

Le marketing et les mégadonnées

Dominique CRIÉ

Professeur des Universités, Université de Lille,
Institut d'administration des entreprises,
Lille Economie Management (UMR CNRS 9221)

Sylvain WILLART

PhD, Risk analyst, Advanzia Bank ; Maître de Conférences
Université de Lille



Dans l'entreprise, c'est souvent le marketing qui ouvre la voie à l'adoption de l'analytique¹ puis à la révolution Big Data en raison de son appétence naturelle pour l'analyse et de sa déjà forte expérience dans le traitement statistique des données dites traditionnelles, stockées dans des entrepôts pouvant être déjà de taille conséquente (Erevelles, Fukawa & Swayne 2016). Certes l'analyse porte alors essentiellement sur des données structurées et de façon synchronique, afin d'offrir une vision comportementale généralement sur un événement précis. C'est l'exemple du ticket de caisse ou de la carte de fidélité qui photographie par instants le consommateur dans ce qu'il fait. Avec le Big Data on passe à une vision holistique, le fameux 360°, qui filme l'individu (et non plus le consommateur car aussi en dehors de ses périodes de consommation) pour saisir ce qu'il est, au sein des réseaux sociaux, dans la tonalité de son discours, dans les multiples traces qu'il laisse sur le web, dans les surfaces de vente ou call-centers, chez les concurrents ou encore dans sa mobilité ou ses horaires de vie...

De façon plus technique et toujours du point de vue marketing, ce qui différencie le plus l'approche « données traditionnelles » du Big Data, c'est la notion de temps réel et de vitesse de traitement. C'est cette caractéristique qui va influencer le plus la pratique marketing. Cette notion de temps réel sous-tend en fait trois types d'éléments : la production et la captation des données qui ne s'analysent plus en stock mais en flux continu, le traitement, l'analyse et l'interprétation instantanés de ces flux de données et enfin l'action marketing simultanée qui est déclenchée. Le second critère d'innovation est sans doute la variété. Une autre différence majeure entre Big Data et données traditionnelles est le passage de données de nature transactionnelle et structurée à des données surtout attitudinales et non structurées par le biais des médias sociaux, où les individus partagent des informations personnelles avec des amis et la famille.

Le Big Data peut également s'analyser comme **un nouveau capital ou une nouvelle valeur patrimoniale immatérielle** (Mayer-Schönberger & Cukier, 2013). Les données fournissent un *insight*² comportemental des consommateurs, le marketing traduit ces *insights* en avantages

1. L'analytique combine l'informatique, le traitement des données, la statistique et la modélisation pour adresser des problématiques managériales, industrielles ou sociales.

2. Un « consumer insight » est une attitude, une attente ou une façon d'envisager les choses par les consommateurs, utile pour la définition des politiques marketing et d'innovation.

marché (Bharadwaj et al., 2013). Par exemple, l'analyse en temps réel et de données en streaming (*Event Stream Processing*) ainsi que la prise en compte de la mobilité (terminaux et géolocalisation) sont des accélérateurs de valeur ajoutée (McAfee & Brynjolfsson, 2012). Dans ce sens le Big Data est bien une nouvelle forme de capital.

Quelques utilisations du Big Data en marketing

De toute évidence, le Big Data peut avoir un impact sur tous les domaines du marketing, que ce soit comme nouvelle source de génération d'idées pour le développement/différenciation des produits ou services, de stratégie prix ou de positionnement concurrentiel, de publicité communication ou encore en matière de distribution... Globalement tous les éléments du mix classique³ peuvent bénéficier d'une approche par les Big Data (Fan, Lau & Zhao, 2015).

Le Big Data initie de nouvelles formes de marketing plus interactives (allant jusqu'à l'affichage dynamique interactif comme dans le film de S. Spielberg, *Minority Report*) ayant pour but d'accroître la productivité marketing et le ROMI⁴ ... Le déluge de données disponibles sur le consommateur sonne la fin du marketing de masse. Il rend la segmentation, le ciblage et le profilage des clients plus pertinents, permettant d'opérationnaliser un marketing hyper-personnalisé et contextualisé allant jusqu'à l'anticipation des besoins et désirs. L'ensemble de ces techniques s'inscrivent dans le « *real-time marketing* » ou le marketing programmatique⁵ ou encore *adtech*⁶, qui grâce à une très grande variété de données, de provenance multiple, en temps réel et à l'ingénierie de l'écosystème Big Data (Demchenko, Ngo, & Membrey, 2013), autorise la proposition d'une offre ou d'une communication personnalisée instantanée. Cette personnalisation de plus en plus fine et tendant vers le marketing individualisé est automatisée par des algorithmes utilisant de nombreuses variables comme les données de navigation et celles relatives au contenu de la page web visitée afin de maximiser la probabilité que l'internaute clique sur la proposition mise en avant.

