

Julien Hirth

LE DATA MARKETING



EYROLLES

LE DATA MARKETING

« Le [data marketing], c'est comme le sexe chez les adolescents : tout le monde en parle, personne ne sait vraiment comment le faire, tout le monde pense que tout le monde le fait, donc tout le monde prétend le faire. » Et les chiffres donnent raison à Dan Ariely, auteur de cette métaphore : selon une récente étude conduite par l'institut Morar, 72 % des entreprises n'exploitent pas les données qu'elles collectent, déboussolées face à la complexité technologique, scientifique et organisationnelle introduite par l'afflux massif de données sur les consommateurs.

Bible du data marketing, cet ouvrage propose de manière inédite une méthode robuste pour exploiter le plein potentiel de la donnée.

Il permettra notamment au marketeur moderne de :

- **comprendre les mécanismes de collecte, de stockage et de réconciliation des données ;**
- **maîtriser les principaux outils du data marketing (DMP, CRM 360, marketing automation, écosystème adtech...) ;**
- **démystifier les concepts mathématiques appliqués au marketing (clustering, scoring, marketing prédictif...).**

© Studio Lenoir



Ingénieur des Ponts et Chaussées, **Julien Hirth** est le co-fondateur de Scibids Technology, solution technologique intégrant l'intelligence artificielle au cœur des stratégies d'achat publicitaires.

Il fut auparavant consultant au sein du cabinet Artefact et a accompagné de nombreuses entreprises dans la définition et la mise en œuvre d'une

stratégie data marketing omnicanale efficace.

Il anime également un compte Twitter (@JulienHIRTH) commentant les dernières tendances du data marketing : un bon complément à la lecture de cet ouvrage !

Le data marketing

Groupe Eyrolles
61, bd Saint-Germain
75240 Paris Cedex 05
www.editions-eyrolles.com

Chez le même éditeur :

- *Réussir sa transformation digitale*, Cindy Dorkenoo, Aurore Crespin, Laura Lombardo, Frédéric Klotz et Hamza Moulim
- *Stratégie digitale*, Cindy Dorkenoo
- *Le guide de la transformation digitale*, Vincent Ducrey et Emmanuel Vivier
- *Le consommateur digital*, Nicolas Riou

En application de la loi du 11 mars 1957, il est interdit de reproduire intégralement ou partiellement le présent ouvrage, sur quelque support que ce soit, sans autorisation de l'éditeur ou du Centre français d'exploitation du droit de copie, 20, rue des Grands-Augustins, 75006 Paris.

© Groupe Eyrolles, 2017
ISBN: 978-2-212-56473-0

Julien Hirth

Le data marketing

La collecte, l'analyse
et l'exploitation des données
au cœur du marketing moderne

EYROLLES

The logo for EYROLLES features the brand name in a bold, black, sans-serif font. Below the text is a horizontal line with a small grey circle centered on it, which tapers out at both ends.

Sommaire

Préface	13
Avant-propos	17
La digitalisation de nos vies...	17
... couplée à l'augmentation de la puissance de calcul.....	18
... changent le métier du marketeur	19
Objectif de ce livre	20

PRÉAMBULE

Concevoir son projet Data

Un vaste champ des possibles	24
Spécifique : quel est le but précis poursuivi par le projet ?	25
Un objectif n'est pas un moyen	26
Un objectif n'est pas une finalité	26
Un objectif doit être précédé d'un constat	27
Un objectif doit être simple à comprendre	28
Une méthode efficace pour déterminer un objectif spécifique : l'arbre d'équivalence	28
Mesurable : avec quels indicateurs évaluer la réussite du projet ?	29
Mesurer l'efficacité du projet : les KPI	30
Déterminer les actions qui pèsent le plus : l'attribution	31
Mesurer la pertinence d'un modèle : la matrice de confusion	33
Comparer les gains par rapport aux coûts : le ROI	34
Atteignable : par quels cas d'usages atteindre son objectif ?	35
De la théorie.....	35
... à la pratique	36

Réalisable: quelles contraintes opérationnelles vont se dresser sur votre chemin ?	38
Les données en silo ralentissent les projets.....	38
Les outils en place ne permettent pas la réalisation de tous les cas d'usages.....	39
L'organisation de l'entreprise n'est pas centrée sur la donnée.....	39
Temporellement défini: et maintenant place à l'action!	41
Viser des livrables régulièrement... ..	42
... et s'inspirer des start-ups!.....	42
Sources	44

PARTIE A

Collecter et stocker

Autodiagnostic	46
Chapitre 1 – Faire connaissance avec ses données	49
Comment classer ses données compte tenu de leur variabilité ?	50
Les critères de classification utilisés par les « data scientists ».....	50
Les critères de classification utilisés par les statisticiens.....	52
Les critères de classification utilisés par les responsables informatiques.....	53
Les critères de classification utilisés en publicité.....	53
Les critères de classification utilisés par le marketing.....	54
Les critères de classification utilisés par les juristes.....	56
Quelles données occupent le plus de volume ?	58
Vos données sont-elles utiles ?	59
Fiabilité et qualité des données: le V de « véricité ».....	59
Obsolescence et capacité à être mobilisé en temps réel: le V de « vélocité ».....	61
Exploitable à travers des cas d'usages: le V de « valeur ».....	61
Sources	63

