Plus important encore, c'est à lui seul qu'on donne à voir telle ou telle version personnalisée du politicien.

Les campagnes usent d'analyses semblables pour identifier de potentiels donateurs et optimiser chacun d'entre eux. Les choses se compliquent ici, car nombre de donateurs réalisent eux-mêmes leurs propres calculs. Ils cherchent le meilleur rendement. S'ils versent d'un coup la contribution maximale, ils savent pertinemment que les équipes de collecte jugeront qu'elles n'ont plus rien à en tirer et qu'ils ne présentent plus d'intérêt. Mais refuser de donner quoi que ce soit les rendra tout aussi inintéressants. Beaucoup distribuent donc leur argent au compte-gouttes, selon qu'ils approuvent ou pas les messages entendus. La gestion d'un candidat s'apparente à leurs yeux au dressage d'un chien à l'aide de friandises. Cet effet est d'autant plus puissant pour ceux qui contribuent aux « super PAC », ces comités d'action politique qui ne fixent aucune limite quant au montant des dons.

Les équipes de campagne connaissent bien entendu parfaitement cette tactique. Le microciblage leur permet d'envoyer à chaque donateur les informations qui ont les plus grandes chances de leur faire déboursier davantage d'argent. Et ces messages varient d'un donateur à l'autre.

\*\*\*

Les tactiques de ce genre ne se limitent pas aux campagnes électorales. Elles contaminent notre vie civique, lobbyistes et groupes de pression utilisant désormais ces méthodes de ciblage pour mener à bien leur sale besogne. En 2015, le Center for Medical Progress, association antiavortement, a publié des vidéos d'un prétendu fœtus avorté dans une clinique du planning familial. Ces vidéos affirmaient que les médecins du planning vendaient des organes de nouveau-nés pour la recherche ; elles déclenchèrent une vague de protestation, ainsi qu'une offensive des républicains visant à priver l'organisme de ses financements.

Des enquêtes montrèrent plus tard que le film avait été trafiqué : le soi-disant fœtus était en réalité la photographie d'un bébé mort-né accouché dans une région rurale de Pennsylvanie. Et le planning familial ne vendait pas de tissus prélevés sur des fœtus. Le Center for Medical Progress a fini

par admettre que sa vidéo contenait de fausses informations. Son attrait auprès du grand public en a été affaibli. Mais grâce au microciblage, les militants antiavortement ont pu continuer à élargir leur audience, en dépit d'un postulat fallacieux, et s'en sont servis pour lever des fonds afin de combattre le planning familial.

Si cette campagne fut exposée à la vue de tous, des centaines d'autres opèrent dans l'ombre, en s'adressant individuellement aux électeurs. Ces campagnes plus discrètes sont tout aussi trompeuses et encore moins contrôlables. Elles lancent des bombes idéologiques, auxquelles les politiciens feront à peine allusion dans leurs déclarations officielles. Comme l'explique Zeynep Tufekci, techno-sociologue et professeur à l'université de Caroline du Nord, les groupes partisans identifient les électeurs vulnérables et en font la cible de campagnes alarmistes, attisant leurs craintes quant à la sécurité de leurs enfants ou la montée de l'immigration clandestine. Mais ils peuvent dans le même temps dissimuler ces publicités aux électeurs qui risquent d'être rebutés (voire écœurés) par de tels messages.

Un microciblage efficace explique en partie qu'en 2015, selon un sondage, plus de 43 % des républicains croyaient encore que le président Obama était musulman. 20 % des Américains pensaient aussi qu'il était né en dehors des États-Unis, et qu'il occupait en conséquence la fonction suprême de façon illégitime. (Bien que les démocrates puissent également répandre par microciblage leur propre désinformation, rien de ce qui a émergé n'égale l'ampleur des campagnes anti-Obama.)

Malgré le développement du microciblage, les campagnes politiques consacrent encore en moyenne 75 % de leurs achats médias à la télévision. On pourrait imaginer que cela produise un effet compensateur, et c'est effectivement le cas. Là où le microciblage agit dans l'ombre, la télévision diffuse un message plus large et vérifiable. Mais elle se tourne aujourd'hui elle-même vers une publicité personnalisée. De nouvelles agences publicitaires, comme Simulmedia, basée à New York, répartissent les téléspectateurs en différentes catégories comportementales pour permettre aux annonceurs de cibler des personnes partageant les mêmes idées, qu'il s'agisse de chasseurs, de pacifistes ou d'acheteurs de véhicules de loisir type SUV, gros comme des chars d'assaut. À mesure que la télévision et les autres médias s'orientent vers le profilage de leurs publics, les possibilités de microciblage politique s'accroissent.

Il deviendra plus difficile, dans ces conditions, d'accéder aux messages politiques que nos voisins peuvent voir – et par suite de comprendre pourquoi ils croient ce qu'ils croient, souvent avec passion. Un journaliste même curieux aura du mal à les dénicher. Consulter simplement le site web du candidat ne suffira pas, car il procède lui aussi au profilage et au ciblage automatique de chaque visiteur, prenant tout en compte, depuis son code postal jusqu'aux liens cliqués sur la page et même aux photos qu'il semble visionner. La création de dizaines de « faux » profils sera tout aussi vaine, car les systèmes associent chaque électeur réel aux connaissances approfondies accumulées à son sujet : historiques d'achats, adresses, numéros de téléphone, participations aux scrutins, voire numéro de Sécurité sociale et profil Facebook. Pour convaincre le système de son authenticité, chaque faux profil devrait s'accompagner de sa propre somme de données. En inventer un de toutes pièces réclamerait dans le cadre d'un projet de recherche beaucoup trop de travail (et dans le pire des cas pourrait valoir à l'enquêteur d'être accusé de fraude).

Ces campagnes souterraines entraînent un dangereux déséquilibre. Les marketeurs politiques accumulent à notre sujet des dossiers détaillés, nous exposent à un mince filet d'informations et mesurent comment nous y réagissons. Mais ils nous maintiennent dans l'ignorance de ce que nos voisins entendent. Cette tactique ressemble à celle couramment utilisée par les négociateurs commerciaux. Ils traitent séparément avec les diverses parties, afin qu'aucune ne sache ce qui est dit à l'autre. Cette asymétrie de l'information empêche leurs interlocuteurs d'unir leurs forces – ce qui est précisément le but d'un gouvernement démocratique.

Cette science en plein essor du microciblage, avec ses profils et ses prédictions, entre impeccablement dans notre sinistre collection d'ADM. Elle agit sur une vaste échelle, se montre opaque et ne rend compte de rien. Elle protège les politiciens et les incite à endosser en fonction des gens de multiples visages.

La notation individuelle des électeurs menace également la démocratie : elle donne de l'importance à une minorité d'entre eux et ne laisse guère plus aux autres qu'un rôle de second plan. Au vu des modèles utilisés lors des élections présidentielles, on a effectivement l'impression que les États-Unis se sont rétrécis. À l'heure où j'écris ces lignes, la totalité de l'électorat qui importe vraiment habite une poignée de comtés de Floride, de l'Ohio, du Nevada et de quelques autres « États charnières ». Au sein de ces comtés

réside un petit nombre d'électeurs dont les opinions pèsent dans la balance. Alors qu'une grande partie des ADM que nous avons étudiées, des publicités prédatrices aux modèles manipulateurs, punissent essentiellement les classes sociales en difficulté, on peut souligner ici que le microciblage politique porte atteinte aux électeurs de tous les niveaux économiques. De Manhattan à San Francisco, riches comme pauvres se retrouvent marginalisés (même si les vrais nantis peuvent bien entendu largement y pallier par le biais de leurs dons).

En tout état de cause, le système politique tout entier – l'argent dépensé, l'attention accordée, les flatteries prodiguées – se tourne vers les électeurs ciblés telle une fleur attirée par le soleil. Les autres sont pratiquement ignorés (sauf pour les collectes de fonds). Les programmes ont déjà prédit notre comportement électoral, et aucune tentative de le changer n'est plus de mise<sup>9</sup>.

Ce mécanisme génère une boucle de rétroaction malfaisante. Les électeurs méprisés risquent de plus en plus de se sentir désabusés. Les privilégiés, eux, savent mener le jeu et en connaissent tous les détails, tandis que l'immense majorité des consommateurs n'en perçoit que des fragments, adaptés au marché.

Il existe de fait une asymétrie supplémentaire. Les gens que l'on s'attend à voir voter mais qui, pour une raison ou pour une autre, manquent un scrutin se trouvent la fois suivante couverts d'attentions. Ils semblent toujours receler un potentiel électoral élevé. Ceux au contraire dont on n'espère pas le vote sont largement négligés. Les systèmes recherchent les voix les moins chères à convertir, celles qui rapporteront le plus pour chaque dollar dépensé. Et les abstentionnistes paraissent souvent trop coûteux. Cette dynamique pousse une certaine classe de gens à rester actifs, et laisse définitivement les autres de côté.

Comme c'est souvent le cas avec les ADM, les mêmes modèles qui sont préjudiciables pourraient être mis à profit. Au lieu de cibler les gens pour les manipuler, ils pourraient leur venir en aide. Lors d'une élection municipale, par exemple, une campagne de microciblage pourrait détecter, via les messages qu'ils envoient, les électeurs mécontents du coût excessif des loyers. Mais si le candidat est informé de leurs griefs, pourquoi ne pas exploiter alors la même technologie pour identifier ceux qui auraient le plus intérêt à bénéficier d'un logement abordable, et les aider à l'obtenir ?



En matière de communication politique, comme avec la plupart des ADM, le cœur du problème réside presque toujours dans l'objectif poursuivi. Si l'on choisit d'aider les gens au lieu de profiter d'eux, l'ADM se trouve neutralisée – et peut même se changer en une force bienfaitrice.

---

1. En France on compte, en 2018, 38 millions d'utilisateurs mensuels de Facebook (dont 22 millions actifs quotidiennement), pour 56,4 millions d'internautes (source : Facebook/Digimind).

2. L'édition originale de cet ouvrage a été publiée en 2016, bien avant que les divers scandales impliquant Facebook, dont Cambridge Analytica, ne soient rendus public, et que le fondateur de Facebook, Mark Zuckerberg, soit entendu par la justice. Le réseau social a depuis retiré des centaines de faux comptes affiliés à l'Iran et la Russie, identifiés comme des outils d'influence politique.

3. Type de saucisse originaire de Pologne. La région de Milwaukee compte beaucoup d'habitants d'ascendance polonaise (*NdT*).

4. Brooklyn abrite une importante communauté juive orthodoxe (*NdT*).

5. L'Iowa est le premier producteur d'éthanol des États-Unis (*NdT*).

6. De la même manière, les commerces en ligne sont nettement plus enclins à proposer des réductions aux gens qui ne sont pas encore connectés sur leur site. Une raison supplémentaire d'effacer régulièrement vos cookies (*NdA*).

7. Les « *swing states* », États qui peuvent changer de camp d'un scrutin à l'autre et faire basculer le résultat final (*NdT*).