Le Big Data en marketing pour...

- **La formulation stratégique⁷** : L'utilisation du big data dans la formulation stratégique est évidente, même si elle n'est pas encore perçue comme essentielle par la majorité des entreprises à l'exception des GAFAM (Google, Apple, Facebook, Amazon, Microsoft) et autres pure players (Gantz & Reinsel, 2011). Nouveau capital, temps d'avance sur les concurrents, innovativité en matière de produits/services, proximité avec les clients, leurs besoins et désirs, accès à une information en temps réel et géolocalisée..., tous ces éléments influencent la structure organisationnelle dans l'objectif d'améliorer ses capacités d'adaptation et permettent à l'entreprise d'être beaucoup plus agile que ses concurrents.

- **Communiquer** : Les utilisations les plus visibles en matière de segmentation/ciblage/profilage des clients sont multiples (Chen et al., 2012) : moteurs de recommandation, *retargeting*⁸ (e.g. Criteo), « *real-time bidding*⁹ » qui permet à l'annonceur d'acheter, dans un système d'enchères automatisé et en temps réel, puis d'afficher des bannières publicitaires contextuellement à la navigation et au profil de l'internaute. L'efficacité apportée par la nature instantanée du ciblage est l'un des critères majeurs du développement de cette technique publicitaire, d'autant qu'elle permet un suivi de la réaction du client et favorise directement les ventes additionnelles. Cette

3. Le mix marketing de base comporte classiquement les 4P : le Produit, le Prix, la Place (circuit de distribution), la Promotion (publicité)

4. Return On Marketing Investment

5. Marketing dont les actions sont automatisées selon des algorithmes qui « réagissent » en fonction des comportements du consommateur

6. Pour advertising technologies, désigne le secteur technologique de la publicité digitale

7. Méthodologie d'élaboration du plan stratégique

8. Le retargeting est le ciblage secondaire de consommateurs ayant manifesté un comportement donné sur le net.

9. Real-Time Bidding ou RTB : Achat d'espaces publicitaires sur le web, par enchères en temps réel.

nouvelle temporalité des données autorise de nouveaux modes opératoires marketing. Par exemple, le *native advertising* (publicité native) est un mode de publicité en ligne qui se fond dans le media support consulté par l'internaute et qui enrichit l'expérience utilisateur. Le message, reprenant les codes et le design des contenus (souvent sous forme vidéo) est alors perçu comme moins intrusif, plus naturel et attire d'autant plus le clic de l'internaute. Cette technique s'intègre parfaitement à l'organisation en flux de l'information des media sociaux ce qui la rend encore plus naturelle vis-à-vis de l'utilisateur. Elle nécessite d'être, dans la forme et le fond, en cohérence avec les sujets d'intérêts et les attentes des consommateurs et ne pas perturber leur expérience utilisateur. L'affichage publicitaire plus traditionnel (*display*) sous forme de bannière et autres formats adaptés à l'internet (*banner, pop-up, rollover ads...*) bénéficie également du Big Data et de ses algorithmes qui agrègent les données personnelles et analysent les traces de navigation émanant de sites tiers afin de mieux cerner nos besoins et désirs et ainsi affiner le ciblage par la connaissance d'un comportement global en temps réel sur le web.

- **Recibler** : Un autre moyen de déclencher une action positive de l'internaute est le retargeting ou publicité comportementale car elle est déclenchée suite à certains comportements, comme l'abandon de panier par exemple. Grâce à une analyse précise de nombreux facteurs internes et externes autour d'une navigation sur un site, le Big Data, par une connaissance élargie des clients, permet de sélectionner les prospects les plus susceptibles de réagir à une nouvelle exposition à une offre donnée et donc d'accroître le taux de conversion. L'offre est donc reproposée en *display* au cours d'une connexion ultérieure quel que soit le site visité (e.g. Criteo).

