



# **MASTER TOURISME ET HÔTELLERIE**

Parcours « Management en Hôtellerie-Restaurant »

## **MÉMOIRE DE PREMIÈRE ANNÉE**

### **LE BIG DATA DANS L'HÔTELLERIE FRANÇAISE**

Présenté par :

**Tiffany ETCHEGORRY**





# **MASTER TOURISME ET HÔTELLERIE**

Parcours « Management en Hôtellerie-Restauration »

## **MÉMOIRE DE PREMIÈRE ANNÉE**

### **LE BIG DATA DANS L'HÔTELLERIE FRANÇAISE**

Présenté par :

**Tiffany ETCHEGORRY**

Année universitaire : **2015 – 2016**

Sous la direction de : **Olivier Laurent**

L'ISTHIA de 'Université Toulouse Jean-Jaurès n'entend donner aucune approbation, ni improbation dans les projets tuteurés et mémoires de recherche. Les opinions qui y sont développées doivent être considérées comme propres à l'auteur(e).

# Remerciements

Je tiens tout d'abord à remercier Monsieur Laurent, mon maître de mémoire qui a su se rendre disponible et être à l'écoute tout au long de ce projet. Ses conseils ont été indispensables à la réussite de ce travail.

Je souhaite aussi remercier tous les professionnels ayant consacré du temps à ce mémoire, que ce soit par le biais du questionnaire ou du guide d'entretien.

Mes remerciements vont également au corps enseignant de l'ISTHIA, pour l'ensemble de leurs cours qui m'ont très souvent aidée et les recommandations qui m'ont été faites pour traiter ce sujet.

Enfin, je souhaite tout particulièrement remercier mes amis qui m'ont soutenue et fourni une aide précieuse tout au long de la rédaction de ce mémoire.

# Sommaire

Remerciements .....	5
Introduction générale .....	8
Partie I. Qu'est-ce que le <i>Big Data</i> ? .....	10
Introduction .....	11
Chapitre 1. Les concepts fondamentaux.....	12
1. Définition.....	12
2. Historique .....	12
3. Dimensions du <i>Big Data</i> .....	15
4. <i>Big Data</i> et concepts associés.....	17
Chapitre 2. La collecte de données: comment et pourquoi? .....	19
1. Les technologies du <i>Big Data</i> .....	19
2. La nature des données.....	25
3. Les sources de production de données.....	27
4. Les grands acteurs du <i>Big Data</i> .....	29
Chapitre 3. Approches disciplinaires .....	31
1. Approche économique.....	31
2. Approche éthique et juridique .....	32
Conclusion .....	34
Partie II. <i>Big data</i> et marketing : quelles opportunités pour les entreprises ? .....	35
Introduction .....	36
Chapitre 1. Le <i>Big data</i> dans les entreprises .....	37
1. Vers un changement des processus décisionnels .....	37
2. Les secteurs impactés .....	38
3. La mise en place d'une démarche <i>Big data</i> .....	39
Chapitre 2. Les effets du <i>Big Data</i> dans la stratégie marketing .....	41
1. La stratégie marketing .....	41
2. Les apports dans les options stratégiques fondamentales .....	44
3. Apports dans le mix marketing.....	47
Chapitre 3. Enjeux et limites.....	51
1. Les enjeux .....	51
2. Les limites .....	55
Conclusion .....	58

Partie III. Méthodologie et terrain d'application .....	59
Introduction .....	60
Chapitre 1. Méthodologie.....	61
1. Déroulement de la recherche .....	61
2. Rappel des objectifs .....	61
3. Les outils utilisés .....	62
4. Les limites .....	64
Chapitre 2. Le parc hôtelier français.....	65
1. Panorama du parc hôtelier français.....	65
2. L'hôtellerie indépendante .....	67
3. L'hôtellerie de chaîne .....	67
Chapitre 3. Analyse des actions des acteurs .....	71
1. Profil des répondants .....	71
2. Constat face à l'utilisation des données .....	71
3. Le stockage des données.....	77
Chapitre 4. Perspectives futures.....	79
1. Possibilités de stage en Master 2.....	79
2. Travail de recherche envisagé .....	79
Conclusion .....	80
Conclusion générale.....	81
Bibliographie.....	83
Annexes .....	85
Table des annexes .....	86
Glossaire.....	106
Table des figures .....	108
Table des tableaux.....	108
Table des matières.....	109

# Introduction générale

L'informatique (contraction d'« information » et d'« automatique »), est au cœur de nos vies depuis plus de quinze ans. Révolution à bien des niveaux, le numérique a entraîné une véritable mutation des comportements, et, avec eux, de nombreuses interrogations. Devenue une réelle nécessité pour la plupart des individus, la technologie ne cesse de progresser pour proposer des usages toujours plus innovants visant à faciliter nos vies.

Phénomène de mode ou perspective d'avenir, il est certain que le *Big data* est désormais un terme incontournable relayé en permanence par les media et le monde professionnel. L'objectif de ce mémoire est de pouvoir comprendre les conséquences d'une nouvelle révolution qui, bien au-delà de son aspect technologique, a également entraîné de profonds changements sociologiques.

L'utilisation permanente de dispositifs reliés à Internet a entraîné une croissance exponentielle du nombre de données produites, pour atteindre plusieurs milliards d'octets par an. Cette masse colossale d'informations constitue autant de traces qui, mises bout à bout, parviennent à décrire les comportements, les habitudes, d'un individu comme d'une population entière. Les entreprises l'ont bien compris et sont de plus en plus nombreuses à intégrer cette composante dans leur stratégie globale et plus particulièrement dans leur plan marketing.