Chapitre 2 – Boîte à outils pour passer de datas éparpillées à des données centralisées	65
Rapatrier ses données : aperçu des principales méthodes de collecte de données et de tracking	66
Le cookie, clé de voûte de la collecte de données web	66
Le Tag Management System (TMS), plus qu'un simple utilitaire	70
L'« API », trois lettres et tellement de possibilités	71
Comment stocker la donnée ainsi collectée ?	73
Structurer la donnée avant de la stocker : la façon « traditionnelle » de procéder	73
Stocker la donnée brute : plongeon dans le « lac de données »	75
SQL vs NoSQL : l'arbre qui cache la forêt	76
Agréger les données autour d'un identifiant unique	77
Un grand nombre d'identifiants se rapportent à un même utilisateur	78
Relier des identifiants entre eux : état de l'art en 2016	80
Sources	85
Chapitre 3 – CRM et DMP : deux outils clés pour dompter la multitude de données	87
Le CRM collecte, stocke et active les données personnelles dans une optique de fidélisation	88
Un CRM peut en cacher un autre...	88
Du CRM au CRM 360 : le grand chantier de la connaissance client	90
La DMP collecte, stocke et active les données anonymes dans une optique d'acquisition	92
Six sources de données alimentent la DMP en cookies	92
Chaque minute, la DMP « classe » des milliers de cookies dans une vaste arborescence : la taxonomie	94
La raison d'être d'une DMP est de créer des audiences...	95
... et de les mettre à disposition des outils d'activation	99
Pas une solution miracle : les limites de la DMP	100
Petite grille d'évaluation pratique pour choisir sa DMP	102
S'organiser dans ses données pas à pas	105

Étape 0: dresser une cartographie de l'existant	105
Étape 1: instaurer un référentiel client unique	105
Étape 2: mettre en place une DMP	105
Étape 3: relier le CRM à la DMP	105
Étape 4: déverser le tout dans un datalake	106
Sources	107

PARTIE B

Extraire de l'information des données

Autodiagnostic	110
Chapitre 4 – Analyse: à la découverte de tendances et de schémas dominants	113
Étape 1 – Découvrir et nettoyer le dataset	114
Étape 2 – Décrire les données	115
Analyser les grandes masses	116
Analyser les évolutions	121
Étape 3 – Comparer les données	124
Comparaison de deux moyennes (ou deux pourcentages)	124
Comparaisons multidimensionnelles	126
Étape 4 – Établir une ou plusieurs segmentations	128
Sources	133
Chapitre 5 – Analyse exploratoire: identifier les liens entre les données	135
La matrice de corrélation identifie les variables reliées entre elles	135
Un peu de théorie	136
Rendre la matrice de corrélation « sexy », c'est possible!	139
Application pratique: étude des corrélations entre ventes et météo	142
L'analyse en composante principale (ACP) réduit le nombre de variables	143
Étape 1 – Création des nouvelles variables	143

Étape 2 – Interprétation des axes factoriels	145
Étape 3 – Projection des individus	146
Le clustering : segmentation en pilotage automatique	147
Le clustering s'appuie sur une notion de distance	148
La méthode de clustering «<hiérarchique>>	149
La méthode de clustering «<non hiérarchique>>	151
Sources	152
Chapitre 6 – Data visualisation et data storytelling	153
Chart chooser : quel type de représentation choisir pour visualiser les données ?	154
Anatomie d'un bon graphique :	
quelques astuces de mise en forme	156
L'échelle	157
Le titre	157
Les couleurs	157
La légende	158
Autres astuces	158
Présentation des résultats sur PowerPoint :	
se concentrer sur l'essentiel	159
Mobiliser et impliquer positivement son audience	159
Structurer sa présentation comme une histoire	160
Vos slides passent-elles le test des cinq secondes ?	161
Le tableau de bord ne présente pas des résultats, il aide à la décision	162
Sources	164
Chapitre 7 – Le marketing prédictif	165
Principes généraux du marketing prédictif	165
Certaines variables sont plus «<prédictives>> que d'autres	166
De nouvelles variables peuvent être créées	168
L'efficacité d'un modèle se mesure et les modèles peuvent être comparés	169
Apprendre, prédire, valider	171
Prédiction explicite à partir des variables :	
les modèles de régression	172

Prédire le gazouillis des criquets	172
Prédire le CA d'un nouveau magasin pour choisir le meilleur emplacement	173
Prédiction d'un état :	
algorithmes de classification	176
Le client va-t-il renouveler son abonnement ?	177
Indications des conversations « tchat » ayant une faible probabilité de satisfaire le client	179
Sources	182