- **Personnaliser/individualiser** : Les GAFAM, mais d'autres géants du net également, ont massivement investi dans les outils de caractérisation des clients et donc de ciblage et de personnalisation tant dans le contenu que dans l'expérience utilisateur. On aboutit alors à un univers commercial (et par là une navigation) complètement différencié avec un affichage, une expérience utilisateur, des recommandations, adaptés et différents pour chaque individu. Ces services ou produits ultra-personnalisés sont moteur d'innovation, de satisfaction client et conduisent à la réalisation de plus grandes marges de profits (Barton & Court, 2012).

- **Automatiser** : Les plateformes de données (Data Management Platforms –DMP-), véritables centres névralgiques de la *customer intelligence* qui permettent l'agrégation de multiples flux de données, forment le cœur du marketing automatisé, envoyant les relances ciblées par des algorithmes en fonction des interactions avec les clients ou prospects et par le media (canal) le plus adapté au moment et à la situation de l'individu. Ces techniques sont également le support d'un parcours client omnicanal fluide dit « sans couture » c'est-à-dire sans que le client ne soit gêné par le changement de canal¹⁰ dans son parcours d'achat.

- **Prédire** : Les données internes aux entreprises sont souvent riches car elles peuvent être issues de l'activité des clients et prospects sur un ensemble de media (*owned media*) créés par/ou associés à l'entreprise (blogs et sites affiliés, pages facebook, comptes twitter, apps, YouTube, newsletters, e-mails...etc...). Cette ingénierie des données permet une connaissance affinée des segments et des individus autorisant une optimisation des contenus des messages, des recommandations, des canaux à utiliser et du tempo adaptés au client/prospect. Une modélisation de l'ensemble de ces données par des techniques prédictives renforce la capacité d'anticipation des besoins et désirs des consommateurs et accroît les taux de transformation (Banker, 2014 ; Ritson, 2014).

10. On entend ici les modalités de contact et de vente, généralement interactifs

Des évolutions substantielles en cours

Le Big Data induit une révolution dans de nombreuses activités marketing comme dans la chaîne de création de valeur ou dans l'obtention d'avantages compétitifs mais aussi en termes de comportement des consommateurs.

- La **recherche marketing et les études** sont les premiers départements à être affectés par l'irruption du Big Data. Traditionnellement ce sont les études qui fournissaient aux entreprises une idée de l'opinion et du comportement de leurs clients vis-à-vis de leurs produits. Grâce au Big Data cette information est directement inférée par l'analytique des réseaux sociaux et autres données comportementales (Lycett, 2013). Le *machine learning* entraîne la disparition des modèles, dans une logique de boîte noire, sans que l'on sache exactement quel a été le « chemin » qui a abouti au résultat, éloignant ainsi le chercheur de la compréhension profonde des mécanismes de prédiction et notamment de l'assurance que des variables essentielles au processus de décision ne soient pas oubliées. Un **renversement de paradigme**, bouleversant le raisonnement scientifique traditionnel d'ajustement des données à des théories préconçues du marché vers l'utilisation des données pour construire de nouvelles théories, est en train de se produire, à savoir une évolution du déductif vers l'inductif (Erevelles, Fukawa & Swayne, 2016). Cette démarche rend le chercheur moins dépendant des connaissances existantes et lui permet de se concentrer davantage sur ce qui est inconnu (Sammut & Sartaoui, 2012), mais l'éloigne de la compréhension intime des phénomènes.

- A un autre niveau, **les objets d'attention et les méthodes d'analyse** changent également de nature. En termes de Business Intelligence et d'analytique, l'évolution des pratiques peut se résumer en trois phases (Chen et al., 2012) : la phase 1.0 où les analyses sont **centrées sur les données**, essentiellement *structurées*, collectées et stockées sur un mode **statique** dans des *entrepôts de données*, gérés dans un système de *bases de données relationnelles* et traitées au moyen d'analyses statistiques classiques et du datamining ; la phase 2.0 où les analyses sont **centrées sur le Web** à partir d'un contenu non structuré (95% des données du Big Data), avec des modes de collecte et de stockage **dynamiques** et traitées par des techniques avancées de *web-mining* et d'analyses spatio-temporelles ; la phase 3.0 où les analyses sont **centrées sur les individus, la géolocalisation et la contextualisation**, les contenus issus des mobiles et objets/capteurs connectés et les analyses sont essentiellement de nature spatio-temporelle et contextuelle focalisées sur les interactions [homme-objets] pour rejoindre le paradigme P.O.S. (Personne-Objet-Situation) dans le cadre du marketing expérientiel, où le consommateur n'est pas que rationnel mais en quête d'émotions, d'expériences, de sensations ou encore de lien social.