Le secteur de l'hôtellerie a souvent été montré du doigt pour son manque de réactivité face au changement. La numérisation des supports est notamment un problème récurrent, car avec des habitudes bien ancrés, certains professionnels ont du mal à s'y adapter. Nous pouvons donc nous poser les questions suivantes : comment les hôteliers français intègrent-ils les nouvelles tendances en matière d'informatique ? Comment adaptent-ils leur stratégie face à l'avalanche de données actuelle ?

Le *Big data* est aujourd'hui un outil majeur en terme de marketing hôtelier et pourtant rares sont les professionnels à vraiment avoir adopté cette composante. Est-ce une question de financement ? De manque d'informations ? Ou simplement d'ignorance face à cette tendance croissante ? Et surtout qui utilise réellement le *Big Data* ? Ce sont autant de questions auxquelles ce mémoire tend à trouver des réponses.

La première partie de ce travail portera sur la conceptualisation de ce vaste sujet afin de pouvoir mieux l'appréhender par la suite. Nous y verrons tout d'abord l'origine de son apparition, les différentes dimensions qui le représentent puis l'ensemble des acteurs, des moyens techniques

et technologiques qui ont permis l'émergence de cette évolution majeure. Enfin, nous l'étudierons sous un angle économique et juridique, ces deux disciplines ayant une importance capitale lors de l'évocation des données.

Nous aborderons en second lieu l'impact du *Big data* sur les entreprises et les conséquences engendrées par son intégration, mais également les opportunités qu'il représente. Ses apports dans le plan marketing hôtelier seront détaillés et s'appuieront sur les informations collectées dans la revue de littérature ainsi que sur l'analyse de l'entretien exploratoire mené, permettant d'apporter quelques éléments de réponses à notre questionnement.

La dernière partie présentera tout d'abord la méthodologie adoptée tout au long de l'année, nous ayant permis d'éclaircir de nouveaux points quant aux hypothèses préalablement construites. Elle fera ensuite le point sur l'étude menée auprès des hôteliers dont l'objectif était d'établir un constat des pratiques de chacun en matière d'exploitation de données.

## **Partie I. Qu'est-ce que le *Big data* ?**

## Introduction

Face à un terme auquel nous sommes de plus en plus exposés, il est indispensable de comprendre l'ensemble des éléments qui compose le *Big data*. Effet de mode ou véritable bouleversement technologique, il est certain que cette expression anglophone ne laisse personne indifférent.

Le premier chapitre de cette partie a pour but d'éclaircir les notions fondamentales liées à cette masse historique de données : ses origines, son évolution, ses caractéristiques qui permettront par la suite d'aborder le reste du sujet de manière plus concrète.

Nous verrons dans un second temps quelles sont les différentes natures de données disponibles et exploitables (ces deux notions ne sont pas forcément associables), ainsi que les acteurs et les technologies qui ont participé à leur création et à leur diffusion au cours des deux dernières décennies.

Enfin, nous aborderons la question du *Big data* sous deux autres angles : économique et juridique. Car l'ensemble des concepts développés s'inscrit, en plus d'une dimension technique indéniable, dans une optique bien plus large qui nous touche de bien des manières, et soulève des interrogations en matière de répercussions sur l'économie mondiale et sur notre vie privée.

## Chapitre 1. Les concepts fondamentaux

De plus en plus fréquemment utilisé par les médias et les entreprises, le terme de *Big data* est au cœur de nombreuses réflexions actuelles. Derrière une définition souvent floue et mal maîtrisée, se cache un concept de grande envergure, en passe de bouleverser notre rapport aux nouvelles technologies et à l'utilisation qui peut en être faite. Car avec la révolution numérique de cette dernière décennie, le *Big Data* concerne chaque personne à travers ses pratiques quotidiennes, qu'elles soient d'ordre public ou privé.

### 1. Définition

Souvent considéré comme un déluge d'informations produites à chaque seconde dans le monde, la définition du *Big Data* (ou métadonnée) est encore vague pour les acteurs de la société actuelle.

La définition donnée par le Mercator (2014, p.480) est la suivante : « Flux de données considérable collectés par le suivi automatique des comportements en ligne : visites de sites, achats de produits, clics sur les bannières et les liens, etc. ». Concrètement, le *Big Data* désigne non seulement le volume considérable de données produites à chaque seconde, mais également sa collecte, son traitement et son stockage. De plus, ces données ne sont pas uniquement issues de la navigation sur des sites Internet par les individus, mais également d'autres canaux et supports utilisés comme les applications sur *smartphones*, la géolocalisation, les transactions bancaires, etc.

### 2. Historique

#### 2.1 L'origine du *Big Data*

La collecte et l'exploitation des données sont loin d'être des faits nouveaux. Les populations, au cours des générations, ont perpétuellement cherché à échanger des informations de la manière la plus efficace possible. L'apparition de l'informatique et les avancées technologiques qui en ont découlé ont permis de passer d'un stock de données physique et difficilement exploitable à un flux fluide et dynamique, en perpétuel mouvement.

Si l'apparition du terme de *Big Data* (ou métadonnée) a été introduit pour la première fois au congrès de l'Econometric Society en 2000<sup>1</sup>, il a toutefois fallu une décennie supplémentaire pour que cette expression rentre dans le langage courant. Le graphique ci-dessous démontre l'engouement récent des internautes pour ce phénomène à travers l'évolution de la recherche de sa signification.