PARTIE C

Activer les données

Autodiagnostic	184
Chapitre 8 – Data et publicité	187
Qu'appelle-t-on « programmatique » ?	187
Il était une fois... la publicité digitale	188
Le programmatique s'impose comme le prochain standard publicitaire	190
Le Real Time Bidding (RTB) change le paradigme de la publicité digitale	193
Avec le RTB, l'écosystème se complexifie	197
La data, au cœur des enjeux de la publicité digitale moderne	201
La data pour créer une audience	203
La data pour optimiser une campagne	210
La data pour personnaliser le message publicitaire	212
La publicité digitale en pratique :	
mesurer la performance en quatre questions clés	214
Où la publicité est-elle diffusée ?	215
Auprès de qui la publicité est-elle diffusée ?	216
Objectif performance ou objectif branding ?	217
Comment se décomposent les coûts ?	218
Sources	219

Chapitre 9 – Marketing direct : du mass marketing au marketing one to one	221
Le concept de « marketing automation » bouleverse le mediaplanning traditionnel	222
Cinq typologies de trigger dans votre boîte à outils pour automatiser le marketing	223
Le profil client est au cœur du processus de personnalisation	226
Individualiser aussi la pression commerciale	228
Borner le nombre de contacts: l'approche « droit au but » ..	228
Modéliser la pression marketing: l'approche « attributive »	230
Jouer la transparence: l'approche « win-win »	230
Le marketing direct one to one en pratique : cinq exemples de campagnes e-mail incontournables ..	232
La relance de panier abandonné	232
La campagne post-navigation	232
La campagne de bienvenue	233
La campagne post-sell	233
La relance des clients inactifs	233
Sources	234
Chapitre 10 – Data et commerce physique : la révolution apportée par le mobile et les objets connectés	235
Comprendre la géolocalisation : grille d'analyse	236
Géolocalisation indoor ou outdoor ?	236
Géolocalisation en mode « pull » ou mode « push » ?	239
Géolocalisation continue ou discrète ?	240
Géolocalisation statique ou bien géolocalisation temps réel ?	241
Comment tirer parti de la donnée de géolocalisation ?	242
Le Drive-to-store pour attirer du trafic en point de vente ..	242
Le Mobile-in-store pour guider le consommateur	243
La création de nouveaux services à valeur ajoutée	244
Des opérations de communication 3.0	244
L'analyse des parcours pour améliorer la productivité	245

Les limites de la géolocalisation : les questions clés à se poser	246
La couverture d'audience est souvent faible	246
L'accord du consommateur est indispensable	247
Les contraintes techniques	248
La donnée de géolocalisation en pratique : cibler les amateurs de tennis de quatre façons différentes	249
Sources	251
Conclusion	253
Réponses aux autodiagnostic	256
Questions partie A – Collecter et stocker	256
Questions partie B – Extraire de l'information des données	258
Questions partie C – Activer les données	260
Glossaire	263
Les sociétés et solutions spécialisées dans la data	275
Cabinets/agences conseils	275
Agences marketing/publicitaires	275
Adtech	276
Martech	277
Autres	278
Remerciements	279
Index	281
Table des figures	287
Table des encadrés	293

Préface

Le digital nous a changés. En nous simplifiant la vie, il a complètement modifié notre rapport aux marques. Alors que nous fêtons seulement cette année le 25^e anniversaire du World Wide Web, nous avons en l'espace de très peu de temps radicalement transformé des siècles de rapports marchands.

Fini l'attente et la patience. Amazon prime nous livre aujourd'hui en moins de 24 heures. Fini les mauvaises affaires et autres manipulations : les comparateurs et les commentaires clients nous rassurent et nous inspirent. Fini l'angoisse du voyage et de la lecture de la carte routière: smartphone en poche, nous sommes invincibles. Le digital a exacerbé les traits de caractères de notre génération Y. Ultra connectés et multi-identitaires, nous sommes impatients, infidèles, voire méchants et rancuniers sur les réseaux sociaux s'il le faut.

Le marketing a dû s'adapter en retour. Le célèbre spot de 30 secondes sur TF1 a fait place à des campagnes digitales ultra segmentées, scénarisées et ciblées. Les annonceurs les plus matures produisent des milliers de contenus différents à chaque campagne pour coller aux plus proches désirs de leurs clients. Ces contenus sont poussés par des algorithmes en fonction de multiples paramètres individuels. Netflix le fait notamment avec ses propres séries. En cas de campagne mal ciblée, la réponse du diginaute est sans appel: toute publicité intrusive ou non pertinente est immédiatement sanctionnée par l'installation d'Ad-block. Les taux d'installation flirtent aujourd'hui avec les 30 % en France.

Le fil directeur de cette évolution est la donnée. C'est cette ressource que certains chercheurs qualifient d'« or noir » du XXI^e siècle, et qui permet aux marques de faire le lien entre toutes nos identités, physique et digitales. C'est elle qui leur permet de savoir proposer intelligemment un matelas confortable à qui vient d'acheter une poussette, ou encore permet de détecter très en avance les signaux « faibles », avant-coureurs d'une résiliation de contrat.

