

Big Data et marketing prédictif. Vers un « consommateur nu » ?

Jean-Luc Giannelloni et Emmanuelle Le Nagard

Co-rédacteurs en chef de *Décisions Marketing*

L'homme nu, du romancier Marc Dugain et du journaliste Christophe Labbé, sorti il y a quelques jours, décrit un monde où l'individu est dépossédé de toutes ses données personnelles, « aspirées » des *smartphones*, tablettes, PC et autres objets connectés au travers de l'Internet par des entreprises privées, notamment les fameux GAFAs (*Google, Apple, Facebook et Amazon*). Une association avec les agences de renseignement américaines prend le contrôle de nos existences au profit d'un nouveau type de dictature, que nous laissons se créer de notre plein gré. L'ouvrage décrit ainsi une civilisation d'« hommes nus », sans mémoire et programmés. Par son titre, cet éditorial y fait explicitement référence, en s'interrogeant sur le rôle du marketing dans ce monde émergent qu'Attali qualifiait en 2014 de « dictature de la consommation »¹.

Le « *Big Data* » (BD), puisqu'il s'agit de cela, n'a qu'environ dix ans, mais il est devenu incontournable aujourd'hui, tant au plan des pratiques qu'au plan académique. Au plan pratique, et d'après une « infographie officielle » publiée au *Congrès Big Data Paris* en mars 2016, le sujet « *big data* » a généré 450 000 messages en 2015². Au plan académique, le nombre de publications consacrées au sujet est également en train d'exploser (Chen, Chiang et Storey, 2012) et il existe même au moins six revues académiques de création récente contenant *Big Data* dans leur intitulé³.

Le BD, que l'on désigne en français parfois par « données massives », a de nombreux défenseurs. Pour Babinet (2015), c'est un moyen de « penser l'homme et le monde autrement », selon le titre de l'un de ses ouvrages. Pour lui, c'est un réel levier de croissance pour les entreprises, en ce qu'il permet de « réveiller » les données détenues par l'entreprise et de les faire travailler en les

1/ Source : <http://blogs.lexpress.fr/attali/2014/08/25/la-dictature-de-la-consommation/> ; consultation du 11 mai 2016.

2/ Source : <https://www.sparklane-group.com/fr/blog/2015-une-annee-de-big-data/> ; consultation du 19 avril 2016.

3/ *Big Data* (Mary Ann Libert, Inc) créée en 2013 ; *Journal of Big Data* (Springer), *Big Data Research* (Elsevier), *International Journal of Big Data Intelligence* (Inderscience) et *Big Data & Society* (Sage), toutes créées en 2014 ; *Big Data Analytics* (BioMed Central), en cours de création. Cette liste ne prétend pas être exhaustive.

croisant et en les corrélant. Cela permet ainsi de « disposer d'une nouvelle lecture de l'entreprise, dans le même temps d'améliorer sa performance de manière significative » et, surtout, de produire de manière plus intelligente. Cela amène Babinet à conclure : « Le XX^e siècle était celui de la production, le XXI^e siècle sera celui de l'intelligence »⁴. Plus généralement, d'autres vont même jusqu'à affirmer que le BD est un outil intelligent qui peut aider à combattre la pauvreté, le crime et la pollution⁵.

Mais le BD peut aussi être considéré comme un « *big brother* » en costume cravate, potentiellementivateur de liberté et annonciateur d'un monde quasi Orwellien⁶. Dans la sphère économique, l'utilisation par les entreprises de données sur les consommateurs pour améliorer leur offre n'est pas récente. Elle est même au cœur de la démarche d'étude, elle-même à la source des pratiques de segmentation qui constituent le cœur du marketing. Elle est également au cœur de modèles prédictifs utilisés, notamment en marketing direct, depuis les années 80. Le changement induit par le BD vient du « saut quantique » effectué dans la quantité et la nature très hétérogène des données disponibles, qui permet désormais de travailler en temps réel ou presque sur des comportements réels à l'échelle des populations.

Disposer d'une masse de données ne suffit pas. Encore faut-il disposer de méthodes de traitement qui permettent réellement d'en extraire du sens. De ce point de vue, sous l'impulsion d'entreprises comme *Google* ou *Amazon*, des algorithmes permettant de traiter efficacement ces données ont été mis au point. Ces algorithmes intègrent également des routines d'apprentissage, le *machine learning*, issues de l'intelligence artificielle. La capacité des marques à prévoir, statistiquement, les comportements des consommateurs s'est par conséquent fortement améliorée, c'est ce que l'on nomme aujourd'hui le marketing prédictif.

Lorsque l'on lit sur ce sujet, quelle que soit la source de littérature, le déséquilibre entre les « pro » et les « anti » BD est frappant, à l'avantage des premiers. Le premier objectif de cet éditorial est de contribuer partiellement et modestement à rétablir l'équilibre entre ces deux positions. Deux idées seront principalement défendues et articulées pour défendre la problématique suivante : le BD n'est pas une réelle nouveauté et, à ce titre, il porte en germe depuis longtemps une vision et une pratique discutables, sinon dangereuses, du marketing, en tout cas incompatibles avec celle exposée dans un précédent éditorial (Le Nagard et Giannelloni, 2016) et défendues également ailleurs sous le terme « Marketing 3.0 » (Kotler, Kartajaya et Setiawan, 2012). Cette articulation donne naissance aux deux parties constitutives de l'exposé. Dans un premier temps, nous tenterons de montrer que le BD et le marketing prédictif sont de « vraies-fausse » innovations. Dans un second temps, nous développerons l'idée selon laquelle leur vision et leur pratique actuelle représentent un réel danger pour une vision du marketing empreinte d'humanisme.