---

<sup>1</sup> Cité par DELORT Pierre, *Le Big data*, 2014, p.5

Figure A - Évolution des recherches associées au terme *Big Data*



Source : Google Trend

Nous pouvons remarquer que l'intérêt vis-à-vis de ce concept débute fin 2011, période durant laquelle de plus en plus d'entreprises ont commencé à se pencher sur l'utilité des données dans la prise de décision, ainsi que sur les conséquences que leur exploitation pouvait potentiellement engendrer.

## 2.2 Évolution

### a) Des usages

Bien qu'ayant longtemps été réservée aux armées et aux organisations gouvernementales (administration, armée, renseignements) puis aux très grandes entreprises, l'informatique s'est peu à peu immiscée dans les foyers du monde entier. Aux débuts de la démocratisation de l'usage d'Internet, cet outil était généralement destiné à des usages simples (recherches basiques, envois d'e-mails, etc.) mais permettait un gain de temps considérable grâce à l'affranchissement de contraintes spatio-temporelles jusqu'alors incontournables (informations disponibles à domicile, délais de transmission de l'information infimes par rapport à un envoi postal, etc.).

Au début du 21<sup>ème</sup> siècle, certaines entreprises ont eu l'idée d'exploiter ce canal à des fins commerciales, proposant leur offre sur la Toile comme Cdicount ou Priceminister. Après des débuts timides, le e-commerce a littéralement explosé, entraînant un bouleversement des modèles économiques traditionnels. En parallèle de cette mutation, nous avons également assisté à un changement profond des interactions sociales avec le web 2.0 et sa capacité à partager des informations sur les réseaux communautaires (MSN, Facebook, Myspace, puis des sites comme Tripadvisor par la suite). La création et la diffusion de contenus sont alors devenues accessibles à tous, et non plus uniquement à une élite ou à quelques acteurs majeurs du web.

Face à cette évolution capitale, le volume de données produites a explosé. Jusqu'alors stockées sans but particulier, le potentiel économique et décisionnel des *data* a peu à peu été mis en exergue pour finalement devenir un bien immatériel de grande valeur.

Le web 3.0 repose d'ailleurs sur ces deux principes. Apparu aux alentours de 2010, il découle de la participation collective des individus dans la création et l'alimentation de la Toile. Le e-commerce est une pratique désormais courante, en passe de détrôner les échanges commerciaux classiques qui semblaient jusqu'alors immuables. C'est l'ère du *crowdsourcing*, des échanges collaboratifs et de la dématérialisation de l'économie, et qui tend à s'accroître encore davantage dans un futur proche. Les données ont fait émerger un marché dont la progression augmente de jour en jour, avec l'apparition dans le monde entier de nouveaux acteurs privés et publics spécialisés dans ce domaine.

Le web 4.0 fait d'ores et déjà son apparition, et avec lui, la banalisation de la consommation de services marchands et non marchands *via* Internet, quelle que soit leur nature. Ce phénomène prend une ampleur telle que les entreprises de secteurs comme le transport ou le tourisme sont directement impactées. Le succès de jeunes *startups* telles que Airbnb, Uber, ou encore Blablacar, démontre un engouement des individus pour ces services produits par des particuliers à des prix ultra compétitifs, qui pousse ainsi les industries à revoir leurs offres et leurs modèles économiques.

## b) Des supports

Les 5 dernières années ont vu apparaître une forte augmentation des supports destinés à se connecter à Internet en raison de l'amélioration des espaces de stockage et de la fluidité de navigation offerte aux usagers.

Au départ, la possession d'un ordinateur n'était réservée qu'à un cercle limité en raison de son coût d'une part, mais également de la place qu'occupaient les dispositifs. Grâce à la miniaturisation résultant des prouesses techniques et technologiques de ces deux dernières décennies, l'ordinateur est aujourd'hui un bien accessible et indispensable à tous. Face aux mutations sociales engendrées par cette révolution, l'accès à l'information doit aujourd'hui être possible partout, tout le temps.

Les supports s'enchaînent et ne cessent d'évoluer depuis, passant il y a quinze ans d'un téléphone portable permettant exclusivement d'émettre et de recevoir des appels, à des *smartphones* capables de nous renseigner sur notre position géographique et d'effectuer n'importe

quelle recherche quel que soit l'endroit dans lequel nous nous trouvons. De plus, les tablettes et les objets connectés envahissent les foyers du monde entier. Les réfrigérateurs, les montres, les voitures, les télévisions sont autant de supports conçus pour rendre nos vies plus faciles et aider à visualiser et à maîtriser nos actions (calories brûlées, tension, localisation, consommation énergétique de la maison, etc.).

L'objectif ici n'est plus uniquement d'utiliser Internet dans le but de rechercher de l'information, mais également d'en produire afin de pouvoir être guidés dans nos choix et nos décisions.

### 3. Dimensions du *Big Data*

Il est difficile de visualiser l'ampleur de cette nouvelle révolution numérique. Pour pouvoir s'en faire une idée un peu plus précise, Meta Group présenta en 2001<sup>2</sup> un rapport faisant état de 3 dimensions à travers lesquelles l'amoncellement et le traitement de données diverses pouvaient être représentés : le volume, la vitesse et la variété. Le terme de *Big Data* n'était alors pas encore apparu et peu de gens étaient conscients que chacune de leur trace laissée sur Internet était enregistrée et stockée. Néanmoins, la définition du *Big Data* vient s'enrichir d'autres termes et interrogations qui en découlent.

#### 3.1 V pour Volume

La numérisation des données a augmenté d'une manière significative : en 2000, seulement 20% des informations collectées étaient numérisées, les 80 autres pourcents étant archivés de manière analogique (documents papiers, photographies, cassettes audio ou VHS). Aujourd'hui, on estime que 98% de nos informations sont stockées de manière non tangible à travers des disques durs, clé USB, ou encore le *Cloud* en raison de la diminution considérable des coûts liés au stockage (TEBOUL et BOUCHER, 2014, p. 127).