Professionnels et passionnés du marketing, pourtant aguerris aux toutes dernières techniques, nous sommes sans cesse impressionnés et devons constamment nous remettre en question pour suivre au mieux les innovations toujours plus ambitieuses mises au point par le milieu du data driven-marketing.

Google, dont la mission ultime est d'organiser et de rendre accessible toute l'information au monde, se donne les moyens de ses ambitions. En achetant, en 2014, la start-up britannique Deepmind, le géant mise sur les techniques les plus pointues d'intelligence artificielle, sans nul doute pour les appliquer au domaine qui constitue sa première source de revenus: la publicité. Deepmind n'a pas attendu longtemps avant de faire ses preuves en battant cette année le champion du monde du jeu, considéré comme l'un des derniers bastions où l'humain domine la machine: le jeu de go, longtemps considéré comme impossible à maîtriser par un algorithme (le nombre de combinaisons jouables est plus de 1 050 fois plus grande que les échecs: même une machine surpuissante est loin de pouvoir calculer tous les coups gagnants). AlphaGO a pourtant réussi l'impossible et a vaincu au cours d'un match à suspense le champion Lee Sedol 4 à 1.

Microsoft n'est pas en reste et a été récompensé par le prix de la data et de la créativité cette année à Cannes. L'algorithme de la société a passé au crible 346 tableaux de Rembrandt (épaisseur des couches de peinture successives, espacement des yeux, position du nez, forme des visages, etc). L'ensemble des caractéristiques a ensuite été reproduit sur une toile inédite, imprimée en 3D, créant le fascinant «Next Rembrandt» sur une toile constituée de plus de 148 millions de pixels. L'illusion est parfaite: les plus grands experts en histoire de l'art ont reconnu qu'il était impossible de différencier le nouveau portrait d'un portrait de Rembrandt.

Jusqu'où ira-t-on dans l'utilisation de la donnée appliquée au marketing ? Le monde de demain sera sans aucun doute un monde de marketing individuel ou «people based». Le contenu autogénéré et absolument unique qui nous sera proposé sera probablement complètement différent du contenu de notre voisin. À l'image de Facebook qui nous propose autant de versions de son newsfeed qu'il y a d'utilisateurs sur le réseau social, les contenus poussés par les marques seront complètement adaptés et pertinents grâce à la masse de données dont elles disposent sur nous.

Nous sommes des fervents défenseurs de la donnée et de son utilisation dans le marketing, persuadés qu'elle nous simplifiera la vie, poussant à transformer le marketing en un service, plutôt qu'en une contrainte. Il est évident que tout cela a un prix. Sans pouvoir ici détailler tous les enjeux, deux risques majeurs me viennent à l'esprit.

La perte du jardin secret. Même si la juridiction sur les données personnelles a intelligemment évolué en ce sens en France (loi Lemaire) ou en Europe (règlement Européen sur la protection des données personnelles), il ne faut pas se tromper de combat. Les GAFAs connaissent déjà TOUT sur nous. Pour prendre l'exemple de Google: nos recherches sont sur Chrome, nos échanges sur Gmail, nos passions sur YouTube, nos déplacements sur Maps ainsi que chez tous les possesseurs d'un téléphone Android. Toutes ces données sont disponibles sous un même identifiant commun: le GoogleID. Le géant a techniquement la possibilité de croiser, recouper à loisir ces données pour établir une connaissance individuelle à laquelle même nos plus proches n'ont pas accès. Il va être quasiment impossible de protéger nos données face à ces géants technologiques sans restreindre les services auxquels nous sommes devenus accros, et nous transformer en ermites.

Le syndrome de panurge. On pourrait croire que la personnalisation exacerbe les goûts et les individualités, mais c'est en fait tout l'opposé qui se produit! Les chercheurs ont montré qu'au lieu de favoriser un effet «long tail» qui aurait pour effet de pousser des contenus de niche, les moteurs de recommandations propulsent les blockbusters. Paradoxalement, le choix diminue alors qu'il devrait être démultiplié. Nous devenons passifs. Plus de curiosité ou d'éveil dans la mesure où c'est maintenant l'algorithme qui choisit à notre place. Un sursaut d'orgueil est attendu: notre créativité et notre individualité se doivent de prendre le pas sur la machine et non l'inverse.

J'ai connu Julien en tant que jeune et brillant consultant au sein de la société Artefact qui est devenue la référence dans le domaine de l'innovation marketing data driven. En plus de ses activités quotidiennes, Julien mettait toute son énergie à construire et à mettre en forme les savoir-faire que nous avons constitués pour former nos jeunes recrues. À une compétence très forte dans nos métiers s'associait très naturellement un talent pour simplifier les concepts les plus complexes.

À ce titre, ce livre éclaire et détaille bien des aspects que tout consultant en marketing ou directeur marketing moderne doit maîtriser. C'est un must-have pour tout professionnel, étudiant ou curieux s'intéressant au monde de la donnée dans son ensemble. Sans être dogmatique, il est pragmatique, pédagogique et apporte de la clarté sur les solutions data driven en vigueur actuellement. Le lecteur intéressé par l'ensemble des problématiques de la donnée dans le marketing pourra (devra!) néanmoins compléter

cette lecture par des ouvrages traitant des impacts de la data sur l'entreprise et les transformations organisationnelles nécessaires.