Big data et marketing prédictif : une « vraie-fausse » innovation ?

Il ne s'agit pas d'être totalement à charge. A certains égards, le BD représente un réel progrès et permet de résoudre des problèmes d'ordre économique et sociétal. Mais l'analyse des méthodes

4/ Source : <https://www.sparklane-group.com/fr/blog/avis-dexpert-gilles-babinet-le-big-data-des-gains-dopportunités-pour-lentreprise/> ; consultation du 19 avril 2016.

5/ Source : http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?_r=1&scp=1&sq=Big%20Data&st=cse ; consultation du 19 avril 2016.

6/ Que Dugain et Labbé (2016) qualifient de « *Big Mother* », pire que « *Big Brother* » par le fait qu'il n'y a aucune coercition visible.

et algorithmes utilisés montre que, dans l'immense majorité des cas, les outils sont en réalité anciens.

Comment définir le Big Data ?

Définir le *Big Data* (BD) n'est pas simple. Les définitions sont d'ailleurs plus managériales qu'académiques. Pour le magazine professionnel *lebigdata.fr*, c'est même impossible car « Etant un objet complexe polymorphe, sa définition varie selon les communautés qui s'y intéressent en tant qu'utilisateur ou fournisseur de services »⁷. Celle proposée par *definitions-marketing.com* est longue de plus d'une page⁸, ce qui la rend très peu opérationnelle. Celle de *Wikipedia* est courte (« *Le big data* (...) désigne des ensembles de données qui deviennent tellement volumineux qu'ils en deviennent difficiles à travailler avec des outils classiques de gestion de base de données ou de gestion de l'information »⁹) mais peu satisfaisante en ce qu'elle n'embrasse que le côté volumineux des données.

Or, plus qu'à de gros volumes de données, l'expression « *big data* » (BD) se réfère avant tout à des technologies logicielles qui permettent de traiter en temps réel des données issues de sources hétérogènes et complexes, selon des protocoles qualifiés de « massivement parallèles ». Les consommateurs ne sont pas, et de loin, les plus gros producteurs des données les plus complexes à traiter. Par exemple, une heure de vol d'un quadriréacteur produit deux téraoctets¹⁰ de données techniques qu'il est essentiel d'interpréter pour repérer les défauts et anticiper les pannes¹¹. Références de géolocalisation, textes dans des courriels ou sur les réseaux sociaux, appels téléphoniques, enregistrements dans des bases de données, images vidéos, relevés de compteurs, données physiques issues de capteurs¹² ... sont des exemples de types de données que les algorithmes actuels permettent d'agrégier et de mettre en relation pour en extraire du sens.

Le Big Data, un réel progrès dans certains cas

Le BD est désormais partout, tout le temps. Dans certains domaines, il a permis de réels progrès. La littérature managériale (e.g. Mayer-Schönberger et Cukier, 2013) et la presse fourmillent d'exemples. En voici trois¹³. Dans le domaine fiscal, la fraude à la TVA intracommunautaire a été pratiquement éliminée en Belgique grâce à l'utilisation d'un logiciel de *data mining* qui détecte automatiquement les escroqueries en croisant les archives de plusieurs ministères. Un formulaire de TVA suspect peut désormais être repéré presque en temps réel. En Italie, un outil similaire croise 128 bases différentes pour repérer d'éventuels décalages entre le train de vie et les revenus déclarés des contribuables. L'apport du BD est moins dans le principe de l'outil, assez ancien, que dans sa capacité à faire les contrôles en temps réel et en mode prédictif, ce qui diminue considérablement les coûts de recouvrement. La santé est un deuxième exemple. Les données traitées sont hétérogènes (imagerie, ordonnances, compte rendus d'interventions, bases internationales, publications scientifiques...). En utilisant un algorithme BD, une équipe de Boston avait détecté l'épidémie de grippe H1N1 avant qu'elle ne se déclare. Couplés à des algo-

7/ Source : <http://www.lebigdata.fr/definition-big-data> ; consultation du 11 mai 2016.

8/ <http://www.definitions-marketing.com/definition/big-data/> ; consultation du 11 mai 2016.

9/ Source : https://fr.wikipedia.org/wiki/Big_data ; consultation du 11 mai 2016.

10/ Un téraoctet = 1000 gigaoctets.

11/ Source : entretien avec François Bourdoncle, *Le Monde*, supplément « Big Data », 21711, 6 novembre 2014.

12/ Un capteur est une interface transformant l'état d'une grandeur physique observée en une information utilisable. Il s'agit d'un élément de base d'un système d'acquisition de données. Source : Wikipédia. Consultation du 26 avril 2016.