8. Le système américain est celui du suffrage universel indirect. Chaque État possède des « grands électeurs » qui forment le collège électoral (538 membres), chargé à son tour d'élire l'un ou l'autre des candidats en lice, démocrate ou républicain. Le nombre de grands électeurs est proportionnel à la représentation de chaque État au Congrès et chaque grand électeur est rattaché à un parti politique. Il est élu par suffrage direct lors d'élections dites primaires. À l'exception du Maine et du Nebraska, les États appliquent le principe du « *winner takes it all* » (« tout au vainqueur ») : l'État donne la totalité de ses voix au candidat présidentiel qui obtient la majorité des votes lors de la primaire (dont celles initialement attribuées au candidat minoritaire).

9. À l'échelon fédéral, l'abolition du système des grands électeurs (le collège électoral) pourrait grandement atténuer ce problème. D'un État à l'autre, c'est le principe mathématique du « tout au vainqueur » qui confère tant de pouvoir à une petite poignée d'électeurs. Sur le plan politique, nous avons d'une certaine manière comme en économie 1 % de privilégiés. Et l'argent des privilégiés économiques finance le microciblage, qui permet de s'assurer les voix des privilégiés politiques. Si l'on supprimait le collège électoral, chaque voix aurait en revanche exactement le même poids. Ce serait un pas vers la démocratie (*NdA*).

# CONCLUSION

1001EBooks

Dans cette traversée virtuelle de l'existence humaine, nous sommes allés à l'école et à l'université, au tribunal et dans l'entreprise, et nous avons même pénétré dans l'isoloir. Nous avons constaté, tout au long de ce parcours, les destructions provoquées par les ADM. Elles promettent efficience et équité mais pervertissent en réalité l'enseignement supérieur, aggravent l'endettement, favorisent une incarcération massive, oppriment en toutes occasions les plus pauvres, et sapent la démocratie. La réaction logique semble être de les neutraliser, une par une.

Le problème, c'est qu'elles se nourrissent les unes des autres. Les gens pauvres ont tendance à présenter un mauvais score de crédit et à vivre dans des quartiers à forte criminalité, entourés d'autres individus pauvres. Une fois ces données digérées, le sordide univers des ADM les bombarde de publicités prédatrices pour des prêts hypothécaires de type *subprime* ou des écoles à but lucratif. Il déploie plus de policiers pour les arrêter et, lorsqu'ils sont condamnés, leur inflige des peines de prison plus longues. Ces données alimentent d'autres ADM, qui attribuent aux mêmes personnes un niveau de risque élevé ou les classent comme des cibles faciles, puis les empêchent de postuler à un emploi, et font grimper dans le même temps les taux de leur emprunt immobilier, de leur prêt voiture, et les tarifs de toutes les assurances imaginables. Cette modélisation abaisse d'autant plus leur cote de solvabilité, générant ainsi rien moins qu'une spirale destructrice. Dans un monde peuplé d'ADM, être pauvre devient de plus en plus dangereux et de plus en plus coûteux.

Les mêmes ADM qui maltraitent les pauvres rangent également les gens aisés dans leurs propres « silos » marketing. Elles les envoient en vacances

dans les tropiques à Aruba et les inscrivent sur la liste d'attente de la prestigieuse Wharton School. Beaucoup d'entre eux peuvent avoir l'impression que le monde devient grâce à elles plus intelligent et plus facile à vivre. Les modèles leur indiquent où acheter du prosciutto et du chianti au meilleur prix, leur conseillent un bon film sur Amazon Prime, ou les guident par GPS jusqu'à un café situé dans un quartier naguère « douteux ». La discrétion et la personnalisation de ce ciblage empêchent les privilégiés de s'apercevoir que les mêmes modèles détruisent aussi d'autres vies, parfois à quelques rues de là.

La devise des États-Unis, « *E Pluribus Unum* », signifie « De plusieurs, ne faire qu'un ». Mais les ADM inversent l'équation. Œuvrant dans l'ombre, elles scindent l'unité en multitude, dissimulant les torts qu'elles causent à nos voisins, proches ou lointains. Et ces torts sont légion. Ils se manifestent lorsqu'une mère célibataire ne parvient pas à organiser assez vite la garde de son enfant pour s'adapter à son planning de travail, ou lorsqu'un jeune en difficulté se voit refuser un emploi payé à l'heure sur la base d'un test de personnalité. On les observe quand un adolescent issu d'une minorité défavorisée est interpellé, brutalisé et mis en garde par la police locale, ou quand un pompiste habitant une commune pauvre doit payer une prime d'assurance plus élevée que les autres. C'est une guerre silencieuse qui frappe d'abord les pauvres mais qui accable également la classe moyenne. La plupart de ses victimes n'ont pas les armes nécessaires – pouvoir économique, accès aux avocats, organisations politiques bien financées – pour livrer bataille. D'où ces atteintes généralisées, qui passent trop souvent pour inévitables.

Il ne faut pas attendre de l'économie actuelle, le libre marché, qu'elle remédie d'elle-même à ces maux. Et pour comprendre pourquoi, comparons donc les ADM à un autre fléau auquel notre société se trouve confrontée : celui de l'homophobie.

En septembre 1996, deux mois avant sa réélection, le président Bill Clinton signait le Defense of Marriage Act. Définissant le mariage comme l'union d'un homme et d'une femme, ce texte lui laissait espérer un soutien plus fort dans les régions conservatrices des États clés, tels que l'Ohio et la Floride. À peine une semaine plus tard, le géant technologique IBM annonça qu'il offrirait une couverture médicale aux partenaires de même sexe de ses employés. On pourrait se demander ce qui avait incité IBM, pilier de l'establishment industriel, à ouvrir cette porte et à s'exposer ainsi à

la controverse, au moment où un président américain supposément progressiste prenait la direction opposée.

La réponse est liée aux profits financiers. En 1996, la ruée vers l'or d'Internet venait tout juste de démarrer et IBM luttait avec Oracle, Microsoft, Hewlett-Packard, ainsi que plusieurs start-up comme Amazon ou Yahoo, pour recruter les meilleurs cerveaux. La plupart de ces autres entreprises proposaient déjà des avantages sociaux aux partenaires de même sexe, et attiraient chez elles les talents gays et lesbiens. IBM ne pouvait se permettre de passer à côté. « En termes de compétitivité commerciale, il s'agissait pour nous d'une démarche judicieuse », indiqua à l'époque un porte-parole de la firme interrogé par *BusinessWeek*.

Si l'on envisage les politiques de ressources humaines d'IBM et d'autres entreprises sous l'aspect d'algorithmes, on constate que leur code a été porteur de discriminations pendant des décennies. La décision de procurer à tous des avantages identiques les a poussées vers un peu plus d'équité. Les communautés gays et lesbiennes ont depuis connu des progrès impressionnants dans nombre de domaines, mais bien entendu les inégalités subsistent. Aux États-Unis, beaucoup de gays, de lesbiennes et de transgenres sont encore victimes de préjugés, d'actes violents, et continuent de subir les assauts des ADM. C'est particulièrement vrai au sein des minorités et des populations pauvres. Pour autant, à l'heure où j'écris ces lignes, c'est un homosexuel – Tim Cook – qui est PDG d'Apple, la plus grosse capitalisation boursière au monde. Et la Constitution américaine lui donne le droit, s'il le souhaite, d'épouser un autre homme.

Puisque l'on a vu à présent que les entreprises possèdent la capacité de corriger une injustice contenue dans leurs algorithmes de recrutement, ne pourraient-elles pas apporter des ajustements similaires aux modèles mathématiques qui provoquent tant de ravages dans notre société – à savoir les ADM ?

Il existe hélas une différence majeure. Les droits des homosexuels ont à bien des égards bénéficié des forces du marché. La communauté gay et lesbienne abritait un vivier de talents hautement diplômés et de plus en plus audibles, dans lequel les entreprises avaient hâte de puiser. Elles ont donc optimisé leurs modèles afin de les séduire, mais en restant focalisées sur leurs profits. Dans la plupart des cas, l'équité ne fut qu'un effet secondaire. Aux quatre coins du pays, des entreprises ont commencé au même moment à s'intéresser aux consommateurs LGBT aisés, en leur proposant des

croisières, des « happy hours » et des émissions TV thématiques. Si cette intégration a sans aucun doute suscité la grogne de certains îlots d'intolérants, elle a également beaucoup rapporté.

Le désamorçage d'une ADM n'offre pas toujours un gain aussi évident. Bien que davantage de justice et d'équité soit certainement profitable à la société dans son ensemble, les entreprises ne sont pas en mesure individuellement d'en récolter les fruits. Pour la plupart, les ADM sont à vrai dire extrêmement efficaces. Des business models tels que celui des universités à but lucratif ou des prêts sur salaire reposent intégralement sur elles. Et lorsqu'un logiciel réussit à cibler des personnes suffisamment désemparées pour payer 18 % d'intérêt par mois, ceux qui engrangent les profits se disent naturellement que ces modèles fonctionnent tout à fait bien.

Les victimes sont évidemment d'un autre avis. Mais la majorité d'entre elles – les travailleurs payés à l'heure et les chômeurs, les gens qui traînent toute leur vie de mauvais scores de crédit – sont pauvres. Les personnes incarcérées n'ont aucun pouvoir. Et dans notre société, où l'argent permet de s'acheter une influence, ces victimes d'ADM sont pratiquement privées de voix. Leur vote ne compte pas. Les pauvres sont en réalité trop souvent tenus pour responsables de leur propre misère, de leurs mauvaises écoles et de la criminalité qui gangrène leurs quartiers. Bien peu de politiciens se donnent par conséquent la peine d'établir des stratégies pour la combattre, de l'avis général, les maux liés à la pauvreté s'apparentent davantage à une maladie, et il s'agit donc – c'est du moins la rhétorique – de l'isoler et d'empêcher qu'elle se propage à la classe moyenne. Une réflexion s'impose, dans notre époque, sur cette manière de désigner des fautifs et les modèles qui alimentent ce cercle vicieux.

Mais les pauvres sont loin d'être les uniques victimes des ADM. Des modèles malveillants peuvent, comme on l'a vu, mettre sur liste noire des candidats à l'embauche parfaitement qualifiés et amputer la rémunération de travailleurs qui ne respectent pas « l'idéal de santé » d'une entreprise. Ces ADM frappent la classe moyenne aussi durement que n'importe qui d'autre. Les riches eux-mêmes se trouvent soumis au microciblage des modèles politiques, et se démènent aussi fort que tout le monde pour satisfaire l'impitoyable ADM qui régit les admissions universitaires et pollue l'enseignement supérieur.