Vincent LUCIANI
Fondateur du cabinet Artefact

Avant-propos

«Or noir», «pétrole brut», «nouvelle révolution industrielle»... les superlatifs ne manquent pas pour décrire la «data». L'analyse et l'exploitation de la donnée ne sont pourtant pas nouvelles. Si les calculatrices à boules de l'Antiquité ou l'horloge Aurorex enregistrant le temps de conduite des flottes de véhicules dans les années 1920 font figure d'anecdotes, le datamining¹ et le scoring² client existent dans les banques et les assurances depuis les années 1970 afin de prédire les «mauvais payeurs», tandis que la business intelligence (BI) a fait les beaux jours des éditeurs dans les années 1980 et le Customer Relationship Management (CRM) ceux des années 2000. C'est également à la fin du xx^e siècle que se démocratisent les data centers³. Alors pourquoi soudainement le terme «data» est-il dans toutes les conversations et que plus de cinquante contenus sont publiés chaque heure sur le sujet en France ? Qu'est-ce qui a tant changé et justifie un tel engouement frisant parfois la folie collective ?

LA DIGITALISATION DE NOS VIES...

La première nouveauté, c'est la digitalisation de notre quotidien. Depuis que vous lisez cette introduction, plus de 250 000 tweets ont été échangés, plus de deux millions de requêtes ont été analysées par Google, plus de trois millions de «likes» Facebook ont été attribués, plus de cent heures de nouvelles vidéos YouTube ont été chargées et plus de 150 000 vidéos Netflix lancées ! Au total en moins d'une minute de l'année 2015 ont été créées autant de données que toutes celles depuis le début de l'humanité jusqu'à 2003 et 90 % des données disponibles dans le monde ont moins de

1. Littéralement «fouille de données», le datamining est un processus faisant appel aux techniques statistiques et mathématiques permettant d'extraire des informations commercialement pertinentes et de modéliser des schémas d'informations cachés au sein de grandes bases de données.

2. Traitement mathématique consistant à attribuer à un individu une probabilité de comportement futur. Les premiers modèles de scoring ont été mis en place dans les banques dans les années 1970 afin d'estimer le risque associé aux crédits.

3. Lieu physique dans lequel sont regroupés les éléments concrets (ordinateurs, serveurs, etc.) constituant le système d'information de l'entreprise.

deux ans¹! Un afflux massif de données qui change complètement le paradigme qui prévalait dans les années 1980 : avec un volume de données à traiter par une entreprise qui double tous les dix-huit mois², le facteur différenciant d'une société aujourd'hui ne réside plus tant dans le fait d'avoir des données que d'être capable de les analyser et de les transformer en information! «La data est abondante, donc ce qui est utile et rare, c'est la capacité à l'exploiter et la rendre opérable», assure ainsi Samir Amellal³, Chief Digital Officer de Publicis.

... COUPLÉE À L'AUGMENTATION DE LA PUISSANCE DE CALCUL...

Justement, **être capable de traiter rapidement et à moindre coût les gros volumes de données est le deuxième changement majeur** à la source de l'essor du data marketing. Alors que les algorithmes statistiques des années 1970 n'étaient pas souvent applicables ou rentables à mettre en place faute de puissance de calcul à disposition, leur mise en production est désormais possible et profitable. Cette évolution tient principalement à quatre phénomènes : la loi de Moore, le calcul distribué, l'open source et le cloud.

La loi de Moore est une façon savante de dire que la puissance des processeurs double tous les dix-huit mois depuis l'invention de l'ordinateur!

Le calcul distribué signifie qu'au lieu de recourir à un énorme serveur très onéreux pour réaliser un calcul complexe, celui-ci est découpé en plusieurs sous-calculs plus simples, réalisés par des serveurs plus conventionnels. Comme pour les comptes d'entreprise, ces micro-calculs sont ensuite consolidés pour trouver le résultat du calcul initial.

L'open source désigne la gratuité des logiciels construits par des «communautés non lucratives». La plupart des architectures Big Data sont composées de tels logiciels en libre accès, le plus célèbre d'entre eux étant Hadoop.

1. *Les données numériques : un enjeu d'éducation de citoyenneté*, Journal officiel de la République française, séance du 13 janvier 2015.

2. Bruno Teboul et Jean-Marie Boucher, *Le Marketing absolu*, éditions Kawa, 2013.

3. Guillaume Serries, «Compétences, gouvernance, vision : SAS détaille les freins du Big Data en France», 5 novembre 2015, zdnnet.fr.

Enfin, **le cloud** a fait diminuer drastiquement le coût nécessaire au traitement des données¹. Sur le modèle de l'économie collaborative, il devient possible à n'importe quelle start-up dans un garage de louer en temps réel autant de machines que nécessaire à Google ou à Amazon!