13/ Source : *Le Monde*, supplément « Big Data », 21711, 6 novembre 2014.

rithmes d'intelligence artificielle (IA), ils aident les médecins à traiter des dossiers extrêmement complexes, par exemple dans le domaine du cancer. L'objectif, à terme, est de pouvoir pratiquer une médecine *one to one* réellement préventive pour ainsi à la fois améliorer les taux de guérison et baisser drastiquement le coût des systèmes de santé. Un troisième et dernier exemple est lié à la sécurité routière. Les automobiles sont équipées depuis longtemps de capteurs qui permettent de connaître l'état physique du véhicule (niveau d'huile, d'eau, température du moteur, etc.). Beaucoup de ces informations peuvent maintenant être communiquées à distance. De plus, l'ajout volontaire d'autres capteurs destinés à analyser le style de conduite (vitesse, nombre et intensité des freinages, distances parcourues, etc.) permet à des compagnies d'assurance de proposer des contrats de type *pay as you drive*. Généralisé à d'autres aspects de la vie des assurés, ceci préfigure le passage d'un modèle de tarification forfaitaire à un modèle fondé sur le mode de vie. Là encore, l'ambition des acteurs est la même : capitaliser sur le BD pour passer à une relation clients nourrie par des modèles prédictifs.

Le Big Data, source d'un marketing prédictif en réalité peu innovant

Ce passage à un marketing prédictif nourri par le BD, pour séduisant qu'il puisse paraître à certains, n'en pose pas moins de nombreuses questions. En premier lieu, est-ce une approche si nouvelle que cela ?

« Le marketing prédictif regroupe les techniques de traitement et de modélisation des comportements clients qui permettent d'anticiper leurs actions futures à partir du comportement présent et passé »¹⁴. La prédiction se fait par des algorithmes spécifiques qui permettent, *ex ante*, de proposer les meilleures offres possibles, au bon moment, par le bon dispositif relationnel. Le marketing prédictif est surtout utilisé par de grandes entreprises soucieuses d'améliorer le ROI de leurs actions digitales.

La prédiction est fondée sur des données et sur des probabilités de réussite et exclut, en principe, tout recours à l'intuition pour choisir des facteurs *a priori* susceptibles d'identifier des segments pertinents (e.g. l'âge pour les vêtements de mode). Il s'agit de « faire parler la data »¹⁵ pour en sortir des profils de consommateurs à qui l'on peut s'adresser de manière très ciblée. Cela revient à remplacer la pêche au filet que constituent la publicité de masse ou les offres promotionnelles indifférenciées, par la pêche à la ligne où la taille de l'hameçon, la résistance du fil et la nature de l'amorce sont spécifiques au poisson ciblé.

Cela étant, la nouveauté ne réside pas dans la nature même des calculs effectués. Ces « algorithmes » prédictifs sont utilisés depuis longtemps : analyses discriminantes et modèles de régression (régression logistique, modèles logit/probit/tobit multinomiaux, analyse de survie, etc.) au mieux enrichis d'algorithmes de lissage exponentiel¹⁶. Parfois, de simples corrélations voire des tris croisés permettent de repérer des profils de comportements similaires. Ce qui est nouveau est l'accès beaucoup plus facile qu'auparavant à des données de géolocalisation et à des bases de données publiques, grâce au développement de l'Internet et des *smartphones*¹⁷. Le comportement

14/ Source : <http://www.definitions-marketing.com/definition/marketing-predictif/> ; consultation du 27 avril 2016.

15/ Source : <http://www.e-marketing.fr/Thematique/digital-data-1004/Big-Data-10021/Breves/Tribune-Marketing-predictif-quand-Big-Data-anticipe-actions-consommateurs-252194.htm> ; consultation du 27 avril 2016.

16/ Le lissage exponentiel consiste à tenir compte de l'écart à la prévision sur des périodes passées pour améliorer la performance future d'un modèle.

17/ Source : <http://www.conseilsmarketing.com/e-marketing/le-marketing-predictif> ; consultation du 27 avril 2016.

des analystes s'est également modifié : le volume de données permet une approche « sérendipitaire » au sens où des corrélations parfois contre-intuitives sont considérées comme une réalité à analyser au lieu d'être rejetées au nom d'une possible approximation d'échantillonnage.

Ces méthodes ne sont pas « intelligentes ». Elles se contentent de repérer des variations dans les comportements passés pour en inférer une probabilité de comportement futur. Mais la masse de données rend les tests statistiques très puissants, au sens de Cohen (1988), et permet de rendre significatives des corrélations dont le caractère fallacieux ne peut être mis en évidence que par l'intelligence ou au moins la compétence humaine. Si les ventes de bière et de lunettes de soleil sont corrélées faut-il en conclure que l'alcool accroît la photosensibilité des yeux ? Non, il y a probablement un épisode de beau temps chaud qui vient de s'installer. Est-il nécessaire de disposer de pétaoctets de données pour savoir qu'un jeune couple trentenaire en train d'accéder à la propriété va très probablement consommer du matériel de bricolage, de l'ameublement, de l'équipement de cuisine, de l'électronique de salon, etc. ? Non. On sent poindre derrière cette escalade technique la volonté de sécuriser toute décision dès lors qu'elle implique un coût, en faisant l'hypothèse que remplacer l'homme par une machine est susceptible d'apporter la sécurité espérée.