Depuis les débuts de l'humanité jusqu'en 2003 on estime que la quantité de données produites sur cette période équivaut à 5 milliards de d'octets. En 2010, cette même quantité était produite en deux jours<sup>3</sup>. À titre d'exemple, la Bibliothèque du Congrès américain (plus grande bibliothèque mondiale), représente à elle seule 200 téraoctets de données compilées dans plus de 80 000 ouvrages, alors que Google en traite près de trois-mille fois plus quotidiennement<sup>4</sup>, si

---

<sup>2</sup> Meta Group. *3D Data Management: Controlling Data Volume, Velocity, and Variety*. [en ligne]. Disponible sur <https://lc.cx/4m6X>. (Consulté le 22-11-2015)

<sup>3</sup> LUPIERI Stéphan. *Big data : Devant nous le déluge*, *LesEchos.fr*, [en ligne]. Disponible sur <https://lc.cx/4m6B>. (Consulté le 10-12-2015).

<sup>4</sup> MAYER-SCHONBERGER Viktor, CUKIER Kenneth. *Big Data: A Revolution that will transform how we live, work and think*. 2013, p.17

bien que l'on parle maintenant de mesures en pétaoctets ou zéttaoctets<sup>5</sup>. Et cette tendance est loin de décroître car selon les estimations, le volume total de données produites est multiplié par deux tous les deux ans.

Cette dimension est celle qui présente la plus grande difficulté et le plus grand enjeu, pour les entreprises notamment, car elle pose la question du choix et du tri des données les plus pertinentes à récolter en vue de les exploiter. Face à la quantité colossale d'informations disponible, il est impératif pour les acteurs économiques de les sélectionner de manière judicieuse pour les aider dans leur prise de décision.

### 3.2 V pour Vitesse (ou vélocité)

L'avantage majeur de la dématérialisation des informations réside dans le fait qu'elles peuvent transiter de manière *quasi* instantanée en s'affranchissant de toute contrainte d'espace et de temps. Lors de l'apparition d'Internet et de sa diffusion au grand public à travers des modems, certes lents comparés à ceux disponibles de nos jours, mais d'une rapidité incroyable pour l'époque, la récupération de nombreuses informations a alors été possible sans délais et sans supports physiques autres qu'un ordinateur. L'apparition de l'ADSL, du haut-débit puis de la fibre n'ont cessé de faire diminuer le temps nécessaire au transit des données, améliorant et accélérant grandement la connaissance des individus. Dans une société évoluant à grande vitesse, cet aspect technologique constitue un atout majeur pour les structures privées comme publiques, désormais capables de prendre des décisions en temps réel.

### 3.3 V pour Variété

Des millions d'octets sont aujourd'hui traités à très grande vitesse permettant d'accéder à la connaissance bien plus rapidement que dans le passé. La numérisation des informations engendrée par l'essor des nouvelles technologies (appareils photo numériques, GPS, *smartphones*, bornes...) permet aujourd'hui de stocker de nombreux éléments de manière intangible quelle que soit la nature de celui-ci (texte, fichier audio, vidéo, emails...). Cette variété de supports représente un véritable souci pour les entreprises qui doivent analyser les données qu'ils produisent afin d'en retirer des éléments pertinents et parvenir à les croiser pour leur attribuer un sens.

---

<sup>5</sup>

### 3.4 Les autres V

La véracité et la valeur sont les deux autres dimensions retrouvées de plus en plus souvent lorsqu'il s'agit du *Big Data*. La véracité tient au fait que face à l'afflux massif de données quotidien, ces dernières deviennent rapidement obsolètes et nécessitent donc un traitement rapide, voire instantané, afin d'être utiles et exploitables. Ce contexte est également à prendre en compte car il peut considérablement faire varier le sens d'une information ou d'une requête (SCOBLE, 2014, p. 90).

La valeur, quant à elle, peut être variable en fonction des informations disponibles. Un profil basique comprenant l'âge, le genre et le lieu d'habitat se monnaie, sur le marché de la donnée, à environ 0.0007 €. Plus le profil est renseigné, plus sa valeur augmente (composition du foyer, catégorie socio-professionnelle, loisir, niveau de santé, etc.) et présente d'intérêt pour les entreprises, notamment les services marketing qui pourront les utiliser dans la segmentation de leur clientèle effective et potentielle, la construction de leur offre, de leur communication, mais également pour leur revente à d'autres organismes, devenant une activité annexe potentiellement fructueuse pour les acteurs économiques de tout bord.

## 4. *Big Data* et concepts associés

### 4.1 L'*open data*

La collecte et la circulation de données anonymes a permis de créer un réseau de partage d'informations communément appelé *open data*. Les données collectées peuvent venir à la fois de l'État, comme des entreprises de service public (musées, transports, etc.), et ont pour but premier de les rendre accessibles à tout un chacun à des fins informationnelles, commerciales, décisionnelles. Par exemple, la ville de Paris a mis à disposition des jeux de données sur la pollution de certains quartiers, les marchés publics ayant été passés ou encore les établissements parisiens dans lesquels le café était à moins d'un euro. À travers son site *data.gouv*, le gouvernement met à la disposition de ses citoyens des jeux de données brutes, mais structurées, permettant de connaître une grande quantité d'informations relatives au territoire (cartographie des établissements Tourisme et Handicap, base de codes postaux, cartographie des musées, santé, etc.). Ces fichiers sont composés de données anonymes et nécessitent un traitement pour leur conférer une réelle utilité dans la prise de décision.