Mais notons-le, nous n'en sommes qu'aux débuts. Les prêteurs sur salaire et leurs semblables s'attaquent naturellement d'abord aux pauvres et

aux immigrés. Ce sont les cibles les plus faciles. Ces derniers ont un accès réduit à l'information, et beaucoup sont aux abois. Mais les ADM, génératrices de fabuleux profits, ne se cantonneront sans doute pas longtemps aux strates inférieures. Ce n'est pas ainsi que les marchés fonctionnent. Elles vont se développer et se propager, en quête de nouvelles opportunités. On le constate déjà, à l'heure où les banques traditionnelles investissent par exemple dans des plateformes de prêt entre particuliers comme Lending Club. Nous sommes tous, en résumé, dans le viseur des ADM. Et tant que rien ne sera fait pour les arrêter, elles continueront de se multiplier et de répandre l'injustice.

L'injustice, qu'elle résulte de la cupidité ou des préjugés, a toujours existé. Et l'on pourrait d'ailleurs soutenir que les ADM n'agissent pas plus mal aujourd'hui que les humains ne le faisaient hier. Il n'était pas rare en effet qu'un agent de prêt ou un recruteur écarte systématiquement des groupes raciaux tout entiers lors d'une demande d'emprunt ou une offre d'emploi, si ce n'est même toutes les personnes du même sexe. Aux yeux de beaucoup de gens, même les pires modèles mathématiques ne sont pas aussi malfaisants.

Quoique souvent viciée, la prise de décision opérée par un humain possède une vertu essentielle. Elle peut évoluer. À mesure qu'ils apprennent et qu'ils s'adaptent, les êtres humains progressent, et leurs processus avec. Les systèmes automatisés, en revanche, restent à jamais figés tant que des ingénieurs ne se penchent pas dessus pour les modifier. Si un modèle d'inscription à l'université basé sur le Big Data avait été mis en place au début des années 1960, peu de femmes accéderaient encore de nos jours à l'enseignement supérieur, car les données d'entraînement se seraient en grande partie fondées sur la réussite des hommes. Si les musées avaient à la même époque codifié les idées prépondérantes en matière de qualité artistique, nous n'admirerions encore pour l'essentiel que le travail d'hommes blancs – financés par de riches mécènes pour créer leurs œuvres. Et l'équipe de football de l'université de l'Alabama serait encore, cela va sans dire, exclusivement blanche.

Les processus reposant sur le Big Data n'inventent pas le futur, ils codifient le passé. Il faut pour cela une imagination morale que les humains sont seuls en capacité de fournir. Nous devons expressément intégrer à nos algorithmes de meilleures valeurs, en créant pour le Big Data des modèles

conformes à nos visées éthiques. Ce qui supposera parfois de placer l'équité au-dessus du profit.

Notre société affronte en un sens une nouvelle révolution industrielle. Et nous pouvons tirer de la précédente certaines leçons. Le tournant du <sup>xx</sup>e siècle fut marqué par de grandes avancées. On a pu s'éclairer à l'électricité et se chauffer au charbon. Les chemins de fer modernes apportaient de l'autre bout du pays viandes, légumes et conserves. Pour beaucoup, l'existence n'allait qu'en s'améliorant.

Mais ce progrès cachait un revers épouvantable. Il se nourrissait de la terrible exploitation des travailleurs, parmi lesquels beaucoup d'enfants. En l'absence de toute réglementation, tant sur le plan sanitaire que sur celui de la sécurité, les mines de charbon constituaient des pièges mortels. Au cours de la seule année 1907, 3 242 mineurs y perdirent la vie. Les ouvriers des abattoirs travaillaient douze à quinze heures par jour dans des conditions insalubres, et mettaient souvent en circulation des produits frelatés. L'entreprise Armour and Co. a ainsi expédié par tonnes entières des boîtes de bœuf avarié aux troupes américaines, en recouvrant la viande d'une couche d'acide borique pour en masquer l'odeur. Des monopolistes rapaces jouissaient de leur côté d'une mainmise totale sur les chemins de fer, les compagnies énergétiques et les services publics, et faisaient grimper les tarifs appliqués aux clients, ce qui revenait à prélever une taxe sur l'économie nationale.

Le libre marché était de toute évidence incapable de réguler ses excès. Lorsque des journalistes comme Ida Tarbell et Upton Sinclair eurent dénoncé entre autres ces problèmes, le gouvernement décida donc d'intervenir. Il instaura en matière d'alimentation des protocoles de sécurité et des inspections sanitaires, et le travail des enfants fut interdit. Avec l'essor des syndicats, et l'adoption de lois défendant leur activité, la société s'orienta vers la journée de huit heures et les week-ends chômés. Ces nouvelles normes protégeaient les entreprises qui refusaient d'exploiter leurs employés ou de commercialiser des produits altérés, car leurs concurrents étaient tenus de suivre les mêmes règles. Et, même si elles avaient sans doute augmenté les coûts de fonctionnement, la société dans son ensemble en a aussi tiré bénéfice. Peu d'entre nous auraient envie aujourd'hui de revenir à l'époque d'avant.

\*\*\*

Par où devrait-on commencer, si nous voulions réguler ces modèles mathématiques qui dirigent de plus en plus nos vies ? Ma suggestion serait de partir des modélisateurs eux-mêmes. À l'instar des médecins, les experts en données devraient prêter une sorte de serment d'Hippocrate, axé sur les mauvais usages et les mauvaises interprétations possibles de leurs modèles. À la suite du krach de 2008, deux ingénieurs financiers, Emanuel Derman et Paul Wilmott, ont rédigé un serment de ce type, énonçant les préceptes suivants :

*Je me souviendrai que je n'ai pas créé le monde, et qu'il ne se conforme pas à mes équations.*

*J'utiliserai sans crainte mes modèles pour estimer des valeurs, mais ne me laisserai pas démesurément impressionner par les mathématiques.*

*Je ne sacrifierai jamais la réalité à l'élégance, sans expliquer pourquoi je l'ai fait.*

*Je ne donnerai pas non plus aux utilisateurs de mon modèle de fausses certitudes quant à sa précision. Je soulignerai au contraire explicitement ses hypothèses et ses omissions.*

*Je comprends que mon travail peut avoir d'énormes effets sur la société et l'économie, dont beaucoup dépassent mon entendement.*

Nous avons là une bonne base philosophique. Mais les nobles valeurs et l'autorégulation ne réfrènent que les gens scrupuleux. Le serment d'Hippocrate ignore, qui plus est, la pression que les experts en données subissent au quotidien, quand leurs supérieurs exigent de leur part des réponses spécifiques. Pour éliminer les ADM, notre corporation ne peut se contenter d'établir de bonnes pratiques. Nos lois doivent elles aussi changer. Et pour y arriver, nous devons réévaluer nos critères de réussite.

À l'heure actuelle, la réussite d'un modèle se mesure souvent en termes de profits, d'efficience ou de taux de défaut. Il s'agit presque toujours de quantités dénombrables. Mais que faudrait-il réellement comptabiliser ? Prenons l'exemple suivant. Quand les gens saisissent une requête sur un moteur de recherche pour obtenir des informations en matière de bons alimentaires, ils voient fréquemment apparaître des annonces d'intermédiaires comme FindFamilyResources, situé à Tempe, dans l'Arizona. Les sites de ce genre présentent un aspect officiel et proposent des liens vers d'authentiques formulaires administratifs. Mais ils recueillent



aussi noms et adresses e-mail au profit de publicitaires prédateurs, et notamment des universités à but lucratif. En fournissant aux gens une prestation superflue, ils engrangent des commissions qui rémunèrent les contacts générés, dont beaucoup reçoivent ensuite très vite des offres ciblées pour des services qu'ils peuvent difficilement se payer.

La transaction est-elle fructueuse ? Tout dépend de ce que l'on compte. Pour Google, le clic sur une publicité rapporte de 25 à 50 *cents*, voire 1 ou 2 dollars. C'est une réussite. Naturellement, le générateur de contacts gagne lui aussi de l'argent. Le système paraît donc fonctionner de manière efficiente. Les rouages du commerce tournent parfaitement.

Du point de vue de la société, en revanche, une simple recherche de services administratifs accroche une énorme cible dans le dos des personnes pauvres, et conduit un certain nombre d'entre elles vers de fausses promesses et des prêts à taux élevés. Même en considérant les choses sous un angle strictement économique, c'est un fardeau pour le système. Le fait même que des gens aient besoin de recourir aux bons alimentaires représente un échec de l'économie de marché. L'État tente de le compenser en puisant dans ses recettes fiscales, dans l'espoir que les bénéficiaires de ces bons seront un jour en mesure de subvenir pleinement à leurs besoins. Mais les agrégateurs de contacts les poussent à d'inutiles transactions, amenant bon nombre d'entre eux à creuser leurs dettes, et les rendant plus dépendants encore de l'aide publique. L'ADM, productrice de revenus pour les moteurs de recherche, les agrégateurs de contacts et les marketeurs, agit aussi comme une sangsue pour l'ensemble de l'économie.

Un système de régulation des ADM aurait pour tâche de mesurer ces coûts cachés, et devrait également prendre en compte tout un ensemble de valeurs non numériques. C'est déjà le cas pour d'autres types de régulations. Bien que les économistes puissent s'efforcer de calculer le coût de la pollution atmosphérique, du ruissellement agricole (écoulement des eaux à la surface du sol) ou de l'extinction de la chouette tachetée, les chiffres ne parviennent jamais à en exprimer la valeur. Il en va souvent de même pour les notions d'équité et d'intérêt général dans les modèles mathématiques. Ce sont des concepts qui n'existent que dans l'esprit des hommes et qui résistent à toute quantification. Et comme ce sont des humains qui s'occupent de concevoir les modèles, ils font rarement l'effort ne serait-ce que d'essayer. Ils considèrent simplement que c'est trop

difficile. Il nous faut pourtant insuffler à ces systèmes des valeurs humaines, même au détriment de l'efficacité.

Un modèle pourrait par exemple être programmé afin de s'assurer que la diversité des appartenances ethniques ou des niveaux de revenus soit bien représentée au sein des différents groupes d'électeurs ou de consommateurs. Ou pour mettre en évidence les cas où les habitants de certaines localités paient certains services deux fois plus cher que la moyenne. Ces approximations paraîtront peut-être grossières, surtout au premier abord, mais elles sont essentielles. Les modèles mathématiques devraient nous servir d'outils, et non pas nous dominer.

Les écarts scolaires, l'incarcération massive et l'apathie électorale représentent aux États-Unis des problèmes majeurs que ni le libre marché ni aucun algorithme mathématique ne sauraient résoudre. Il convient donc en premier lieu de tempérer notre utopie technologique, cet espoir sans borne et injustifié dans ce que les algorithmes et la technologie pourraient accomplir. Avant d'en exiger mieux, nous devons d'abord admettre qu'ils ne peuvent pas tout.

Pour désamorcer les ADM, il nous faut également mesurer leur impact et auditer leurs algorithmes. La première étape, avant de plonger dans le code du logiciel, consiste à effectuer des recherches en considérant l'ADM comme une boîte noire, qui absorbe des données et émet des conclusions : cette personne présente un risque moyennement important de commettre un autre crime ; celle-ci a 73 % de chances de voter républicain ; les performances de cet enseignant-là le classent dans le décile inférieur. En étudiant ces résultats, on pourrait alors reconstituer les hypothèses sous-jacentes et en évaluer l'équité.