Il s'agit bien de *data mining* pour l'instant. Le « *machine learning* » pose d'autres questions, abordées plus loin. Dans le *data mining*, La décision humaine est encore nécessaire. Certains pensent que seuls des spécialistes de systèmes d'information et/ou de statistiques sont nécessaires et pourraient ainsi remplacer les *marketers*. Mais de nombreux spécialistes s'accordent à dire que le profil souhaité, au contraire, est bien un *marketer*, dont les compétences en statistiques seraient renforcées¹⁸. Pour Provost et Fawcett (2013, p. 52), par exemple, « la science des données implique bien plus que des algorithmes de *data mining*. Des *data scientists* efficaces doivent être capables d'aborder les problèmes managériaux du point de vue de la donnée ». De ce point de vue, les managers rejoignent les chercheurs : « (...) les interventions des *data scientists* ne remplaceront ni les intuitions ni l'inspiration du marketing pour proposer des offres pertinentes aux consommateurs. Au-delà des moyens technologiques, le marketing ne doit pas non plus négliger son investissement relationnel dans les réseaux sociaux, et encore moins la transparence dont il doit faire preuve dans l'utilisation des données client. Ces "algo" n'ont pas non plus encore les capacités à traduire l'éventail des nuances du comportement humain, ou à créer des dictionnaires normés facilement exploitables »¹⁹. Ajoutons que, parce qu'ils s'appuient uniquement sur des comportements passés, ces algorithmes sont, par nature, incapables de prédire des ruptures, des changements brusques bouleversant la vie des consommateurs. Par ailleurs, une information ne serait-ce que partiellement incomplète ou erronée (« Les gens mentent », dirait *Dr House*), est potentiellement génératrice d'aberrations. Nous recevons assez fréquemment, de la part d'*Amazon*, des publicités pour des ouvrages que nous venons juste d'acheter. C'est une erreur mineure mais qui peut prétendre en connaître, à terme, les effets délétères sur l'image d'*Amazon* ? Une erreur beaucoup moins amusante consisterait à trouver dans le courrier d'un parent décédé une offre pour une convention obsèques.

En dehors de la capacité qu'il offre à disposer de volumes considérables de données hétérogènes et à les traiter en temps quasi-réel, le BD, cantonné à cette dimension prévisionnelle n'est donc en rien innovant et ne saurait remplacer l'intuition d'un « vrai » spécialiste de marketing. Par contre,

18/ Vaste question, tant les formations en marketing ont « abandonné » le terrain face aux réticences des étudiants confrontés aux matières quantitatives.

19/ Source : <http://www.latribune.fr/opinions/tribunes/20140612trib000834781/marketing-predictif-remede-miracle-ou-poudre-aux-yeux-.html> ; consultation du 28 avril 2016.

son évolution prévisible pose un certain nombre de questions d'ordre éthique abordées dans la partie suivante.

Big Data et marketing prédictif : réel danger pour une pratique humaniste du marketing ?

La thèse défendue ici est que le BD enrichi d'algorithmes de type « *machine learning* » représente au plan éthique un danger potentiel, incompatible avec une vision humaniste du marketing.

Big Data et marketing prédictif : des enjeux éthiques

En dehors des problèmes classiques d'intrusion dans la vie privée, déjà présents, l'intelligence artificielle (IA) couplée aux nanotechnologies amène à poser la question de la liberté de décision. L'utilisation croisée des *cookies*, des puces GPS intégrées dans les *smartphones* (et/ou dans d'autres types d'objets connectés), des cartes de crédit et des caméras de surveillance urbaine permet aujourd'hui de savoir presque tout ce que chacun fait, quand et où. Le combat titanesque, au sens mythologique, de *Samaritain* contre *La Machine* dans *Person of Interest*²⁰ est là pour nous rappeler deux choses. D'une part, la notion de vie privée n'a plus grand sens dans un tel monde. D'autre part (mais pour combien de temps ?), il y a nécessairement une main humaine derrière les machines et celle-ci agit pour le meilleur ou pour le pire.

Là non plus, il n'y a rien de bien nouveau. Depuis des millénaires, les régimes politiques s'évertuent à glaner des informations sur les citoyens. Les mieux intentionnés diront que c'est pour répondre aux aspirations des citoyens en matière de services publics, les moins bien intentionnés le feront dans une logique de contrôle totalitaire au nom de telle ou telle idéologie. On retrouve au fil des siècles des citations qui expriment le lien évident entre la connaissance d'un objet et le contrôle que l'on peut avoir sur cet objet²¹. Ce qui change, aujourd'hui, est évidemment l'échelle à laquelle cette collecte peut se faire. Entre les quatre millions de fiches de la STASI, pour une population de seize millions d'habitants et les écoutes de la NSA, il y a un changement d'envergure d'ordre quantique²².