---

<sup>6</sup>CALIXTE Laurent. *Internet : calculez vous-même la valeur de vos données personnelles*, juillet 2013, Challenges.fr [en ligne]. Disponible sur <http://urlz.fr/3grn>. (Consulté le 12-01-2016).

## 4.2 Le *crowdsourcing*

Le *crowdsourcing* ou production participative fait appel à l'intelligence et à la créativité collective dans une optique d'innovation, de résolution d'un problème, d'accomplissement d'une tâche ou d'acquisition de données par le biais du web et de techniques de collaboration sociale, qui seraient autrement inaccessibles<sup>7</sup>.

On peut citer le cas des *start-up* comme BlablaCar ou Uber pour le transport, Airbnb pour l'hébergement, dont le fonctionnement repose sur ces principes. Le marketing moderne fait d'ailleurs de plus en plus appel à cette forme de participation des internautes dans la conception de leurs offres, permettant à la fois d'intégrer le consommateur dans le processus de création, de connaître ses attentes et ses freins en créant une véritable relation avec celui-ci.

## 4.3 Le *cloud computing*

Face au besoin croissant en termes de stockage, les entreprises et les particuliers sont de plus en plus nombreux à délocaliser l'enregistrement de leurs données sur des supports externes. Ce système se traduit par la « location » d'espaces de stockages auprès de structures spécialisées matérialisées par des *data center* (ou entrepôts de données) qui offrent d'une part des supports dont la mémoire est bien supérieure à ceux disponibles sur le marché public, mais assurent aussi une protection et une sécurité des données bien plus élevés que les dispositifs personnels.

Cette tendance croissante est appelée le *cloud computing* (ou informatique en nuage) et séduit de plus en plus d'entreprises qui confient leurs données à ces structures spécialisées dont nombre a considérablement augmenté en réponse à une demande en augmentation constante.



Nous avons pu voir au cours de ce premier chapitre les notions théoriques de base entourant le terme de *Big data* : ses origines, ses caractéristiques principales (Volume, Vitesse, Variété) ainsi que les différents concepts qui lui sont associés et qui sont désormais entrés dans le langage courant. Il convient à présent de s'attacher au cœur du sujet, les données : leurs différentes formes, les moyens de les collecter et de les traiter en vue de les exploiter à des fins décisionnelles. L'ensemble de ces concepts sera développé dans le chapitre ci-après.

---

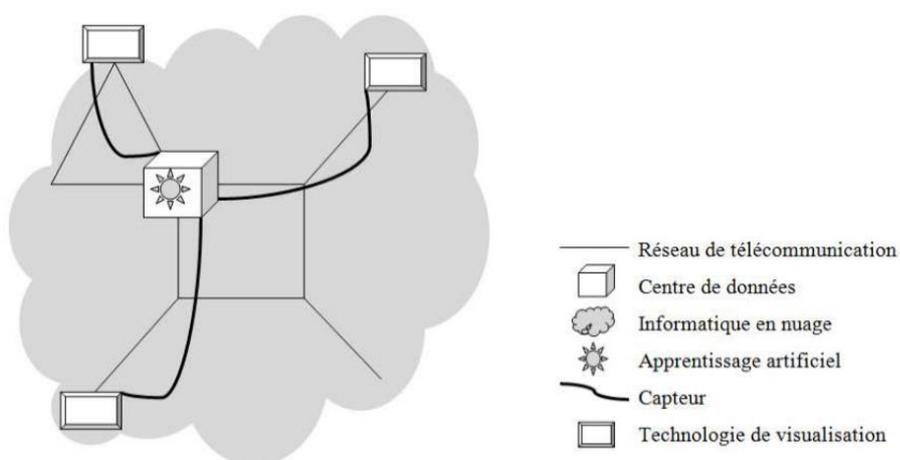
<sup>7</sup> Gartner. IT Glossary: crowdsourcing [en ligne]. Disponible <http://urlz.fr/3isT>. (Consulté le 03-02-2016).

## Chapitre 2. La collecte de données: comment et pourquoi?

Les progrès réalisés en matière de Technologies de l'Information et de la Communication (TIC) ont permis de diffuser l'usage d'Internet aux citoyens du monde entier mais également aux entreprises pour lesquelles ces technologies présentent un enjeu majeur. Pour rendre ces actions efficaces, différents acteurs sont indispensables : les utilisateurs, les organisations ainsi que toutes les entités intervenant dans la collecte, le stockage et le traitement de ces données.

Le schéma ci-dessous permet de mieux comprendre les procédés permettant de les récolter :

Figure B - Système de de production et de captation des *Big Data*



Source : Vayre, 2014

Ce schéma expose, de manière simplifiée, comment nos traces parviennent à être consignées en vue d'une exploitation potentielle. Ce chapitre sera consacré aux diverses technologies et techniques mises en œuvre pour récolter, analyser et stocker les données de différentes natures ainsi que les différents protagonistes intervenant dans ce processus complexe.

### 1. Les technologies du *Big Data*

#### 1.1 Le processus d'exploitation des données

Afin d'être utilisables à des fins décisionnelles, les données doivent subir un traitement en amont donnant lieu à des indications sur un ou plusieurs points :



Ce cheminement correspond au traitement classique de l'information. La seule différence, bien qu'elle ait une importance capitale, porte sur le raffinement qui correspond au tri des données, en

vue de ne sélectionner que celles qui revêtent une importance particulière. La valorisation, quant à elle, désigne le fait de modéliser et mettre en forme les informations.