... CHANGENT LE MÉTIER DU MARKETEUR

Fini les «4 P»² si chers aux professeurs et aux livres du xx^e siècle! Au placard les études et panels annuels! **76 % des marketeurs considèrent que le marketing a davantage changé en deux ans que lors des cinquante dernières années**³.

Le digital et les progrès informatiques ont en effet bouleversé le parcours d'achat du client. D'un être passif à séduire à grand renfort de messages commerciaux, le consommateur est devenu informé et critique (comparaison des prix et avis consommateurs en temps réel), mobile et connecté (multitudes de points de contact avec la marque), méfiant et lassé (inondé de messages commerciaux – plus de 300 par jour⁴). Une vraie diva insaisissable et exigeante, au parcours d'achat de plus en plus complexe et de moins en moins linéaire!

Pour rester compétitif, il devient dès lors nécessaire de centrer le marketing sur le client et non plus sur le produit⁵, ce qui signifie concrètement reconnaître le client à chaque point de contact et personnaliser l'interaction, à la fois par rapport à l'individu et à son parcours.

Et pour relever ce défi, le marketeur moderne a besoin de données: des données sur l'individu en tant que tel («une adolescente de 14 ans»), sur son comportement («cet individu est venu trois fois sur le site cette semaine, c'est un prospect chaud»), sur ses centres d'intérêt... La planification média et l'analyse stratégique laissent dès lors place à d'autres questions au sein des réunions entre professionnels du marketing: quelles données peuvent aider à augmenter la conversion de l'étape une à deux dans le parcours

1. Le prix du mégaoctet était estimé à 300 \$ en 1980. Il est estimé aujourd'hui à moins de 0,0002 \$ (source: *Les données numériques: un enjeu d'éducation de citoyenneté*).

2. Produit, Prix, Place, Promotion.

3. Digital Marketing Symposium d'Adobe, 2014.

4. Arnaud de Baynast et Jacques Lendrevie, *Publicitor*, 2014 (8^e édition).

5. Passage d'une logique «product centric» à une logique «client centric».

d'achat du client? Comment les collecter? Que signifie vraiment cette analyse remise par le statisticien? La solution imaginée est-elle concrètement réalisable? Comment prouver l'efficacité de son idée?

Plus scientifique, plus technologique, le marketing moderne se fait aussi plus créatif: auparavant dépendant du système d'information (SI), l'apparition d'outils en mode «Software as a Service» (SAAS)¹ libère le potentiel opérationnel du marketing. Une liberté qui a toutefois un prix: le changement d'habitudes de travail et l'acquisition de nouvelles compétences et de nouveaux savoir-faire.

OBJECTIF DE CE LIVRE

«La data, c'est comme le sexe chez les adolescents: tout le monde en parle, personne ne sait vraiment comment le faire, tout le monde pense que tout le monde le fait, donc tout le monde prétend le faire.» Cette métaphore culottée de Dan Ariely, professeur de psychologie et d'économie comportementale, exprime le flou qui entoure aujourd'hui le phénomène «data». Le data marketing reste en effet encore très obscur pour beaucoup de directions d'entreprises, quasiment «magique». Les attentes sont souvent élevées, au niveau du degré de «magie» mis dans cette nouvelle approche et ces nouvelles technologies «miracles». Bercés la nuit par les promesses de la data («le bon message au bon moment à la bonne personne») et heurtés le jour par les difficultés opérationnelles (aussi bien techniques qu'organisationnelles), les dirigeants d'entreprises ont aujourd'hui le tournis. 20 % seulement d'entre eux, interrogés lors d'une récente étude de la société Bizo², estiment avoir réussi à prendre le tournant de la data, tandis que 72 % des entreprises n'exploitent pas les données qu'elles collectent selon une étude conduite en juillet 2015 par l'institut Morar pour le compte du constructeur Purestorage³.

1. Logiciel mis à disposition à distance par un fournisseur et accessible par le biais d'une URL web. Le logiciel est loué, au mois ou à l'usage. Les mises à jour sont automatiques.

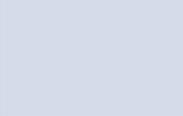
2. Start-up rachetée 175 millions de \$ par LinkedIn en 2014.

3. *Big Data's Big Failure: the struggles businesses face in accessing the information they need.*

Face à ces constats, ce livre a pour vocation d'aider les directions marketing à :

- mettre en œuvre une stratégie data robuste et à industrialiser un ensemble de cas d'usages marketing « data driven » ;
- avoir une vision holistique des données à disposition et un descriptif des manières de les collecter, de les stocker et de les réconcilier entre elles ;
- approfondir leurs connaissances sur les principaux outils du data marketing (DMP, CRM 360, marketing automation, écosystème adtech...);
- comprendre les bases de la data science appliquée au marketing (clustering, scoring, marketing prédictif...).