Il est par conséquent logique que des entreprises souhaitent collecter de l'information sur leurs clients, leurs prospects et les consommateurs en général. Les problèmes de « *privacy* » posés par cette collecte sont d'une double nature. En amont, la collecte elle-même soulève des questions d'ordre moral ou de légitimité. Le libre arbitre est-il toujours possible lorsqu'en tant que citoyen ou consommateur on a une telle épée de Damoclès au-dessus de la tête ? En aval, ce sont les actions marketing menées par les entreprises et leurs marques suite au traitement de ces informations qui posent un problème majeur d'intrusion forcée dans la vie privée. La communication non sollicitée (le SPAM, mais plus généralement toute intrusion publicitaire susceptible de détourner

20/ *Person of Interest* est une série américaine diffusée sur TF1. Un informaticien met au point un algorithme d'IA (appelé « *La Machine* ») qui, couplé à tous les dispositifs de surveillance et de communication actuels (caméras, e-mails, SMS...) permet de prévoir un danger de mort imminent pour un individu donné et ainsi lui sauver la vie. Une organisation politique vole cet algorithme et le transforme (en le baptisant *Samaritain*) pour, sous couvert de prévoir les menaces terroristes imminentes, prendre le contrôle sur le gouvernement américain et, par ricochet, sur l'ensemble de la société. Les deux IA finissent par se combattre par humains interposés, un combat titanesque allégorie de l'opposition entre le bien et le mal.

21/ Par exemple dans le domaine militaire « Qui connaît son ennemi comme il se connaît, en cent combats ne sera point défait » (Sun Tzu, V^e siècle avant J.-C.) ou philosophico-scientifique « Savoir, c'est pouvoir » (Francis Bacon, 1597).

22/ On ne connaît pas exactement le nombre de documents extraits des serveurs de la NSA par le seul Edward Snowden, mais il est estimé à plusieurs millions.

l'attention d'une tâche en train d'être accomplie) en est la forme la plus visible. Face à une telle asymétrie d'information, peut-on encore parler de « relations équitables » dans l'échange, au sens de la définition récemment proposée par l'afm ? (Le Nagard et Giannelloni, 2016, p. 6-7)

Heureusement, la médaille a son revers, plutôt favorable aux consommateurs dans les pays où l'Internet n'est pas contrôlé ou bloqué d'une manière ou d'une autre. L'utilisation croissante d'*ad-blockers* et de logiciels de cryptage montre que les utilisateurs du Web ont pris conscience de l'importance de la confidentialité de leurs données et de la protection de leur vie privée. Cela dit, les consommateurs ne sont pas à une contradiction près : d'un côté ils adoptent des conduites parfois irréfléchies en matière de divulgation de données personnelles et, de l'autre, ils sont très inquiets de l'usage qui peut en être fait. Mais, en s'appuyant sur la théorie du *privacy calculus* (Laufer et Wolfe, 1977), Lancelot-Miltgen et Lemoine (2015) montrent, par exemple, que dans certaines conditions les consommateurs considèrent que donner leurs informations personnelles à une entreprise peut avoir de la valeur pour eux.

Mais face à ces stratégies de protection de la part des consommateurs, et en particulier des internautes, les outils de l'IA (principalement le traitement du langage naturel, la reconnaissance des formes et le *machine learning*) sont déjà très présents, sans forcément que l'on s'en rende compte. On les retrouve dans tout ce qui permet d'identifier des voix et des visages ou de répondre à des questions comme *Siri d'Apple* ou *Echo d'Amazon*. L'IA progresse vite et sort désormais de son giron d'origine (*Google, Apple, Facebook, Amazon*). Elle sera au cœur de l'automobile autonome de demain, avec la commande vocale comme interface privilégiée. *IBM*, avec *Watson*, propose également des solutions de relation client innovantes, dont l'utilisation se développe. Par exemple, *Under Armour* propose aux 160 millions de membres de sa communauté une application qui collecte les données fournies par les vêtements connectés de la marque et propose ensuite des conseils de *coaching* adaptés à l'usage de chaque individu. La marque est non seulement capable de personnaliser ses recommandations en fonction des habitudes de chaque sportif, de ses objectifs, de ses paramètres physiques ou de son régime alimentaire, mais également capable d'apprendre des réactions du sportif à chaque recommandation, de son vocabulaire ou de son humeur²³.

Le danger potentiel majeur de l'IA est paradoxalement sa capacité d'apprentissage. *SIRI* s'enrichit quotidiennement des millions de questions qui lui sont posées dans le monde et devient maintenant un réel assistant personnel capable de traiter des problèmes complexes dans un registre croissant de situations différentes²⁴. *Amazon* a déposé fin 2013 un brevet pour un système permettant de prévoir les commandes futures des clients et ainsi d'anticiper leur livraison, avant même que la commande ait réellement été passée. Fondé sur les algorithmes de suggestion existants, le système se nourrit de toutes les commandes passées, croisées avec les profils des clients concernés et du temps plus ou moins long à consulter les pages des produits concernés²⁵.

Pour l'instant, ce système ne semble pas avoir été mis en place par *Amazon*. On peut faire l'hypothèse d'une mise au point plus longue que prévue, en raison d'un taux d'échec important. Parmi les problèmes susceptibles de se poser figurent bien entendu ceux d'ordre technique, au premier rang desquels l'envoi d'un livre n'ayant finalement pas été commandé. Mais on peut également

23/ Source : <https://www.linkedin.com/pulse/en-2016-votre-strat%C3%A9gie-marketing-adoptera-julien-bich> ; consultation du 28 avril 2016.