- Le Big Data ne sert naturellement pas exclusivement à mieux satisfaire les attentes des consommateurs, l'objectif est d'augmenter le volume d'affaires avec eux dans les meilleures conditions possibles pour l'entreprise. C'est par exemple l'objet du **yield management** (tarification en temps réel) et de sa généralisation le **dynamic pricing**¹¹ (tarification dynamique) qui trouvent avec le Big Data un instrument d'optimisation extraordinaire et font basculer la notion de prix vers celle de consentement individuel à payer... une autre révolution induite par le Big Data s'il en est et qui permet à une organisation de mettre en œuvre une stratégie de tarification flexible, réactive et différenciée et de maximiser les profits. Le prix consenti résulte alors d'un calcul complexe fait par le consommateur en fonction de ses besoins instantanés ou anticipés, d'influenceurs ou prescripteurs et de son interaction avec le vendeur. Uber utilise

11. Yield management ou gestion des capacités de service par le prix. Le dynamic pricing c'est la modulation du prix en fonction des conditions du marché et des caractéristiques du client. Le yield management prend en compte des données exclusivement liées au marché et à la demande, le dynamic pricing prend en compte des données liées au consommateur, notamment son consentement à payer qui dépend de la connaissance qu'il peut avoir du marché.

une forme de ce principe (*surge pricing*) pour ajuster parfaitement l'offre et la demande en cas de forte demande et ainsi satisfaire les acteurs des deux versants de son marché, clients et chauffeurs.

- **Le marketing devient créateur de normes sociales.** Par l'observation des comportements individuels, des secteurs entiers de l'économie vont réviser leurs standards et changer de paradigme pour aboutir à la création de nouvelles normes par le marketing. Par exemple, dans l'assurance, le Big Data d'une part et les objets connectés d'autre part permettent d'affiner de façon considérable le calcul du risque au niveau individuel. Allianz propose d'équiper les conducteurs de boîtiers connectés qui permettent de les segmenter en «communautés comportementales» tarifaires (Grandin de l'Eprevier, 2016), ce qui remet fondamentalement en cause le principe de mutualisation des risques, mais qui joue un rôle d'éducation ou de norme pour la partie des automobilistes qui pensaient être de bons conducteurs mais qui n'ont obtenu aucune réduction de prime (e.g. youdrive pour les jeunes conducteurs).

- **Le data driven behavior** (comportement guidé par les données) suppose que le consommateur (ou l'individu) se comporte selon une norme implicite créée par les données et les métriques qu'elles sous-tendent, souvent différente des normes comportementales habituelles et participant à la génération d'autres normes. Dans une logique de comparaison et de feedback social, il existe un effet réflexif des données sur les comportements, un *datadriven behavior* qui mesure le réel pour mieux le transformer, la donnée devient alors un outil ou une extension de soi (e.g. quantified-self, programme MesInfos...).

- **L'intimité biologique sollicitée.** Des multi-capteurs portables permettent aux entreprises de recueillir une toujours plus grande variété de données (Hardy, 2012). Certaines entreprises (e.g. Proteus, Sano Intelligence, MC 10) ont commencé à collecter, stocker et analyser des données physiologiques en temps réel telles que la fréquence cardiaque, l'activité cérébrale et la température du corps, en utilisant des capteurs portables. Ces données physiologiques en temps réel, combinés avec des mesures comportementales telles que les visites en magasin dérivés de l'information géospatiale ou de *beacons*, offrent la possibilité d'inférer des *insights* consommateurs à la fois comportementaux et émotionnels. Jusqu'où aller dans la compréhension du consommateur et est-ce judicieux ? Pourra-t-on prédire les achats d'impulsions, les changements subits de marque ou de produit ? N'est-on pas dans une spirale qui peut aboutir au contraire à un contrôle plus important, de la part du consommateur, sur les données qu'il émet et qui aboutira à une stérilisation du processus?

Les réactions des clients...