Dans un esprit résolument pratique et orienté « projet », de nombreux exemples viennent illustrer la théorie, démystifier les buzzwords qui envahissent les conférences et les discours de certains fournisseurs de solutions en plein « data washing » et amèneront, nous l'espérons, le lecteur de la confusion à la clarté !



Préambule

Concevoir son projet Data

D'après une étude CapGemini¹ menée en 2015, seuls 35 % des projets data d'envergure sont qualifiés de réussites ou de francs succès. Un constat qui n'est pas sans rappeler les débuts du CRM, dont les projets connaissaient près de 70 % d'échecs² !

La cause profonde des échecs dans la plupart des projets data n'est pas dans le manque de créativité des applications possibles, ni dans le manque de données, ni dans le manque d'outils technologiques, ni même dans le manque de savoir-faire pour collecter et analyser les données. **La cause réside le plus souvent dans la fragilité de la phase de conception:** attentes irréalistes, objectif business non précisément défini, «oubli» de la composante humaine, feuille de route trop superficielle... Il faut dire qu'avec la data, il est facile de se perdre dans l'immensité des possibles et de s'engluer dans des projets qui impliquent une multitude de personnes dans l'entreprise!

Sans apporter une réponse «miracle», ce préambule propose un canevas en cinq étapes pour éviter la majorité des erreurs de conception et aborder de façon structurée un projet data marketing:

- définir un objectif business spécifique;
- établir les critères de succès du projet;
- décomposer l'objectif en cas d'usages;
- étudier la faisabilité, technique et organisationnelle;
- déterminer une feuille de route.

Autrement dit, votre objectif doit être **Spécifique, Mesurable, Atteignable, Réalisable et Temporellement défini**, ce qui mis bout

1. «Cracking the data conundrum: how successful companies make Big Data operational».

2. Étude Butler Group, 2002. Ce taux était estimé à 47 % en 2009 par Forrester Research.

à bout constitue l'acronyme SMART, célèbre méthode créée en 1981 par George Duran et toujours massivement utilisée en management, en coaching, en gestion de projet... et dans ce préambule d'introduction à la conception d'un projet data marketing! Il s'agit d'un canevas «d'école» qui souffre comme la plupart des frameworks¹ de beaucoup d'exception, mais qui a le mérite de poser les bonnes questions très en amont.

UN VASTE CHAMP DES POSSIBLES

Quel est le cap mon capitaine? Avant de se lancer tête baissée dans la mise en œuvre d'un chantier de «data management», il est primordial de définir très précisément les objectifs de ses initiatives. À chaque but correspondront en effet des mécanismes variés de collecte, d'analyse et d'activation de la donnée.

Une des difficultés du data marketing est que la discipline interagit avec un très grand nombre d'activités historiques du marketing: prospection, acquisition, conversion, fidélisation... dans tous ces grands classiques du marketing conventionnel puis du marketing digital, de nouvelles perspectives s'ouvrent avec la data! Par exemple:

- **en matière de notoriété et d'acquisition**, il devient possible de cibler des profils «jumeaux statistiques» de ses meilleurs clients², d'exclure ses clients des campagnes d'acquisition, de piloter algorithmiquement les enchères Adwords, de faire levier sur la donnée de géolocalisation pour cibler selon les habitudes de déplacement...;
- **pour convertir** les prospects en leads³ puis les transformer en acheteurs, le marketing programmatique couplé à des algorithmes de personnalisation permet de délivrer par e-mail ou par bannières display des centaines de scénarios différents selon le profil du lead;
- **la fidélisation** n'est pas en reste et un des projets de data marketing les plus en vogue dans les milieux bancaires est le calcul d'un «score d'attrition», prédisant la propension du client à

1. Anglicisme pour désigner «canevas».

2. Procédé dit du «look alike audience». Nous reviendrons dessus plus en détail dans le chapitre 4.

3. Terme anglais fréquemment utilisé pour désigner un prospect intéressé par l'offre. Typiquement un internaute qui laisse ses coordonnées pour être rappelé est un «lead».

rompre son contrat. L'analyse sémantique des «verbatim»¹ clients au sein du service client ou sur les réseaux sociaux pour identifier des ambassadeurs de la marque est également une des applications phares de l'exploitation de la data à des fins marketing.

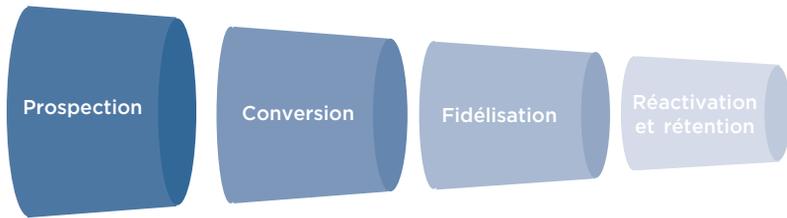


Figure 1 - La data peut être exploitée à tous les étages du tunnel marketing

Au-delà de l'optimisation de l'entonnoir marketing (en anglais, *funnel marketing*), le Graal du data marketing est de comprendre derrière l'achat le projet global du consommateur (exemple : ce consommateur achète une lampe, car il déménage ; ce consommateur achète ce thé vert, car il suit une cure d'amaigrissement) pour réinventer la manière de communiquer avec ce consommateur. Cette ambition est devenue accessible en croisant ses données avec des données externes.