24/ Source : http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?_r=1&scp=1&sq=Big%20Data&st=cse ; consultation du 28 avril 2016.

25/ Source : <http://www.lefigaro.fr/secteur/high-tech/2014/01/21/32001-20140121ARTFIG00232-amazon-veut-expedier-des-produits-avant-qu'ils-soient-commandes.php> ; consultation du 29 avril 2016.

supposer, au moins chez certains clients, que cette façon de procéder induira des réactions affectives plus ou moins négatives et, en tout état de cause, de la réactance psychologique en raison d'un sentiment de perte de libre-arbitre dans l'acte d'achat. Et qu'en est-il des consommateurs qui, par bienveillance ou simple inertie, conserveront un livre dont ils n'avaient pas forcément envie et qu'ils ne liront probablement jamais ? Face à une telle pression, y a-t-il encore échange de valeur entre l'entreprise et le consommateur ou, à tout le moins, cet échange est-il équitable ? Bref, avons-nous encore affaire à du marketing ?

On peut toutefois imaginer que l'objectif des entreprises, lorsqu'elles utilisent ces outils, est d'anticiper les attentes et les désirs de leurs clients pour améliorer et personnaliser autant que possible la relation qu'elles ont avec eux. C'est en tout cas le discours qu'elles tiennent et nous n'avons pas identifié d'éléments objectifs qui permettraient de leur faire un procès d'intention à ce sujet.

Mais, pour parvenir à cette proximité et à cette réponse en temps réel, voire avant même que le client ait exprimé sa demande comme chez *Amazon*, on utilise toutes sortes d'avatars qui, d'une certaine manière, déshumanisent la relation. Et on peut se demander si, à une époque où la quête de sens et de lien est permanente, le développement à grande échelle de la relation client médiatisée par des robots, quelle que soit leur nature, ne contient pas, en germe, les clés de son propre échec. Dans *Face aux feux du soleil*, Asimov décrit un monde où des humains ayant colonisé l'espace ont fait des robots le pilier de leur civilisation. A tel point qu'ils ont développé au fil du temps un mode de vie qui exclut toute forme de contact physique entre eux.

Evidemment, tout ceci peut sembler nous éloigner du marketing. Cela étant dit, un monde dans lequel les adolescents communiquent par SMS d'un bout à l'autre de la table familiale ou dans lequel des collègues partageant le même *open space* communiquent par e-mail n'est-il pas une préfiguration de *Solaria*, le monde décrit précédemment ? Le marketing doit-il être à l'origine ou même simple partie prenante d'un tel monde ?

Big data et marketing prédictif sont-ils compatibles avec une vision humaniste du marketing ?

Il nous semble que la conception du marketing développée dans l'éditorial précédent (Le Nagard et Giannelloni, 2016) est peu compatible avec un développement généralisé du marketing prédictif fondé sur des « robots ». Aussi complexe soit-il, un robot ne peut pas, en tout cas pour l'instant, intégrer toute la complexité des paramètres à la base de la décision humaine. A notre connaissance, aucun robot n'éprouve de sentiments, d'émotions, d'humeurs et autres formes d'affect, dont on sait qu'il influence fortement et souvent de manière très contextuelle la plupart de nos décisions. Ceci est à la base de nombreux échecs de prédiction, dont certains exemples circulent actuellement dans la presse ou sur Internet. Deux exemples sont développés ici.

Le premier relève du *data mining* classique. Un article du *New York Times* datant de 2012 développe de manière approfondie comment *Target*, cinquième distributeur généraliste aux Etats-Unis, exploite les données qu'il possède sur ses clients pour prévoir certains événements de leur vie. La démarche est fondée scientifiquement sur la recherche liée au rôle de l'habitude dans la prise de décision. L'article s'achève sur un exemple développé par le *data scientist* de *Target*. Certes annoncé comme hypothétique, l'exemple montre comment certains achats effectués par une jeune femme pourraient faire en sorte que l'algorithme de prévision en infère une forte probabilité de grossesse. Chez *Target*, ceci déclencherait un envoi automatique d'offres promotionnelles portant sur des articles liés à la grossesse et aux nouveau-nés. Pour des adolescentes ou très jeunes femmes résidant encore chez leurs parents, rien n'empêcherait alors le courrier

d'être intercepté par les parents et ainsi générer des sources de conflits potentiellement graves au sein du foyer²⁶. Dans certains pays, des jeunes femmes sont assassinées par leurs parents pour beaucoup moins que ça.