Dès le départ, certaines ADM apparaissent parfois très clairement comme des outils primitifs, qui simplifient à outrance une réalité complexe et facilitent la tâche des managers lorsqu'ils souhaitent licencier des groupes d'individus ou offrir des remises à d'autres. Le modèle de la valeur ajoutée utilisé dans les écoles publiques new-yorkaises – celui qui, une année, avait décerné à Tim Clifford la désastreuse note de 6, puis l'année suivante un faramineux 96 – constitue par exemple sur le plan statistique une véritable farce. Si l'on reporte sur un graphique les scores annuels successifs, ils se répartissent de manière presque aussi aléatoire que les atomes d'hydrogène dans une pièce. Beaucoup d'élèves de ces mêmes établissements pourraient en un quart d'heure étudier ces statistiques et en conclure avec certitude que

les scores ne mesurent rien. Année après année, les bons enseignants ont tendance en effet à le rester. À la différence, par exemple, des lanceurs de relèbe dans le base-ball, il leur arrive rarement de connaître coup sur coup des saisons formidables puis catastrophiques. (Et contrairement à eux, leurs performances résistent à l'analyse quantitative.)

Un modèle aussi fruste que celui de la valeur ajoutée est impossible à corriger. L'unique solution consiste en l'occurrence à abandonner ce système injuste. Renonçons, au moins pour les dix à vingt années qui viennent, à concevoir des outils destinés à mesurer l'efficacité d'un enseignant. C'est une chose trop complexe à modéliser, et les seules informations dont on dispose proviennent de données indirectes, grossières et imprécises. Le modèle n'est tout simplement pas encore assez bon pour guider la prise de décisions importantes au sujet de ces personnes auxquelles nous confions l'éducation de nos enfants. C'est une tâche qui réclame de la subtilité et une perception correcte du contexte. Même à l'ère du Big Data, un tel problème ne peut être résolu que par des êtres humains.

Ces analystes humains, qu'il s'agisse du chef d'établissement ou des personnels administratifs, devront bien entendu examiner quantité de données, y compris les résultats des tests passés par les élèves. Des boucles de rétroaction positives, cousines angéliques des boucles de rétroaction néfastes que nous connaissons désormais si bien, devront être intégrées au processus. Une boucle positive procure à l'expert en données (ou au système automatique) des informations utiles permettant d'améliorer le modèle. Dans le cas présent, il s'agira simplement de demander aux enseignants ainsi qu'aux élèves si les évaluations leur paraissent sensées, et s'ils en comprennent et en acceptent les fondements. Sinon, quelles modifications pourrait-on y apporter ? L'unique espoir d'améliorer l'enseignement grâce aux données passe par la création d'un écosystème basé sur des boucles de rétroaction positives. Tant qu'il n'existe pas, l'évaluation conserve un caractère punitif.

Comme les défenseurs des données s'empressent de le souligner, l'esprit humain met en œuvre il est vrai ses propres modèles internes, empreints souvent de préjugés ou d'égoïsme. L'équité de ses conclusions – en l'espèce, des évaluations d'enseignants – doit donc elle aussi être vérifiée. Et ces vérifications devront être conçues et testées avec soin par des êtres humains, avant d'être automatisées. Les mathématiciens pourront, en

parallèle, élaborer des modèles pour aider les enseignants à mesurer eux-mêmes leur efficacité et à s'améliorer.

D'autres audits se révèlent beaucoup plus compliqués. Prenons les modèles de prédiction du récidivisme criminel, que les magistrats de nombreux États consultent avant de prononcer une peine. La technologie étant assez nouvelle, on peut facilement étudier ici un avant et un après. Depuis qu'ils reçoivent une analyse de risque générée par cette ADM, les tendances des juges en matière de sentences ont-elles changé ? On constatera sans doute qu'un certain nombre d'entre eux appliquaient dans leur tête des modèles tout aussi inquiétants bien avant l'arrivée du logiciel, infligeant des sanctions plus sévères aux détenus pauvres et issus des minorités. Dans certains cas, le logiciel aurait vraisemblablement adouci leur verdict. Dans d'autres, non. Avec un volume de données suffisant, des schémas pourront néanmoins se dégager, qui nous permettront d'évaluer la puissance et l'influence de l'ADM.

Si l'on découvre (comme des études l'ont déjà montré) que les modèles de récidivisme systématisent les préjugés et pénalisent les pauvres, il conviendra alors de jeter un œil aux données d'entrée. Elles comporteront quantité de rapprochements fondés sur le principe du « qui se ressemble s'assemble », et prédiront le comportement d'un individu en fonction des gens qu'il connaît, de son emploi et de son score de crédit – autant d'informations irrecevables devant un tribunal. Pour restaurer l'équité, il faudra s'en débarrasser.

« Une minute ! », répondront beaucoup de spécialistes. Allons-nous, au nom de l'équité, sacrifier l'exactitude du modèle ? Devons-nous rendre nos algorithmes moins intelligents ?

Dans certains cas, oui. Si nous voulons être égaux face à la loi, ou bénéficier d'un traitement identique en tant qu'électeurs, nous ne pouvons tolérer que des systèmes nous rangent dans différentes castes et nous considèrent différemment les uns des autres<sup>1</sup>. Des entreprises comme Amazon et Netflix peuvent répartir leurs clients en une multitude de petites catégories et les optimiser tant qu'elles le souhaitent. Mais un tel algorithme ne saurait être vecteur de justice ou de démocratie.

Diverses démarches visant à auditer les algorithmes sont déjà sur pied. À l'université Princeton, par exemple, des chercheurs ont lancé le Web Transparency and Accountability Project (Projet pour la transparence et la responsabilisation du Web). Ils créent des robots logiciels qui se font passer

sur Internet pour des individus de toute nature – riches, pauvres, hommes, femmes, ou personnes atteintes d’une pathologie mentale. En étudiant la manière dont ces robots sont traités, les chercheurs peuvent détecter les partis pris des systèmes automatisés, depuis les moteurs de recherche jusqu’aux sites de recrutement. Des initiatives similaires sont également en train de naître dans d’autres établissements comme Carnegie Mellon ou le prestigieux MIT (Massachusetts Institute of Technology).

Le soutien des universitaires revêt une importance cruciale. Pour mettre au pas les ADM, nous avons en effet besoin de gens dotés des compétences nécessaires pour les créer. Leurs outils de recherche permettent d’en reproduire l’immense portée et d’extraire des quantités de données suffisantes pour révéler les déséquilibres et les injustices ancrées au cœur des modèles. Ils ont aussi la possibilité de mettre en œuvre des campagnes de participation collaborative (« *crowdsourcing* »), afin que tout citoyen, quelle que soit sa place dans la société, puisse fournir des informations détaillées sur les messages reçus en provenance des publicitaires ou des politiciens – ce qui permettrait de faire la lumière sur les pratiques et les stratégies des campagnes de microciblage.

Ces dernières ne recèleraient pas toutes un aspect malfaisant. Au lendemain de l’élection présidentielle américaine de 2012, l’organisme ProPublica a élaboré par exemple ce qu’il a appelé la Message Machine (« machine à messages »), pour disséquer grâce au *crowdsourcing* le modèle de ciblage publicitaire de la campagne d’Obama. Il s’est avéré que différents groupes d’électeurs avaient entendu à propos du président des remarques élogieuses émanant de diverses célébrités, qui s’adressaient vraisemblablement chacune à un public spécifique. Il n’y avait là rien de coupable. Mais en apportant des informations et en dissipant le mystère qui entourait ce modèle, la Message Machine a fait taire (au moins un peu) les soupçons et les rumeurs les plus sordides. Et c’est une excellente chose.

Si l’on envisage les modèles mathématiques comme les moteurs de l’économie numérique, ce qu’ils sont à maints égards, le travail des auditeurs consiste alors à soulever le capot et à nous montrer comment ils fonctionnent. C’est une étape vitale, pour nous permettre d’équiper ces puissants moteurs d’un volant – et de bons freins.

Les auditeurs affrontent toutefois souvent une certaine résistance de la part des géants du Web – qui représentent, en matière d’information, l’équivalent le plus proche des distributeurs de gaz et d’électricité. Google a

par exemple interdit aux chercheurs de créer de multiples faux profils afin de pointer les partis pris de son moteur de recherche<sup>2</sup>. L'entreprise effectue elle-même ce genre d'audits, mais préfère les mener en interne. Elle peut ainsi dissimuler aux regards les rouages profonds de son algorithme et ses préjugés. Ceux qui travaillent en son sein, sujets comme tout le monde au biais de confirmation, auront tendance à ne découvrir que ce qu'ils s'attendent déjà à trouver. Ils ne poseront peut-être pas les questions les plus poussées. Et s'ils constatent des injustices qui semblent avoir pour effet d'accroître les revenus de Google, cela risque d'entraîner des discussions gênantes qu'ils souhaiteront à coup sûr ne pas dévoiler. Il existe donc pour une entreprise de solides arguments en faveur du secret. Mais à l'heure où le public en sait de plus en plus au sujet des ADM, et exige de ces firmes qu'elles rendent davantage de comptes, j'ose espérer que Google n'aura guère d'autre choix que d'ouvrir la porte aux observateurs extérieurs.

Et Facebook pareillement. La politique inflexible du réseau social, consistant à associer aux utilisateurs leur véritable nom, limite considérablement les recherches que des tiers peuvent mener dessus. Cette politique du « vrai nom » s'avère à bien des égards remarquable, en particulier parce qu'elle pousse les utilisateurs à assumer la responsabilité de leurs messages. Mais Facebook doit lui aussi rendre compte vis-à-vis de nous tous – ce qui suppose d'ouvrir davantage la plateforme aux auditeurs externes.

L'État a bien entendu un rôle important à jouer en termes de réglementation, comme ce fut le cas face aux excès et aux tragédies de la première révolution industrielle. Il pourra commencer par adapter les lois déjà en vigueur, puis veiller à leur application.

Comme on l'a vu dans le chapitre sur les scores de crédit, le Fair Credit Reporting Act (FCRA) et l'Equal Credit Opportunity Act (ECOA) visaient à garantir l'équité de ces évaluations. Le FCRA assure au consommateur un droit de regard sur les données utilisées et la possibilité de faire corriger d'éventuelles erreurs, tandis que l'ECOA interdit de rattacher le score d'une personne à son sexe ou à ses origines ethniques.

Ces réglementations ne sont pas parfaites et ont grand besoin d'être actualisées. Les plaintes des consommateurs sont souvent ignorées et rien n'interdit explicitement aux agences de notation d'utiliser les codes postaux en remplacement des origines ethniques. Elles constituent néanmoins un bon point de départ. Nous devons en premier lieu exiger de la transparence.

Chacun devrait pouvoir être averti lorsqu'un score de crédit est utilisé pour le juger ou pour vérifier sa situation. Et chacun devrait avoir accès aux informations prises en compte pour calculer ce score. Si elles sont fausses, nous devrions avoir le droit de les contester et de les rectifier.