Calculer plus finement le ROI de chaque levier marketing ou monétiser sa donnée sont également des projets data marketing à forte valeur ajoutée rencontrés régulièrement sur le marché.

Afin de rendre davantage concrètes les possibilités offertes par l'utilisation de la donnée à des fins marketing, nous considérerons dans la suite de ce préambule les exemples d'un assureur et d'un site éditorial d'actualités.

SPÉCIFIQUE : QUEL EST LE BUT PRÉCIS POURSUIVI PAR LE PROJET ?

Face au vaste choix des possibles, déterminer précisément des objectifs spécifiques est un des plus grands défis du professionnel

1. Mots et phrases employés par un consommateur lorsqu'il s'adresse directement à une entreprise (service client, réseaux sociaux, etc.).

du marketing. Les cinq règles d'or ci-dessous devraient aider à le relever.

Un objectif n'est pas un moyen

«Mieux comprendre mes clients», «segmenter ma base clients», «personnaliser mes e-mails», «faire du prédictif» sont des mauvais objectifs entendus en permanence dans la bouche de certains professionnels du marketing. Il s'agit en effet ici de moyens. C'est grâce aux analyses des bases de données, aux modélisations mathématiques ou à la personnalisation de la communication que seront atteints les véritables objectifs d'augmentation de la fréquence d'achat, d'augmentation du panier moyen, de baisse de l'attrition¹ de la base client, etc.

Une assurance souhaite réduire les ruptures de contrats. En analysant les départs à la concurrence, l'assurance constatera peut-être que dans 40 % des cas ces départs sont liés à un déménagement. Elle étudiera alors tous les cas de déménagement et arrivera à la conclusion que lors d'un déménagement, dans 75 % des cas le client rompt son contrat.

Dans la même veine, mettre en place une «Data Management Platform» (DMP)² ou tout autre outil data marketing n'est pas un objectif data. Le véritable objectif business qui justifie le projet DMP pourra être, par exemple, «réduire mes coûts d'acquisition digitaux qui m'asphyxient».

Exemple de l'assurance: une DMP sera un moyen de détecter les gens qui déménagent grâce à des partenariats data³ avec d'autres sites (fournisseurs d'énergie, seloger.com...).

Un objectif n'est pas une finalité

L'autre écueil fréquent dans la conception d'un projet data est la confusion entre objectif et finalité. «Augmenter le CA», «augmenter le taux de conversion du site», «acquérir des clients» ou «mieux fidéliser» sont les conséquences d'un projet data bien

1. Perte de clientèle, à la suite, par exemple, de la résiliation d'un abonnement. Le terme anglais «*churn*» est également très souvent utilisé.

2. Plateforme technologique facilitant la collecte de données anonymes depuis plusieurs sources (navigation web, campagnes média, marketing direct, partenaires...), la construction d'audiences et l'activation de ces audiences, principalement sur l'écosystème publicitaire.

3. Partenariats dits «second party».

mené, mais ne peuvent constituer un objectif spécifique associé à un projet data. Il y a en effet des dizaines de façons d'augmenter la conversion ou de fidéliser les clients, chacune pouvant faire l'objet d'un projet à part entière.

Exemple de l'assurance: l'objectif spécifique du projet est de détecter parmi la base de clients ceux qui ont un fort risque de déménager et ainsi de les appeler pour faire le point sur leur situation. Ce passage d'un marketing proactif à un marketing réactif aura pour finalité une baisse du taux d'attrition.

Exemple d'un site éditorial d'actualités: un site d'actualités souhaite augmenter ses revenus publicitaires. L'objectif du projet data est d'augmenter le nombre moyen de pages vues par internaute lors d'une visite.

Un objectif doit être précédé d'un constat

Il s'agit d'une «best practice» dont il faut user et abuser pour asseoir la crédibilité du projet. Très souvent des projets data marketing sont lancés alors que rien ne le justifie vraiment! Du coup personne ne sait exactement ce qu'il faut attendre du projet et il y a toutes les chances que le projet ne dépasse pas le stade du «Proof of Concept» (POC).

Par exemple dans le cas de l'assurance souhaitant lutter contre l'attrition:

Mauvaise pratique: «En tant qu'assureur, on n'a pas le choix: puisque l'assurance est obligatoire, grandir signifie forcément “piquer” des clients à ses concurrents. Le projet vise à lutter contre l'attrition grâce à la donnée.»

Bonne pratique: «Quand un client déménage, dans 30 % des cas il part à la concurrence, ce qui représente pour nous une perte de trois millions d'euros par an. Nous souhaitons grâce à la donnée:

- 1. déceler la probabilité de déménagement chez nos clients afin d'agir avant qu'il ne soit trop tard;*
- 2. optimiser notre offre “déménagement” pour aboutir à celle évitant au maximum le départ du client.»*