### 1.2 Les *data center* et les *datawarehouse*

Les données ont beau être intangibles, leur existence revêt une dimension matérielle incontournable. Car si les informations transitent rapidement, leur acheminement et leur stockage ne peuvent être dissociés de supports physiques. Les *data centers* sont de vastes structures dans lesquelles fonctionnent de manière continue des disques durs de capacité incomparablement supérieure à ceux dédiés au grand public. Ce sont dans ces établissements que les données sont enregistrées, triées, les rendant ainsi disponibles à leurs détenteurs qui « louent » des espaces de mémoire. On estime à 3625 le nombre de ces structures disséminées dans le monde entier<sup>8</sup>, dont 139 en France, en troisième place du classement européen des pays disposants du plus d'espaces de stockage numérique.

Les *datawarehouse* sont quant à eux des architectures de stockage conservant les données extraites de systèmes de transaction, de données opérationnelles et de sources externes, fonctionnant généralement sous forme de clusters d'entreprises<sup>9</sup>. Ces systèmes permettent également un traitement de leur contenu grâce à des outils de *business intelligence*. Face à l'augmentation exponentielle de données, beaucoup d'entreprises ont dû externaliser la sauvegarde de leurs fichiers dans ces espaces de stockage plus puissants et plus sécurisés, donnant ainsi naissance au *Cloud*.

### 1.3 Les bases de données

« Une base de données (BDD) est un conteneur structuré de données, intégré dans un ensemble de logiciels permettant de décrire, de créer, de modifier, de supprimer et de consulter les données. La base de données et ses logiciels associés constituent le système de gestion de base de données »<sup>10</sup>. Une grande partie d'entre elles fonctionne grâce à un langage informatique normalisé appelé SQL (*Structured Query Language* ou Langage de Requête Structurée) donnant la possibilité de combiner des données opérationnelles (ventes, clients, fournisseurs, etc.) avec des données transactionnelles (paiements, stocks, etc.). La massification et la diversification perpétuelles de données liées au concept de *Big data* soulignent néanmoins les limites de cet outil. En effet, elles rendent l'analyse instantanée difficile et longue en raison du volume

---

<sup>8</sup> CAULIER Sophie. *Les data center, clés de voûte du réseau*, Le Monde Économie, mai 2015, [en ligne]. Disponible sur <https://lc.cx/4m62>. (Consulté le 18-01-2016).

<sup>9</sup> Gartner. *IT Glossary : Data Warehouse* [en ligne]. Disponible sur <http://urlz.fr/3itf>. (Consulté le 19-01-2016).

<sup>10</sup> BRASSEUR Christophe. *Enjeux et usages du Big Data : technologies, méthodes et mise en œuvre*. Paris : Éditions Hermes Science Publications, 2013, p.101

considérable d'informations à traiter. De plus, l'augmentation de la quantité de données non structurées, constituant 80 % du web, a mené à repenser le stockage afin de le rendre plus intelligent. Le langage NoSQL (Not Only SQL) est ainsi apparu en 1998, permettant de prendre en charge de manière plus aisée les informations présentant des structures différentes. Aucune de ces méthodes ne revêt cependant de caractère magique faisant apparaître des solutions et des recommandations toutes prêtes. L'enjeu majeur des développeurs est désormais d'en « faire le meilleur en fonction des situations et des besoins »<sup>11</sup> et impose une réflexion quant à leur modélisation et la représentation visuelle qui en découle.

#### 1.4 Les logiciels et algorithmes

De nombreux scientifiques ont œuvré depuis le siècle dernier afin de rendre les machines capables d'interagir avec l'être humain de la manière la plus naturelle possible. L'amélioration permanente des logiciels et des algorithmes ont permis de créer des machines capables de « penser » par elles-mêmes et de réagir différemment en fonction du contexte dans lequel elles évoluent. L'intelligence artificielle fascine autant qu'elle effraie et, au vu des progrès techniques et technologiques actuels, est en passe de surpasser ses créateurs. Son utilité indéniable passe par les différentes dimensions développées ci-dessous.

##### a) Le *data mining*

L'exemple de *data mining* le plus utilisé pour illustrer ce concept est celui de la chaîne de supermarchés Wal-Mart qui, en analysant ses ventes (contenu des paniers, fréquence et moment d'achat), s'est rendue compte qu'il existait un lien entre l'achat de bière et de couches le vendredi soir. Cette corrélation a été expliquée par le fait que les jeunes pères de familles, avant de rentrer chez eux pour regarder les matchs de football du week-end à la télévision, faisaient un détour par le supermarché le plus proche pour y acheter des bières. Ayant des enfants en bas âge, les jeunes mères leur demandaient d'acheter des couches par la même occasion. Afin d'augmenter ses ventes, l'enseigne a alors décidé de rapprocher ces deux produits dans le rayonnage, augmentant de manière significative son chiffre d'affaires pour ces deux produits (GORUNESCU, 2011, p. 23-24).

Au niveau marketing, cette découverte a permis de mettre en avant non seulement le couplage de ces deux produits, mais également le contexte dans lequel était effectué l'achat, permettant

---

<sup>11</sup> KHAN Azharuddin. *Difference between SQL and NoSQL: A comparison*. [en ligne]. Disponible sur <https://lc.cx/4m66>. (Consulté le 08-03-2013)

ainsi de repenser son rayonnage (ou *merchandising*) de façon à optimiser le passage de chaque client dans le magasin (stratégie extensive). Cette expérience a démontré l'importance et l'impact que peut avoir l'étude du comportement des consommateurs. Grâce à des algorithmes de plus en plus pointus, il est désormais possible pour les entreprises y ayant recours de déterminer de manière fine le profil des consommateurs, offrant la possibilité d'adapter leur communication et de proposer des offres en rapport avec leurs attentes, leur contexte de consommation et leurs aspirations.