Le second exemple relève du *machine learning*. La forte concurrence dans ce domaine amène les acteurs du secteur à mettre sur le marché des produits pas encore totalement maîtrisés²⁷. *Tay*, l'IA conversationnelle de *Microsoft* a fait deux apparitions sur *Twitter*, puis a été retirée dans la précipitation parce qu'elle s'est mise, en très peu de temps, à tenir des propos racistes, antiféministes et néonazis, par mimétisme avec des propos de hackers malintentionnés. Ce type d'IA affine ses réponses par accumulation d'expériences et devient ce que la majorité de ses utilisateurs font d'elle. Une IA doit donc être éduquée, mais par qui ? Quelles normes, fondées sur quel type de morale, fixer et quelle autonomie donner à l'IA par rapport à ces normes en cas de choix « Cornélien » ?²⁸ Autrement dit, il faudrait inventer l'équivalent des lois de la robotique développées par Isaac Asimov pour éviter à terme un monde post-apocalyptique à la *Terminator*. Mais cela impliquerait, comme dans le domaine de l'environnement, une réelle coopération au niveau mondial de l'ensemble des acteurs concernés, avec tous les enjeux géopolitiques que cela suppose. Autant dire que cela relève *a priori* de l'utopie.

Dans un contexte marketing, on peut imaginer une IA « mal éduquée » se comporter comme le pire des vendeurs en manipulant des personnes fragiles pour les forcer à acheter des produits ou services inadaptés voire inutiles. Pour éviter ce type de dérive, les entreprises doivent disposer en interne d'une réelle expertise et, par conséquent, recruter des personnes aguerries aux mathématiques et aux technologies d'analyse prédictive. Il ne s'agit ainsi pas, comme certains le recommandent, de « déceler pro-activement les segments de clientèle les plus captifs, une approche évitant à tout jamais le gaspillage publicitaire »²⁹, ce qui serait du « *nonmarketing* » au sens de Gaski (2013), comme nous l'avons développé dans notre précédent éditorial (Le Nagard et Giannelloni, 2016). Ce type de ressource permettrait au contraire de développer des algorithmes conçus pour éviter toute dérive comparable à celles qui viennent d'être exposées.

De fait, le marketing prédictif, par sa dimension déshumanisante et par les comportements potentiellement imprévisibles des algorithmes reposant sur de l'IA, nous semble incompatible avec la vision du marketing développée par Kotler, Kartajaya et Setiawan (2012). Au sens de ces auteurs, le marketing du futur est un marketing des valeurs, qu'elles soient fonctionnelles, émotionnelles ou spirituelles. Le client est un être humain complexe et multidimensionnel. Il est intègre, il a une raison, un cœur et un esprit et sa collaboration est essentielle. En retour, les clients choisissent les marques et produits qui satisfont sa demande marquée de participation, de créativité, de communauté et d'idéalisme.

Le développement de cette nouvelle forme de marketing est poussé par l'ubiquité liée aux technologies mobiles et aux réseaux sociaux, la mondialisation (économique, politique et socio-culturelle), la prise de conscience écologique et les échanges culturels à l'origine de la modification de nos systèmes de valeurs et par l'émergence de la classe des créatifs culturels qui prennent leurs distances avec la consommation et privilégient un retour à la frugalité (Giannelloni et Tissier-Desbordes, 2013).

26/ Source : http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=2&hp ; consultation du 4 mai 2016.

27/ Si tant est que l'on puisse, par définition, maîtriser une intelligence, fut-elle artificielle.

28/ Source : Laurent Alexandre, *Le Monde*, supplément Science&Médecine, 22171, 27 avril 2016.

29/ Source : <http://www.latribune.fr/opinions/tribunes/20140612trib000834781/marketing-predictif-remede-miracle-ou-poudre-aux-yeux-.html> ; consultation du 27 avril 2016.

De nombreuses pratiques liées au marketing prédictif semblent incompatibles avec cette conception du marketing. En tout cas, elles génèrent une forme de résistance, probablement nourrie par de la réactance psychologique, ce qui suggère que les « victimes » de ces pratiques développent une attitude négative à l'égard de l'entité (la marque, le plus souvent) identifiée comme étant la source des pratiques. Par exemple, l'adoption croissante des *Adblocks*, encouragée par leur facilité d'installation et par le faible niveau d'engagement induit, est sans doute une réponse au caractère intrusif de certains modes de publicité. Un autre mode de protection à faible engagement est tout simplement le mensonge, ce qui a pour effet de polluer les flux de données. Les outils VRM demandent, quant à eux, un niveau plus élevé d'engagement mais permettent le contrôle de ses données personnelles. Face à cette réticence des consommateurs, et pour rétablir leur confiance, il est nécessaire de bannir ce qui est mensonger, trompeur, erroné, intrusif, indiscret, ou risqué.

Pour (re)devenir compatible avec une perspective de marketing 3.0, cette action doit venir de la marque elle-même. Elle doit s'engager dans une relation donnant-donnant, d'une part en restituant les données collectées de manière personnalisée sous forme d'alertes ou de recommandations et, d'autre part, en proposant elle-même à ses clients les moyens d'autocontrôler leurs données. Cette réciprocité est un système de paiement symbolique qui permet de maintenir la confiance. Comme le souligne Christophe Benavent, « Ce donnant-donnant équitable appelle aussi à une autre approche des données », vues comme « les inputs d'un processus de service » dans une relation conçue comme une boucle de feed-back, plus que comme de simples informations³⁰.