Les réglementations devraient d'autre part être étendues aux nouveaux types de sociétés de crédit comme Lending Club, qui utilisent de curieux e-scores pour prédire le risque de défaut des emprunteurs. Elles ne devraient pas être autorisées à opérer dans l'ombre.

Le Americans with Disabilities Act (ADA), qui protège des discriminations professionnelles les personnes souffrant de problèmes de santé, a lui aussi besoin d'être modernisé. La loi interdit à l'heure actuelle l'usage d'examens médicaux pour sélectionner les candidats à l'embauche. Mais il faut désormais tenir compte des tests de personnalité, des scores ayant trait à la forme physique et la réputation, basés sur l'exploitation du Big Data. Ils échappent tous à la législation, ce à quoi il faudrait remédier. Une possibilité déjà à l'étude serait d'étendre le périmètre de l'ADA aux « prédictions » médicales. En d'autres termes, si l'analyse de son génome montre qu'un sujet présente un risque élevé de développer un cancer du sein ou la maladie d'Alzheimer, cette information ne devrait pas servir à lui refuser un emploi.

Il faut aussi élargir le champ du Health Insurance Portability and Accountability Act (HIPAA), qui protège nos informations médicales, afin qu'il s'applique également aux données collectées aujourd'hui par les employeurs, les applications de santé et d'autres entreprises du Big Data. Toutes les données en rapport avec la santé recueillies par des courtiers, comme les recherches de traitements effectuées sur Google, doivent être également protégées.

Si nous voulons frapper fort, nous pourrions songer à nous orienter vers le modèle européen, qui stipule que la collecte de n'importe quelle donnée doit d'abord être approuvée par l'utilisateur. Il interdit en outre qu'elle soit réutilisée à d'autres fins. Le consentement préalable est très souvent contourné, en obligeant l'utilisateur à cocher une case rédigée dans un jargon juridique incompréhensible. Mais la clause de « non-réutilisation » est en revanche très ferme : elle rend la vente des données illégale et les préserve ainsi des courtiers, dont les dossiers alimentent le calcul d'e-scores toxiques et nourrissent les campagnes de microciblage. Grâce à cette clause,

les courtiers européens se voient imposer des limites beaucoup plus strictes – si tant est qu’ils respectent la loi.

Enfin, les modèles qui exercent un impact conséquent sur nos vies, ceux notamment des scores de crédit et des e-scores, devraient être ouverts et accessibles au public. L’idéal serait de pouvoir les gérer sur nos téléphones à l’aide d’une application. Un consommateur qui connaît une fin de mois difficile pourrait par exemple comparer avec cette application l’impact sur son score de crédit des factures de téléphonie et d’électricité impayées, et constater à quel point un score plus faible pourrait affecter son projet d’acheter une voiture. La technologie existe déjà ; seule manque la volonté.

\*\*\*

Un jour d’été 2013, j’ai pris le métro jusqu’à la pointe sud de Manhattan et me suis rendue dans un grand bâtiment administratif situé face à l’hôtel de ville de New York. Je voulais concevoir des modèles mathématiques qui viendraient en aide à la société – le contraire des ADM. Je m’étais donc inscrite comme stagiaire bénévole dans une équipe chargée de l’analyse de données, au sein de la direction du logement et des services sociaux. À New York, on avait atteint le nombre de 64 000 sans-abri, dont 22 000 enfants. Je devais contribuer à la création d’un modèle qui permettrait de prédire le temps qu’une famille sans domicile allait passer dans un hébergement d’urgence et de confier chacune d’entre elles aux services adéquats. L’idée était d’apporter aux gens ce dont ils avaient besoin pour prendre soin d’eux-mêmes et de leur famille, et pour trouver un logement fixe.

Mon travail revenait par bien des aspects à élaborer un modèle de prédiction de la récidive. À l’instar des analystes qui avaient conçu le modèle LSI-R, je m’intéressais aux forces qui conduisaient les gens à retourner en foyer et à celles qui les menaient vers un logement stable. Contrairement à l’ADM utilisée pour infliger des sanctions pénales, notre petite équipe axait cependant le produit de ses recherches sur l’aide aux victimes, pour réduire la précarité et le désespoir. L’objectif consistait à créer un modèle tourné vers l’intérêt général.

Dans le cadre d’un projet distinct mais apparenté, un autre chercheur avait découvert une corrélation extrêmement forte, qui laissait percevoir une solution. Un certain type de familles sans domicile avait tendance à disparaître des foyers d’accueil et à ne jamais y revenir : celles qui, en vertu



d'un programme fédéral appelé « Section 8 », avaient reçu un coupon pour accéder à un logement abordable. Cela n'avait rien de bien surprenant. Si l'on procure un logement social aux familles sans abri, peu d'entre elles choisiront de rester à la rue ou dans un foyer sordide.

Mais cette conclusion pourrait avoir embarrassé le maire de l'époque, Michael Bloomberg, et son administration. La municipalité avait annoncé avec éclat qu'elle voulait apprendre aux familles à se passer de ce dispositif. Elle institua un nouveau système baptisé « Advantage », qui limitait la durée des allocations à trois ans, l'idée étant que l'expiration annoncée de la prestation inciterait les pauvres à gagner davantage et à payer de leur propre poche. Cette vision se révéla très optimiste, comme les données l'ont clairement montré. À la même époque, l'essor du marché immobilier new-yorkais poussait les loyers à la hausse, et rendait la transition d'autant plus ardue. Les familles privées des coupons du programme Section 8 reprirent le chemin des foyers d'urgence.

Les conclusions du chercheur furent mal accueillies. En vue d'une réunion avec d'importants responsables publics, notre groupe avait préparé une présentation PowerPoint sur le problème des sans-abri à New York. L'affichage d'une page sur les statistiques de récidive et l'efficacité de la Section 8 déclencha aussitôt un bref débat, extrêmement inconfortable. Quelqu'un exigea qu'elle soit retirée. La ligne officielle du parti devait prévaloir.

Bien que le Big Data, s'il est géré avec sagesse, puisse fournir de précieux éclairages, beaucoup d'entre eux s'avèrent disruptifs. Le Big Data vise après tout à repérer des schémas que nos yeux sont incapables de percevoir. Pour les experts en données, le défi consiste à bien appréhender les écosystèmes qu'ils explorent, et à présenter non seulement les problèmes qui se posent mais également les solutions possibles. La simple analyse d'un processus de travail pourra mettre en évidence que cinq employés sont apparemment superflus. Mais si l'équipe chargée de l'analyse fait appel à un expert, il aidera peut-être alors à imaginer une version plus constructive du modèle. Il pourrait suggérer dans le cadre d'un système optimisé d'autres tâches susceptibles de leur être confiées, et identifier la formation dont ils auraient besoin pour occuper ces postes. Parfois, la mission d'un expert en données est de comprendre qu'on n'en sait pas assez.

À mesure que j'étudie l'économie des données, je vois émerger quantité de modèles mathématiques qui pourraient être mis au service d'une bonne cause, et un nombre tout aussi important capables de produire d'excellents résultats – sous réserve d'être bien utilisés. Prenons le travail de Mira Bernstein, qui s'emploie à traquer les situations d'esclavage. Docteur en mathématiques diplômée de Harvard, Bernstein a créé un modèle permettant d'inspecter de vastes chaînes d'approvisionnement industrielles – comme celles mises en œuvre pour la production de téléphones mobiles, de chaussures de sport ou de voitures de type SUV – afin d'y déceler des indices de travail forcé. Elle a conçu ce modèle à destination d'un organisme à but non lucratif, *Made in a Free World*, qui souhaite s'en servir pour aider les entreprises à retirer de leurs produits les composants fabriqués par des esclaves. L'hypothèse est qu'elles auront à cœur de se débarrasser de ce fléau, d'une part parce qu'elles sont vraisemblablement contre l'esclavage, mais aussi parce que le fait d'y être associé pourrait avoir un impact dévastateur en termes d'image de marque.

Bernstein a collecté des données issues de multiples sources, notamment les données sur le commerce des Nations unies, les statistiques sur les régions du monde où l'esclavage était le plus répandu, et des informations détaillées sur les composants utilisés dans la fabrication de milliers de produits industriels, et les a injectées dans un modèle qui évalue sous forme d'un score la probabilité qu'un produit donné, en provenance d'une région donnée, ait été fabriqué en recourant au travail forcé. « L'idée est que l'utilisateur contacte son fournisseur et lui dise : "Donnez-moi de plus amples renseignements sur l'origine des composants suivants, présents dans vos ordinateurs" », a-t-elle expliqué au magazine *Wired*. Comme beaucoup de modèles responsables, le détecteur d'esclavage ne va pas plus loin. Il se contente d'indiquer les aspects suspects et laisse la dernière partie de la traque aux êtres humains. Certaines entreprises constateront certainement que le fournisseur soupçonné est parfaitement en règle (tout modèle génère en effet des faux positifs). Ces informations remontent jusqu'à *Made in a Free World*, où Bernstein peut alors étudier le feed-back ainsi obtenu.

Un autre modèle propice à l'intérêt général a fait son apparition dans le domaine de l'action sociale. C'est un modèle prédictif, qui identifie les ménages dans lesquels les enfants ont le plus de chances d'être maltraités. Développé par Eckerd – organisme non lucratif d'aide à l'enfance et aux familles, basé dans le sud-est des États-Unis –, il a été lancé en 2013 dans le

comté d'Hillsborough, en Floride, où se situe la ville de Tampa. Au cours des deux années précédentes, neuf enfants de la région avaient succombé à de mauvais traitements, dont un bébé jeté par la fenêtre d'une voiture. Les modélisateurs ont saisi dans leur base de données 1 500 cas de maltraitance infantile, y compris ceux qui avaient abouti à un décès. Ils ont découvert un certain nombre de marqueurs, tels que la présence d'un petit ami au domicile, des antécédents de toxicomanie ou de violence domestique, ou le fait qu'un des parents ait été placé enfant en famille d'accueil.

S'il s'agissait d'un programme destiné à repérer de potentiels criminels, on perçoit aussitôt à quel point il serait injuste. Avoir vécu en famille d'accueil ou héberger chez soi un partenaire hors mariage ne devrait pas être un motif de suspicion. Le modèle aurait de surcroît beaucoup plus tendance à cibler les pauvres – et à laisser passer d'éventuelles maltraitances perpétrées dans les quartiers riches.

Si l'objectif consiste en revanche non pas à punir les parents, mais à secourir les enfants qui pourraient en avoir besoin, cette ADM en puissance devient alors salutaire. Elle canalise les ressources vers les familles à risque. Selon le *Boston Globe*, dans les deux ans qui ont suivi la mise en place du modèle, le comté d'Hillsborough n'a d'ailleurs déploré aucun nouveau décès d'enfant par maltraitance.