Pour Stéphane TUFFERY<sup>12</sup>, professeur de *data mining* à l'ENSAI (École Nationale de la Statistique et de l'Analyse de l'Information), « *le data mining est l'art d'extraire des informations, voire des connaissances à partir de données [...] et permet de limiter la subjectivité humaine dans les processus de décision, et aussi, grâce à la puissance grandissante des outils informatiques, de traiter de plus en plus rapidement de grands nombres de dossiers* ».

Cet outil est de plus en plus répandu dans le cadre du marketing stratégique et opérationnel car il permet d'agir sur de nombreuses dimensions en dehors de la découverte d'associations inattendues entre deux produits<sup>13</sup> :

- Création de packages et de promotions ;
- Conception de nouveaux produits ;
- Optimisation de la tarification ;
- Politique de la fidélisation de clientèle ;
- Segmentation de la clientèle ;
- Adaptation de la communication à chaque segment de clientèle.

## b) L'analyse prédictive

L'analyse prédictive découle directement du *data mining*, qui, à travers l'analyse de données, permet de déduire de nouvelles informations de nature qualitative (*scoring*) ou quantitative (prédiction). Il s'agit de d'anticiper un comportement probable par l'agrégation de variables explicatives présentes dans le système de stockage (lieu d'achat, fréquence, quantité...). Pour Larry Maisel, fondateur du cabinet de conseil DecisionVu<sup>14</sup>, le succès d'un tel outil repose sur

---

<sup>12</sup> TUFFERY Stéphane. *Data mining et statistique décisionnelle. L'intelligence des données*. Paris, Éditions tecnip, 2012, p. 1

<sup>13</sup> *Ibid.* p. 10

<sup>14</sup> ALPERN Peter. *Q&A: DecisionVu's Larry Maisel on predictive Business Analytics*. *Business Finance* [en ligne]. Disponible sur <http://urlz.fr/3gcR>. (Consulté le 07-02-2016)

les compétences des décideurs, les technologies sur lesquelles il s'appuie et le processus d'analyse continu des performances antérieures de l'entreprise. La maîtrise de l'ensemble de ces facteurs permet à la fois d'avoir une vision claire des scénari pouvant survenir et prendre les décisions les plus efficaces et efficientes possibles, voire de les automatiser.

Le recours à ces programmes a tendance à se généraliser dans les grandes firmes et découle de la massification des données, de l'amélioration des dispositifs analytiques et surtout du succès de son application. Il nécessite néanmoins des connaissances pointues rarement détenues par les dirigeants et qui imposent de faire appel à des professionnels ou des entreprises spécialisées dans l'analyse et le traitement de données.

### c) Le *machine learning*

De manière générale, le *machine learning* correspond à l'apprentissage automatisé des ordinateurs et de leurs programmes, sans intervention ou assistance humaine, en vue de mener une action face à un évènement<sup>15</sup>. La question de l'intelligence artificielle occupe les esprits de nombreux scientifiques depuis longtemps maintenant. Le premier à avoir mis en évidence les capacités de cognitives des machines est Arthur Samuel qui mit au point dans les années 50 un programme permettant de battre un joueur humain au jeu de dames. Cette prouesse a été rendue possible par l'analyse de toutes les combinaisons possibles de jeu par le programme, permettant ainsi de déterminer quelle combinaison était la plus judicieuse pour le mener à la victoire.

Les principes de cette technique ont été depuis encore améliorés pour constituer un des éléments de base du e-commerce, à l'image d'Amazon et de son système de recommandation dont les algorithmes restent un bien gardé secret. Il en va de même avec le moteur de recherche Google qui propose à ses utilisateurs les résultats qu'il juge les plus pertinents en fonction des requêtes précédemment effectuées.

### d) La sémantique

D'autres individus comme le mathématicien Alan Turing se sont intéressés de près à cette innovation technologique et ont mis en évidence la capacité d'apprentissage des machines. Le test de Turing reste un exemple très parlant des prouesses dont elles sont capables. Lors d'une expérience menée elle aussi dans les années 50, le scientifique a mis en évidence l'aptitude d'un ordinateur à imiter le langage naturel d'un être humain, au point que son interlocuteur ne puisse

---

<sup>15</sup> SCHAPIRE Rob. *Theoretical Machine Learning*, 2008, 6 p. [en ligne]. Disponible sur <http://urlz.fr/3gcT>. (Consulté le 28-12-2015)

pas détecter qu'il converse avec une machine. La sémantique reste un pôle de recherche et d'innovation primordial pour les acteurs du *Big Data*, notamment en ce qui concerne l'analyse des contenus échangés sur la Toile qui permettent d'avoir une meilleure approche du client en ayant connaissance du champ lexical adopté, de ses sujets favoris, de ses impressions vis-à-vis d'une prestation ou d'un produit consommé. Bien que nombreux progrès aient été faits en la matière, le développement de cette technique relève du challenge de par la complexité et la diversité des structures (langue, orthographe parfois aléatoire, nouvelles expressions, etc.) ainsi que des supports employés (réseaux sociaux, forums, nature des requêtes effectuées, etc.).