Conclusion

D'une certaine manière, cet éditorial a abordé sous un autre angle la question de la « digitalisation » du marketing, déjà évoquée dans un précédent numéro (Tissier-Desbordes et Giannelloni, 2014). Ici, nous avons souhaité creuser l'intérêt, certes, mais aussi et surtout les conséquences potentiellement négatives du développement du *Big Data* et de son usage par le marketing prédictif. Les idées développées peuvent se résumer assez simplement : si le marketing ne se conçoit que dans un espace de liberté (Le Nagard et Giannelloni, 2016) alors certaines des voies prises par le marketing prédictif sont potentiellement dangereuses pour le libre-arbitre des consommateurs et il convient, dans une perspective de marketing réhumanisé, que les marques elles-mêmes décident et organisent un usage raisonné et sincère des données collectées dans une logique d'échange équitable de valeur avec les consommateurs. Nous suggérons deux pistes.

En premier lieu, mettre en place une écoute proactive des consommateurs, au travers de la multitude des interactions entre ces derniers et leur sphère sociale, en particulier sur Internet et les réseaux sociaux. Il s'agit ensuite d'utiliser cette information avec une réelle volonté d'instaurer un dialogue transparent avec le client en lui offrant les moyens, par exemple par des outils VRM, de contrôler lui-même le volume et la nature des informations qu'il met à disposition. En donnant au consommateur la possibilité de faire part de ses attentes et de définir le degré de pression marketing qu'il peut tolérer, la marque lui permet de reprendre la main. Les conséquences potentielles en sont le rétablissement de la confiance, l'arrêt des comportements de résistance et, à terme, une plus grande valeur réciproque échangée.

En second lieu, il s'agit de prêter attention aux compétences nécessaires pour instaurer cette relation. Il nous semble incompatible avec la logique développée jusque-là qu'elle le soit uniquement

30/ Source : <https://www.linkedin.com/pulse/privacy-pas-seulement-prot%C3%A9ger-1%C3%A9thique-de-la-christophe-benavent> ; consultation du 29 avril 2016.

par des robots, quelle que soit la forme que prennent ces derniers. Aujourd'hui, les entreprises disposent soit de *data scientists* rompus aux mathématiques, aux statistiques et/ou aux systèmes d'informations soit de *marketers* davantage issus des sciences humaines. Nous pensons que, sauf exception, seuls des professionnels du marketing peuvent comprendre et maîtriser l'indispensable humanité de la relation entre une marque et les consommateurs. Mais il est évident que la maîtrise du *big data* aujourd'hui appelle à repenser les compétences que doivent posséder les *marketers*. En particulier, les formations au marketing, dans les universités comme dans les écoles, ont depuis longtemps expurgé leurs programmes des matières quantitatives, quand elles ne les ont pas complètement supprimées. Nous sommes peut-être allés trop loin dans cette logique et, bien entendu sans affaiblir l'enseignement de la compréhension et de l'empathie à l'égard des consommateurs, ADN du *marketer*, il peut paraître aujourd'hui nécessaire de (re)donner aux futurs professionnels de notre discipline les outils mathématiques et informatiques nécessaires pour traiter et analyser ces flux de données.

Références

- Babinet G. (2015), *Big Data, penser l'homme et le monde autrement*, Paris, Le Passeur.
- Chen H., Chiang R.L.H. et Storey V.C. (2012), Business intelligence and analytics: from big data to big impact, *MIS Quarterly*, 36, 4, 1165-1188.
- Cohen J. (1988), *Statistical power analysis for the behavioral sciences*, Lawrence Erlbaum Associates.
- Dugain M. et Labbé C. (2016), *L'homme nu. La dictature invisible du numérique*, Paris, Plon.
- Gaski J.F. (2013), To serve man: A marketing manifesto (and an article that should not have been necessary), *Journal of Public Policy and Marketing*, 32, 1, 6-17.
- Giannelloni J.-L. et Tissier-Desbordes E. (2013), Entre consommateurs frugaux et consommateurs résistants aux discours écologiques, quelle(s) voie(s) pour le marketing ?, *Décisions Marketing*, Éditorial, 71, juillet-septembre, 5-11.
- Kotler P., Kartajaya H. et Setiawan I. (2012), *Marketing 3.0*, Bruxelles, De Boeck.
- Laufer R. et Wolfe M. (1977), Privacy as a concept and a social issue: a multidimensional developmental theory, *Journal of Social Issues*, 33, 3, 22-42.
- Lancelot-Miltgen C. et Lemoine J.-F. (2015), Mieux collecter les données personnelles sur Internet. Une étude qualitative auprès d'internautes français, *Décisions Marketing*, 79, juillet-septembre, 35-52.
- Le Nagard E. et Giannelloni J.-L. (2016), « Ce n'est que du marketing » !, *Décisions Marketing*, Editorial, 81, janvier-mars, 5-8.
- Mayer-Schönberger V. et Cukier K. (2013), *Big Data. A revolution that will transform how we live, work and think*, New York, Houghton Mifflin Harcourt.
- Provost F. et Fawcett T. (2013), Data science and its relationship to Big Data and data-driven decision making, *Big Data*, 1, 1, 51-59.
- Tissier-Desbordes E. et Giannelloni J.-L. (2014), Repenser le marketing à l'ère du numérique, *Décisions Marketing*, Éditorial, 73, janvier-mars, 5-8.