Devant la pression publicitaire ou comportementale exercée par ces technologies, les consommateurs se rebellent en utilisant par exemple des outils permettant de bloquer la publicité (e.g. Adblock). Sans que les professionnels le reconnaissent vraiment, un grand nombre de leurs annonces ne sont pas affichées sur le terminal du client ciblé. Des stratégies plus générales d'offuscation consistent à produire de la fausse information de manière à masquer ou rendre ininterprétable l'information pertinente comme par exemple cliquer automatiquement sur toutes les annonces publicitaires ou générer des requêtes aléatoires sur Google (e.g. TrackMeNot) afin que celles réalisées par l'internaute n'aient aucun sens vis-à-vis de ses centres d'intérêts, ou encore effacer toute trace de navigation sur un site (Disconnect). D'autres stratégies sont possibles, le recours à des pseudonymes multiples, le brouillage d'IP (e.g. IP Hider) ou encore le recours aux solutions de VRM¹² (e.g. Onecub) qui constituent un écran entre le commerçant et le client.

12. Le VRM ou Vendor Relationship Management est le pendant du Customer Relationship Management (CRM). Ici le client prend le contrôle de la relation qu'il veut entretenir avec l'entreprise.

Un cadre juridique à écrire

A qui appartiennent les données ? À ceux qui les produisent ou à ceux qui les récoltent ? Derrière cette question se cache celle de la réelle valeur de ces données personnelles et donc celle de leur monétisation, notamment vis-à-vis des entreprises dont le business model est exclusivement adossé aux données et traces du consommateur et qui monétisent les données, les modèles, les actions commerciales et leurs résultats. Aucun dispositif législatif ou réglementaire n'encadre encore la collecte et l'utilisation des données, même si la récente Loi Numérique pose quelques jalons en France. Par contraste, le droit à l'image, (qui n'est autre qu'un ensemble de données ou pixels), permet à quiconque de s'opposer à l'utilisation, commerciale ou non, de son image, au nom du respect de la vie privée. Pour Russel W. Belk (2013) : « les données que nous possédons reflètent l'identité de leur propriétaire, elles contribuent à notre identité, elles sont un bien virtuel ».

Références

- Banker, S., (2014), "Amazon and anticipatory shipping: A dubious patent?" *Forbes* (consulté le 8 mai 2016)
- Barton, D. & Court, D. (2012), « Making Advanced Analytics Work for You », *Harvard Business Review*, Vol. 90, n° 10, p.78-83.
- Belk, R. W., (2013), "Extended self in a digital world", *Journal of Consumer Research*, 40, 3, 477-500
- Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N., (2013), « Digital business strategy: Toward a next generation of insights », *Management Information Systems*, 37(2), 471-482
- Chen H., Chiang R. H. L., & Storey V. C., (2012), "Business Intelligence and Analytics: From Big Data to Big Impact", *MIS Quarterly*, Special Issue: Business Intelligence Research, 36, 4, 1165-88
- Demchenko Y., Ngo, C., & Membrey, P., (2013), Architecture framework and components for the big data ecosystem, *Journal of System and Network Engineering*, pp 1-31
- Erevelles S., Fukawa N. & Swayne L., (2016), "Big Data consumer analytics and the transformation of marketing", *Journal of Business Research*, 69, 897-904
- Fan S., Lau R.Y.K. & Zhao J. L., (2015), « Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix », *Big Data Research*, 2, 28-32
- Gantz J. & Reinsel D., (2011), "Extracting Value from Chaos", International Data Corporation (IDC) IVIEW, June 2011, disponible à : <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>
- Hardy, Q., (2012), « Big data in your blood », *New York Times*, 7 September 2012
- Grandin de l'Eprevier Jade, (2016), « Trier les bons et les mauvais conducteurs, le pari connecté d'Allianz », *Le Monde*, 13.04.2016
- Lycett, M., (2013), « 'Datafication': Making sense of (big) data in a complex world », *European Journal of Information Systems*, 22(4), 381-386
- Mayer-Schönberger, V., & Cukier, K. (2013), *Big data: A revolution that will transform how we live, work, and think*, New York: Houghton Mifflin Harcourt
- McAfee, A. et Brynjolfsson, E. (2012), "Big Data: The Management Revolution", *Harvard Business Review*, Vol. 90, Issue 10, p60-68
- Ritson, M., (2014), « Amazon has seen the future of predictability », *MarketingWeek*, <http://www.marketingweek.co.uk/opinion/ritson/amazon-has-seen-the-future-of-predictability/4009154.article>, consulté le 8 Mai 2016
- Sammut, G., & Sartawi, M., (2012), « Perspective-taking and the attribution of ignorance », *Journal for the Theory of Social Behaviour*, 42(2), 181-200