Les modèles comme celui-ci vont se multiplier dans les années à venir ; ils évalueront notre risque d'ostéoporose ou d'infarctus, voleront à l'aide des élèves qui rencontrent des difficultés en mathématiques, et prédiront même quels sont les gens les plus susceptibles de faire une chute invalidante. Beaucoup d'entre eux, à l'instar de certaines des ADM dont nous avons parlé, seront animés des meilleures intentions. Mais ils devront également faire preuve de transparence, en divulguant aussi bien les données utilisées que les résultats de leur ciblage. Ils devront permettre en outre qu'on les soumette à des audits. Ce sont après tout de très puissantes machines : nous devons garder l'œil dessus.

Ni les données, ni les ordinateurs et encore moins les mathématiques ne sont appelés à disparaître. Les modèles prédictifs sont des outils auxquels nous aurons de plus en plus recours pour faire fonctionner nos institutions, déployer nos ressources et gérer nos vies. Comme j'ai essayé de le montrer tout au long de ce livre, leur élaboration ne repose toutefois pas uniquement sur la collecte de données mais dépend aussi des choix que nous effectuons, pour décider de celles qu'il faut prendre en compte et de celles qu'il

convient d'écarter. Ces choix ne se réduisent pas à des questions de logistique, de profits et d'efficacité. Ils revêtent un caractère fondamentalement moral.

Si nous battons en retraite et que nous envisageons les modèles mathématiques comme une force neutre et incontournable, semblable aux conditions météorologiques ou au mouvement des marées, nous abdiquons alors notre responsabilité. Avec pour conséquence, on l'a vu, l'apparition d'ADM qui nous traitent sur le lieu de travail comme les pièces d'un engrenage, excluent certains employés et se régale des injustices. Nous devons tous ensemble nous unir pour contrôler ces ADM, les dompter et les désamorcer. Je caresse l'espoir que l'on s'en souvienne demain comme on se souvient des mines de charbon meurtrières du siècle passé, reliques des premiers jours de cette nouvelle révolution, avant que l'homme n'ait su imposer à l'ère des données le respect de l'équité et l'obligation de rendre compte. Les mathématiques méritent bien mieux que des ADM, et la démocratie aussi.

---

1. On pourrait penser qu'un audit impartial poussera à éliminer de l'analyse des variables telles que l'origine ethnique. Mais si nous voulons mesurer l'impact d'une ADM, ces données sont nécessaires. À l'heure actuelle, la plupart des ADM évitent de prendre directement en compte ce critère car c'est souvent illégal. La mise en évidence de discriminations raciales s'avère cependant plus facile en matière d'emprunts hypothécaires que de prêts automobiles : les organismes de crédit sont en effet tenus dans le premier cas de s'informer sur l'ethnie du demandeur, mais pas dans le second. Comme l'a noté l'informaticienne Cynthia Dwork, si l'on inclut ce critère dans l'analyse, on peut alors quantifier l'injustice raciale partout où on l'observe ; on peut la rendre publique, débattre du problème éthique qu'elle pose et suggérer des remèdes. Cela étant dit, la race est une construction sociale et reste à ce titre difficile à cerner même quand on s'y attèle, comme tout individu métis pourra vous le confirmer (*NdA*).

2. Google a exprimé le désir d'œuvrer pour éliminer les biais de son algorithme, et quelques-uns de ses employés m'en ont brièvement parlé. L'un des premiers conseils que je leur donne, c'est d'ouvrir un peu plus la plateforme aux chercheurs venant de l'extérieur (*NdA*).

# POSTFACE

1001EBooks

**S**i vous êtes comme moi, vous avez des amis qui avant l'élection présidentielle américaine de 2016 ont passé des soirées entières à actualiser les pages du site d'information FiveThirtyEight ou de celui du *New York Times*, impatients de voir qui allait l'emporter. Beaucoup d'entre nous ont par la suite été choqués que les algorithmes de ces organes d'information n'aient pas réussi à prédire correctement l'issue du scrutin.

Dans les sondages électoraux, on demande aux gens s'ils pensent aller voter et pour quel candidat ils penchent. Les algorithmes agrègent ensuite les résultats de multiples enquêtes afin de pronostiquer le vainqueur. Ils analysent pour ce faire les sondages réalisés à divers moments lors des périodes électorales antérieures, en identifiant ceux qui sont tombés le plus juste.

Mais alors que les modèles standard du Big Data s'actualisent de jour en jour, voire à intervalle de quelques heures, la boucle de rétroaction des modèles utilisés pour les élections de mi-mandat ou les scrutins présidentiels ne reçoit de nouvelles données qu'une fois toutes les deux ou quatre années. Beaucoup de choses peuvent se produire entre-temps, et beaucoup de choses se sont bien entendu produites entre 2012 et 2016. La popularité grandissante de Donald Trump et des politiciens pro-Brexit au Royaume-Uni a coïncidé avec un scepticisme croissant vis-à-vis des médias en général et des sondeurs en particulier. Pour que les sondages – et par conséquent leurs algorithmes – fonctionnent, les individus interrogés doivent former un échantillon représentatif de l'électorat. Or, en 2016, les Blancs non diplômés ont été encore plus réticents que d'habitude à répondre

aux questions des sondeurs. Et ce groupe, comme on le sait maintenant, s'est exprimé à une écrasante majorité en faveur de Trump.

Depuis des dizaines d'années, la participation aux sondages ne cesse il est vrai de décliner. Les gens sont tous davantage occupés et beaucoup d'enquêtes passent encore par un appel sur téléphone fixe, alors même que 41 % des ménages – et deux tiers des *millennials* – n'en possèdent plus. Cette tendance globale est néanmoins devenue relativement prévisible, et les analystes sont capables d'en tenir compte dans leurs modèles. Le problème survient lorsqu'un biais d'un nouveau genre fait son apparition, en particulier quand il affecte de manière inégale le corps électoral. Au cours des quatre dernières années, aux États-Unis et dans une grande partie de l'Europe, les politiciens ont exploité de plus en plus une rhétorique anti-immigration, attisant les craintes à l'égard des réfugiés syriens présentés comme une armée d'envahisseurs. Les personnes convaincues par ces messages anti-immigrés ont pu éprouver une réticence supplémentaire à répondre avec franchise aux sondeurs, sachant que beaucoup les catalogueraient comme racistes. L'effet anti-sondage et l'effet anti-immigrés s'ajoutant l'un à l'autre, un sondage mené en 2016 présentera donc des biais très différents d'un sondage mené en 2012, ce qui signifie aussi qu'il déstabilisera n'importe quel algorithme. Quand le résultat d'une élection se joue à quelques points de pourcentage, ces « inconnues inconnues » peuvent alors faire toute la différence entre un raz-de-marée électoral et une défaite.

Que les sondages de 2016 n'aient pas su refléter la réalité semble déjà suffisamment regrettable. Mais ce qui est encore plus grave, c'est que les sondages peuvent en eux-mêmes influencer l'issue d'un scrutin. Pour commencer, les enquêtes réalisées au tout début d'une période électorale sont notoirement inexactes et tendent à défavoriser les candidats les moins connus. À ce stade, il s'agit davantage d'un concours de popularité : les vraies questions politiques sont rarement sinon jamais abordées, et les personnes interrogées se fient donc plutôt aux noms qu'elles reconnaissent. Malgré cela, si un candidat récolte à l'occasion de ces tout premiers sondages des scores décevants, les observateurs n'hésiteront pas à affirmer qu'il n'a « aucune chance » de l'emporter. Ayant entendu ce message, les électeurs qui auraient peut-être voté pour lui se montreront moins enclins à donner de leur temps ou à soutenir financièrement sa campagne. Les résultats des sondages deviennent une prophétie autoréalisatrice, qui

favorise fortement les candidats connus ou soutenus par un parti, au détriment de l'étude attentive des enjeux politiques.

Les sondages réalisés quelques jours avant une élection posent eux aussi problème, tout comme les sondages de « sortie des urnes », effectués à proximité immédiate des isoloirs. Si un électeur apprend à ce stade avancé que son candidat est clairement parti pour gagner, ou au contraire qu'il n'a pas la moindre chance, il aura moins envie d'aller voter. Ce mécanisme opère également à l'échelle d'une campagne ; l'algorithme de simulation électorale d'Hillary Clinton (baptisé « Ada » par son équipe, en hommage à la mathématicienne Ada Lovelace) avait supposé qu'elle emporterait facilement les États du Wisconsin et du Michigan. En conséquence, Clinton n'a jamais fait campagne dans le premier et s'est très peu souciée du second. Elle a finalement perdu les deux États, avec au total moins de 35 000 voix de retard.

Les sondages politiques constituent-ils donc une arme de destruction mathématique ? Ils exercent à coup sûr un grand pouvoir et une influence importante, à la fois sur la manière dont nous choisissons nos candidats et sur les sujets dont nous-mêmes, ainsi que les médias, décidons de débattre. Beaucoup de leurs algorithmes sont également opaques – sinon secrets, en tout cas désespérément complexes. Reste la question de leur caractère destructeur. Bien qu'on puisse selon moi argumenter dans les deux sens, j'ose espérer, compte tenu de l'échec absolu des sondages en 2016, qu'on leur accordera moins d'attention lors du prochain cycle électoral. Nous leur avons donné du pouvoir, et nous pouvons le reprendre.

Au cours de cette saison électorale, la défaillance des sondages ne fut pas la seule, en matière d'algorithmes, à faire la une des journaux. Alors que les deux camps s'appuyaient de plus en plus, pour partager leurs messages, sur des réseaux sociaux comme celui de Facebook, les algorithmes qui gouvernent ces plateformes ont fait l'objet d'une remise en cause croissante.

En 2014, Facebook a dévoilé la fonctionnalité « Trending Topics », qui mettait en avant des articles d'actualité populaires sur le site. Cette catégorisation ne reposait pas sur un processus exclusivement algorithmique. Du moins pas au départ. Dans un premier temps, un algorithme rudimentaire (dit « naïf ») faisait remonter une liste d'articles fréquemment partagés et donc potentiellement populaires, classés par ordre d'intérêt. Cette liste était ensuite confiée à un petit groupe de journalistes,

engagés par Facebook pour sélectionner les Trending Topics. Leur travail consistait en premier lieu à trouver un titre et à résumer les articles, mais ils supprimaient aussi les éléments redondants, les canulars manifestes, ou les contenus qui n'étaient pas corroborés par une « liste blanche » de rédacteurs. Ils pouvaient aussi effacer des articles ou les mettre à l'index.

Lorsque des signalements ont commencé à circuler, indiquant que ses journalistes avaient à plusieurs reprises supprimé des articles orientés à droite, Facebook a renvoyé l'équipe et placé les Trending Topics entre les mains d'un nouvel algorithme, afin vraisemblablement d'éliminer les partis pris supposés. Une fois aux commandes, cet algorithme s'est néanmoins comporté de manière encore plus erratique que les journalistes humains, et de nombreux rapports ont fait rapidement état de fausses nouvelles (*fake news*) s'immiscant parmi les Trending Topics.

Ce fiasco soulève d'intéressantes questions. Premièrement, qu'entend-on par « parti pris » ? Quand on voit à quel point nos journaux et nos échanges publics se sont polarisés – les faits eux-mêmes et les limites d'un débat acceptable suscitent aujourd'hui de profonds désaccords –, quelle politique Facebook pourrait-il adopter, qui serait reconnue de tous comme équitable ? Même si Facebook traitait ouvertement et publiquement de son influence et de sa responsabilité (ce qui n'est absolument pas le cas), on aurait du mal à répondre.

Une politique de validation des articles cohérente et ouverte s'avérerait sans aucun doute délicate à mettre en œuvre, et prêterait le flanc à la controverse et aux récriminations. Ce serait néanmoins toujours mieux que de s'appuyer sur un algorithme opaque, capable si facilement de désinformer et d'induire les gens en erreur. Les processus algorithmiques reflètent au même titre que tout processus humain certaines valeurs et certains principes éthiques ; s'ils paraissent plus « propres », c'est uniquement parce qu'ils savent mieux le cacher.

En nouant des partenariats avec des vérificateurs externes, Facebook a consenti quelques efforts en faveur d'une plus grande transparence. Et bien que cela puisse paraître dérisoire, en comparaison des actions menées pour censurer la nudité ou traquer la propagande de Daech, c'est un pas dans la bonne direction.

De manière générale, les entreprises qui développent et déploient les algorithmes devraient être les premières à en répondre. Elles devraient assumer la responsabilité qu'implique leur influence et démontrer que leur



activité ne cause aucun préjudice, tout comme les industriels de la chimie sont tenus d'apporter la preuve qu'ils ne polluent pas autour d'eux les cours d'eau et les nappes phréatiques. Il ne s'agit pas de dire que les algorithmes devraient être tous interdits ou contraints de divulguer leur code source, mais cela signifie en revanche que la charge de la preuve revient aux entreprises, et qu'il faudrait les obliger à vérifier régulièrement la légalité, l'équité et l'exactitude de leurs algorithmes.

Les entreprises ont rechigné jusqu'ici à endosser cette responsabilité supplémentaire. Mais des journalistes et des chercheurs s'en sont récemment emparés. Chez ProPublica, Julia Angwin et son équipe ont pris l'initiative d'expliquer sérieusement au public comment les algorithmes fonctionnent, les partis pris qu'ils sont susceptibles de renfermer ou les discriminations tarifaires qu'ils peuvent engendrer. Dans l'un de leurs articles les plus novateurs, ils ont étudié le système d'évaluation du risque de récidive COMPAS, conçu par l'entreprise Northpointe et utilisé dans le comté de Broward, en Floride. Pour mener cet audit, Angwin et son équipe ont déposé sur une période de trois ans de multiples requêtes en vertu du Freedom of Information Act (loi garantissant l'accès aux documents administratifs) et rendu également leur code public.

Ils ont pu constater que le système de notation COMPAS était correctement calibré ; autrement dit, que la proportion de scores indiquant un risque élevé correspondait bien, tant parmi les Noirs que parmi les Blancs, au taux de récidive réellement observé. C'est un moyen courant de définir l'équité, et celui sur lequel Northpointe s'appuyait pour optimiser son modèle. Cela veut dire en clair que le système avait autant de chances de « bien » juger du niveau de risque des inculpés noirs que de celui des Blancs.

De gros problèmes survenaient cependant quand le système se trompait. L'équipe d'Angwin découvrit qu'il y avait deux fois plus de faux positifs chez les Noirs que chez les Blancs, et deux fois plus de faux négatifs chez les Blancs que chez les Noirs. Ce qui signifie que les Noirs avaient deux fois plus de chances de se voir attribuer un risque élevé sans pour autant retourner en prison, et les Blancs deux fois plus de chances d'être gratifiés d'un risque faible mais d'être en fin de compte à nouveau incarcérés dans les trois ans. Cela paraît assez grave, surtout lorsqu'on considère qu'un faux positif entraîne souvent une détention prolongée.

Comment le système pouvait-il être à la fois bien calibré et injuste ? On a là une situation classique, dans laquelle des points de vue différents conduisent à des définitions distinctes du succès : l'objectif primordial de Northpointe consistait à détecter les malfaiteurs aussi souvent et aussi précisément que possible, tandis que l'étude de ProPublica mettait en lumière les atteintes contre les droits individuels des personnes indûment accusées.

Nous n'aurons sans doute jamais une définition simple et universellement acceptée de ce qui caractérise un algorithme juste. Mais heureusement, nous en débattons enfin. Il nous appartient désormais de poursuivre cette discussion de façon réfléchie, méticuleuse et inclusive, afin de fixer une norme pour l'avenir de la responsabilité algorithmique. La question n'est, comme on l'a vu, pas uniquement technique – elle est aussi éthique. Et elle exige la contribution d'une grande variété d'individus, notamment de ceux qui seraient affectés par l'algorithme en question, pour identifier les preuves à même de les convaincre qu'un algorithme donné « fonctionne ».

Nous devrions en tant que scientifiques proposer pour l'audit d'un algorithme des méthodes statistiques judicieuses et exhaustives. En tant que citoyens, il nous faudra contrôler avec le plus grand soin la pertinence et la clarté de ces audits.

C'est un débat public qui est d'ores et déjà entamé, mais auquel je voudrais apporter une modeste suggestion. Reformulons la question de l'équité : au lieu de batailler sur l'indicateur à utiliser pour déterminer l'équité d'un algorithme, essayons plutôt d'identifier les parties concernées et de mesurer leur préjudice relatif.

Dans le cas de l'algorithme du risque de récidive, il s'agirait de comparer le préjudice d'un faux positif – une personne affligée à tort d'un score de risque élevé et injustement incarcérée – et le préjudice d'un faux négatif – une personne relâchée par erreur et qui *pourrait* commettre un crime.

Les algorithmes employés pour repérer les maltraitances infantiles offrent un autre exemple intéressant. Plusieurs administrations locales à travers les États-Unis ont expérimenté l'utilisation de l'analyse de données afin de déployer plus efficacement les services d'aide à l'enfance. Elles reçoivent les appels d'enseignants, de voisins ou de médecins qui s'inquiètent du bien-être de tel ou tel enfant. Compte tenu de leurs ressources restreintes, impossible de dépêcher pour chaque appel un travailleur social. D'où la question : comment choisir ceux auxquels on va répondre ?

Le problème du préjudice relatif ne saurait se poser d'une manière plus saisissante. Un faux positif correspondra ici à l'envoi d'un travailleur social au sein d'une famille qui n'avait aucun risque de maltraiter son enfant. Elle pourra dans certains cas s'en voir retirer inutilement la garde – un traumatisme pour lui, et une tragédie pour toute la famille. Les tendances historiques laissent penser en outre que cette issue négative frappe plus volontiers les familles pauvres et les minorités ethniques.

De l'autre côté de l'équation, nous envisageons le préjudice d'un faux négatif, c'est-à-dire d'un enfant en danger immédiat mais que le système ignore et laisse par suite sans protection. Tragédie là aussi, cette fois-ci pour l'enfant victime. Comment comparer les deux ? Si nous voulons bâtir correctement cet algorithme, c'est tout simplement indispensable.

Il est important de noter, lorsqu'on s'efforce d'appréhender la notion de préjudices relatifs, qu'ils dépendent entièrement du contexte. Si le score de risque élevé attribué à un prévenu donné le rendait par exemple éligible à un programme de réinsertion, qui pourrait l'aider à trouver un emploi à sa sortie de prison, les faux positifs nous inquiéteraient beaucoup moins. Dans le cas de la maltraitance infantile, si nous étions sûrs qu'un score de risque élevé aboutisse à une étude impartiale et approfondie de la situation au domicile, nous craindrions moins que des enfants soient inutilement retirés à leurs parents. La *manière* dont un algorithme sera utilisé devrait en définitive impacter la façon de le construire et de l'optimiser.

En plus de solliciter pour l'élaboration des algorithmes l'apport des parties concernées en premier lieu, et de prendre en compte leurs préjudices relatifs, il nous faut également définir des normes pour les surveiller une fois mises en place, afin de nous assurer qu'ils fonctionnent comme prévu. Même si cela peut paraître évident, nous avons récemment constaté les impacts d'algorithmes qui s'étaient vu accorder une confiance bien trop aveugle.

Entre 2013 et 2015, par exemple, l'agence pour l'emploi du Michigan avait déployé un algorithme destructeur ironiquement baptisé Midas. Destiné à débusquer les demandes d'indemnités chômage frauduleuses, il s'était emballé et avait incriminé à tort plus de 20 000 personnes, avec plus de 93 % de faux positifs. Il avait ruiné l'existence de ces gens, en leur infligeant jusqu'à 100 000 dollars d'amende au moment précis où ils avaient le plus besoin d'argent.

Examinons cet exemple dans le contexte des préjudices relatifs. Le système était conçu pour veiller à ce que l'État du Michigan ne verse pas d'indemnités injustifiées, mais négligeait totalement la possibilité qu'une personne puisse être injustement accusée de fraude. S'il n'y a guère de raisons de penser que les faux positifs étaient intentionnels, on ne saurait non plus les réduire à de simples incidents techniques. Le seul fait que ce système automatique ne comporte aucun mécanisme de responsabilité témoigne que l'on avait très peu songé aux potentielles victimes. En d'autres termes, le ratio de « préjudice relatif » était soit nul, soit infini, selon la manière de le calculer.

Les exemples de ce genre montrent à quel point la question des normes de responsabilité devient aujourd'hui un enjeu critique. Quand des systèmes algorithmiques tels que le système Midas souffrent de défauts rédhibitoires, volontaires ou non, ils finissent par causer encore plus de torts que les systèmes humains qu'ils remplacent. Et comme on l'a vu à plusieurs reprises dans ce livre, la souffrance qui en résulte n'est pas équitablement répartie, mais échoit plutôt aux citoyens les plus vulnérables. Les gens qui ne peuvent précisément pas se permettre d'engager de coûteux avocats doivent affronter directement la machine. C'est une lutte déloyale, et de tels exemples plaident clairement pour que la charge de la preuve incombe à ceux qui conçoivent et mettent en œuvre les algorithmes.

Il faut cesser en bref de s'appuyer sur une foi aveugle et entreprendre enfin de conférer à la science des données un caractère réellement scientifique. Nous devons acquérir le scepticisme nécessaire pour évaluer, comprendre et justifier les données qui alimentent nos algorithmes, et les résultats qu'ils produisent.

Tout part des réalités concrètes. Prenons l'algorithme auquel nous sommes si nombreux à faire confiance au quotidien : le moteur de recherche de Google. Il fonctionne la plupart du temps magnifiquement bien, et affiche en réponse à nos requêtes des documents pertinents. Mais ces dernières années, nous avons pu aussi en observer les conséquences néfastes. Au lendemain de l'élection présidentielle, par exemple, si l'on demandait à Google « qui a remporté le vote populaire ? », le premier résultat pointait vers un blog conspirationniste affirmant que c'était Donald Trump. Quant au jeune Dylann Roof – qui avait abattu en 2015 neuf paroissiens dans la ville de Charleston –, il s'était de son propre aveu

radicalisé après avoir cherché les termes « crimes de Noirs sur des Blancs », et pris pour argent comptant les résultats obtenus.

Cela ne devrait pas nous surprendre. Les algorithmes d'apprentissage automatique et d'intelligence artificielle n'intègrent pas une modélisation du monde capable de distinguer de façon fiable la vérité du mensonge. La technologie actuelle est extraordinairement performante dès qu'il s'agit de parcourir des univers clos, dotés de règles et de principes bien définis. Raison pour laquelle les champions d'échecs, de dames, de poker ou de go peuvent être mis en difficulté, voire vaincus, par des algorithmes. Inutile en revanche d'espérer que l'algorithme de Facebook puisse à court terme déterminer par ses propres moyens si un article donné sert de propagande ou dit la vérité. L'ordinateur Watson lui-même, célèbre pour avoir battu tous les concurrents du jeu *Jeopardy !*, n'avait pu réussir que parce que ses données d'entraînement – c'est-à-dire la totalité du contenu d'Internet – étaient en majorité exactes du point de vue des faits. Mais si on lui avait posé des questions à propos de l'Holocauste ou d'autres sujets dominés sur le Web par les tendances conspirationnistes, Watson aurait produit de bien étranges réponses.

Quand les spécialistes – les *data scientists* – parlent de « qualité des données », ils font généralement référence à leur nombre ou à leur netteté. Y en a-t-il assez pour entraîner un algorithme ? Les chiffres correspondent-ils à ce que l'on attend ou sont-ils aléatoires ? Or, nous n'avons en l'occurrence aucun problème de qualité : les données sont disponibles, et d'ailleurs pléthoriques. Elles sont simplement fausses. Mais elles sont néanmoins présentées comme fiables – Google Home et Alexa vantent ainsi leurs services domestiques comme une extension de nos cerveaux, plus pratique et plus sûre qu'un dictionnaire de cuisine. Nous devrions exiger en matière de communication publicitaire davantage d'honnêteté, avant de déplorer d'autres Pizzagates<sup>1</sup> ou d'autres tueries dans les églises.

Les individus aux penchants conspirationnistes ne sont toutefois pas les seuls à se laisser berner. La communauté technologique est également en faute, et doit se mettre au travail. Nous devons nous assurer que les données utilisées offrent une représentation effective et complète du monde – et, plutôt que de l'influencer, surtout si c'est pour aggraver les situations de détresse, qu'elles témoignent comme nous l'espérons de ses souffrances. Pour y parvenir, le rôle des *data scientists* aurait besoin d'être élargi.

Il ne s'agit pas d'affirmer qu'ils puissent tous être experts dans tous les domaines, ou qu'ils doivent le devenir. Mais cela veut dire que nous devons examiner avec scepticisme d'où viennent nos données, et si elles reflètent bien toute la complexité de la vie réelle.

Dans le cas de la police prédictive, il s'agirait de refuser d'associer les données relatives aux arrestations et les données concernant la criminalité. Bien que sa consommation s'établisse à un niveau comparable chez les Noirs et chez Blancs, nous savons déjà que les Noirs ont beaucoup plus de chances d'être arrêtés pour possession de cannabis. L'écart varie en fonction du commissariat. Et si le nombre d'arrestations reflète non seulement le taux de criminalité mais aussi les pratiques de la police, cela signifie que le modèle prédictif de maintien de l'ordre fait de même. Mais à quelle proportion ? Nous l'ignorons.

Ce n'est là qu'une question parmi tant d'autres. Les accusés noirs sont-ils plus souvent placés en détention ? Ont-ils plus de mal à payer leur caution, et plaident-ils de ce fait plus fréquemment coupables d'une inculpation qu'un coûteux avocat aurait pu leur éviter ? Comment ces facteurs influencent-ils les données qui alimentent le modèle ?

Nous devons apprendre non seulement à questionner nos algorithmes, mais aussi notre processus de collecte des données. Cette démarche difficile réclame en règle générale des compétences et des ressources qui vont au-delà du travail habituel d'un *data scientist*. Cela ne veut pas dire pour autant qu'il faille abandonner nos efforts. S'agissant de situations où notre activité impacte profondément la vie des gens, nous devrions au contraire aspirer à des normes plus strictes, même s'il faut pour cela faire appel à des experts externes pour nous mettre sur la bonne voie.

Ce qui nous mène à une question plus vaste, sur la nature de la preuve elle-même. S'il y a bien une chose que j'ai apprise, au cours des recherches menées pour la rédaction de ce livre, c'est que chaque discipline intellectuelle possède une conception très différente de ce qui constitue une preuve, des aspects d'un raisonnement qui doivent être « démontrés », et du moment à partir duquel les indices accumulés forment un argument probant. Ce qui est considéré comme une vérité observable en droit ou en philosophie ne le sera pas forcément en psychologie.

Compte tenu de ces énormes différences, il est donc normal qu'un élément jugé convaincant par l'expert d'un domaine donné paraisse quelquefois bancal aux yeux de quelqu'un d'autre. Les mathématiciens

s'appuient sur des preuves logiques, qui semblent de fait assez infaillibles. Mais la réalité est qu'on y découvre aussi des erreurs, parfois plusieurs années après qu'elles ont été admises par la communauté mathématique. Aucune discipline ne détient la vérité absolue. Au lieu de renâcler face à nos désaccords, nous devrions tirer de ces multiples points de vue une force supplémentaire.

Dans le contexte politique actuel, celui de la « post-vérité » et des « faits alternatifs », on pourrait estimer que c'est un vœu pieux. Si des politiciens, partout à travers le monde, dénigrent activement les données statistiques et sapent la confiance que le public leur accorde, comment espérer que l'industrie du Big Data clarifie les choses, au lieu d'ajouter au bruit ambiant ?

Il le faut, car il n'y a pas d'alternative. Aujourd'hui, des géants comme Google, Amazon et Facebook exercent une fantastique emprise sur la société parce qu'ils contrôlent les données. Ils engrangent de gigantesques profits, tout en laissant à d'autres la responsabilité de vérifier les faits. Que Steve Bannon<sup>2</sup>, qui s'emploie à ébranler la confiance du public vis-à-vis des sciences et des preuves scientifiques, siège en même temps au conseil d'administration de Cambridge Analytica (une société de données politiques qui s'est attribué le mérite d'avoir fait gagner Trump, en se targuant de mener en secret des campagnes de « dissuasion électorale ») n'a rien d'une coïncidence.

Tout cela participe d'une tendance plus large à la privatisation des données, utilisées dans une optique de profit et d'influence – le public se trouvant quant à lui exclu du processus et enjoint de se comporter convenablement et de faire confiance aux algorithmes. Mais il est temps désormais de se rebeller.

Les algorithmes deviendront dans les années à venir de plus en plus omniprésents. Nous devons exiger que les systèmes garantissant leur responsabilité se généralisent eux aussi. Commençons dès maintenant à bâtir un cadre, pour s'assurer à long terme que les algorithmes rendent des comptes. Posons comme base la démonstration de leur légalité, de leur équité et de leur ancrage factuel. Et continuons au fil du temps de préciser ce que ces critères signifient, en fonction du contexte. Ce sera un travail collectif, et nous aurons besoin d'autant d'avocats et de philosophes que d'ingénieurs. Mais en concentrant nos efforts, il sera possible d'y arriver.

Nous ne pouvons pas nous permettre de rester spectateurs.

- 
1. Le Pizzagate est une théorie conspirationniste, selon laquelle une pizzeria de Washington était au cœur d'un vaste réseau de pédophilie impliquant l'ancien directeur de campagne d'Hillary Clinton (*NdT*).
  2. Directeur exécutif de la campagne présidentielle de Donald Trump, nommé ensuite conseiller stratégique du président des États-Unis puis limogé le 18 août 2017 (*NdT*).



# REMERCIEMENTS

**M**erci à mon mari et à mes enfants pour leur incroyable soutien. Merci également à John Johnson, Steve Waldman, Maki Inada, Becky Jaffe, Aaron Abrams, Julie Steele, Karen Burnes, Matt LaMantia, Martha Poon, Lisa Radcliffe, Luis Daniel et Melissa Bilski. Merci enfin à ceux sans qui ce livre n'existerait pas : Laura Strausfeld, Amanda Cook, Emma Berry, Jordan Ellenberg, Stephen Baker, Jay Mandel, Sam Kanson-Benanav et Ernie Davis.

**L'EXEMPLAIRE QUE VOUS TENEZ ENTRE LES MAINS A ÉTÉ RENDU POSSIBLE  
GRÂCE AU TRAVAIL DE TOUTE UNE ÉQUIPE.**

**ÉDITION :** Flore Gurrey  
**COUVERTURE :** Quintin Leeds  
**CONCEPTION GRAPHIQUE :** Sara Deux  
**MISE EN PAGE :** Pixellence  
**RÉVISION :** Emmanuel Dazin et Sarah Ahnou  
**FABRICATION :** Maude Sapin  
**COMMERCIAL :** Pierre Bottura  
**PRESSE/COMMUNICATION :** Isabelle Mazzaschi, Jérôme Lambert et Adèle Hybre  
**RELATIONS LIBRAIRES :** Jean-Baptiste Noailhat

**DIFFUSION :** Élise Lacaze (Rue Jacob diffusion), Katia Berry (grand Sud-Est), François-Marie Bironneau (Nord et Est), Charlotte Jeunesse (Paris et région parisienne), Christelle Guillemot (grand Sud-Ouest), Laure Sagot (grand Ouest), Diane Maretheu (coordination) et Charlotte Knibiehly (ventes directes), avec Christine Lagarde (Pro Livre), Béatrice Cousin et Laurence Demurger (équipe Enseignes), Fabienne Audinet (LDS), Bernadette Gildemyn et Richard Van Overbroeck (Belgique), Nathalie Laroche et Alodie Auderset (Suisse), Kimly Ear (Grand Export)

**DISTRIBUTION :** Hachette

**DROITS FRANCE ET JURIDIQUE :** Geoffroy Fauchier-Magnan  
**DROITS ÉTRANGERS :** Sophie Langlais  
**ENVOIS AUX JOURNALISTES ET LIBRAIRES :** Patrick Darchy  
**LIBRAIRIE DU 27 RUE JACOB :** Laurence Zarra  
**ANIMATION DU 27 RUE JACOB :** Perrine Daubas  
**COMPTABILITÉ ET DROITS D'AUTEUR :** Christelle Lemonnier, Camille Breynaert et Christine Blaise  
**SERVICES GÉNÉRAUX :** Isadora Monteiro Dos Reis

ISBN papier : 978-2-35204-980-7  
ISBN numérique : 978-2-7112-0043-6  
Dépôt légal : novembre 2018