### 1.5 Les principales solutions *Big data*

Des entreprises comme Google ont dû, pour pouvoir perpétuer leur activité, trouver des solutions pour répondre à leurs besoins vis-à-vis des multitudes de données qu'elles récoltaient. Elles ont alors conçu leurs propres logiciels et leurs propres modèles de programmation pour pouvoir y parvenir et les mettent aujourd'hui à disposition du grand public<sup>16</sup>. Les deux principaux sont MapReduce et Hadoop, mis respectivement au point par Google et la fondation Apache, deux entités étroitement liées dans la mesure où les technologies de ce dernier ont été développées à partir de celles du géant de l'Internet. De nombreuses grandes firmes comme Amazon, IBM, Spotify, LinkedIn ou encore le New York Times ont déjà fait le choix de ces solutions dans leur fonctionnement quotidien car elles permettent de gérer des masses considérables de données et présentent de nombreux avantages en étant :

- Souples, car elles permettent de traiter une grande variété de données ;
- Rentables, car elles permettent d'optimiser les coûts grâce à une meilleure mobilisation des ressources ;
- Résilientes, car elles sont à même de continuer le traitement même si une des composantes du système tombe en panne ;
- Évolutives, car elles s'adaptent aux besoins de manière transparente.

Le principe de fonctionnement repose sur plusieurs étapes dont les deux principales sont le *mapping* et le *reducing*. La première consiste à accomplir une tâche spécifique sur chaque élément alors que l'autre combine les éléments en fonction d'un algorithme particulier pour pouvoir fournir les résultats.

---

<sup>16</sup> CORINUS Michaël, DEREY Thomas, MARGUERIE Jérémie. *Rapport d'étude sur le Big data*, 2012, 54 p. [en ligne]. Disponible sur <http://urlz.fr/3gr0>. (Consulté le 08-03-2016)

## 2. La nature des données

### 2.1 Les données structurées et non-structurées

L'ensemble des données disponibles sur le web n'est pas exploitable tel quel. Pour certaines d'entre elles ayant déjà fait l'objet d'un traitement, c'est-à-dire qui ont été classées et organisées en lignes et en colonnes (par le biais de tableurs notamment), les rendant ainsi utilisables directement, on parle de données structurées. En pratique, il peut s'agir d'indications concernant les clients (nom, âge, CSP, chiffre d'affaires réalisé, historique de ventes, etc.), un site Internet (trafic, sexe des internautes, heure de connexion, etc.), ou type d'interaction laissant des traces numériques.

Les données non-structurées dominent néanmoins largement la Toile. Présentes majoritairement sous forme de textes écrits, *via* des *posts*, des commentaires clients, des emails échangés au sein d'une organisation, leur contenu nécessite une analyse poussée en vue de les rendre exploitables par les organisations. Le défi majeur réside dans le fait de pouvoir extraire de la connaissance à partir de ces données, très difficilement classifiables pour plusieurs raisons (langues différentes, fautes d'orthographe, phrases longues et complexes, supports variés, etc.) mais qui débouchent sur des décisions plus fines.

Elles sont très souvent précieuses dans le cadre de l'amélioration de la relation clientèle ou du marketing relationnel car elles offrent la possibilité d'avoir une idée approfondie des attentes des consommateurs, leur opinion vis-à-vis d'une marque, leurs centres d'intérêt ainsi que leur comportement sur la Toile.

### 2.2 Données personnelles et données anonymes

Les données personnelles sont partout. Nous laissons chaque jour des dizaines d'empreintes numériques (boîte mail, utilisation de GPS, navigation Internet, etc.) franchissant l'accès à l'information de toute limite spatio-temporelle. En effet, les informations sont productibles et accessibles, de manière persistante et en dehors des cercles *a priori* légitimes que sont l'environnement familial, professionnel et amical. Selon la CNIL (Commission Nationale Informatique et Liberté) les données personnelles sont définies de la manière suivante :

"Constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre

son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne ". (Article 2 de la loi Informatique et libertés)

Les données personnelles sont peuvent être directement nominatives d'une part (nom, prénom, adresse mail, etc.) mais aussi indirectement nominatives avec les numéros de téléphone, l'adresse postale ou encore les liens cliqués et peuvent être classées en deux sous-catégories supplémentaires: les données objectives et les données subjectives<sup>17</sup>.

Les données anonymes sont quant à elles des données, qui *a priori*, ne permettent pas de retrouver un individu de manière nominative. Or dans la réalité, il est prouvé qu'il est possible d'identifier une personne seulement à partir de son sexe, sa date de naissance et son lieu de résidence.

### 2.3 Les données référentielles, transactionnelles et opérationnelles

D'autres types de données peuvent également présenter une importance particulière pour les opérateurs et sont classées en 3 catégories<sup>18</sup> :

- **Les données référentielles** sont souvent disponibles dans les bases de données des entreprises et sont à la base de la connaissance. Elles sont généralement considérées comme immuables en raison de leur fréquence de changement faible, comme par exemple les informations liées à la géographie, au patrimoine, aux normes administratives ou encore les fichiers clients, les services ou les produits proposés. Ces données constituent la base du système d'information des organisations.
- **Les données transactionnelles** résultent de la traçabilité des individus et relèvent de pratiques quotidiennes, tant au sein de l'entreprise qu'à l'extérieur (code barre d'un billet de train, achats auprès des fournisseurs, ventes aux clients, etc.). Elles sont accessibles et manipulables par les logiciels de gestion de relation client (CRM pour Customer Relationship Management), et permettent de prendre des décisions granulaires. Elles présentent néanmoins un souci éthique inhérent à intrusion dans la vie privée.
- **Les données opérationnelles** sont la traduction d'activités publiques qui changent fréquemment (fréquentation d'un site, déplacements de la population, etc.) et représentent un volume considérable. Elles sont utiles à la prise de décisions quotidiennes

---

17

Annexe A : Tableau de conversion des unités de mesure de mémoire informatique

<sup>18</sup> CLAVERIE Bruno. *Économie numérique*. Cours de Master 1 MHR, ISTHIA, Université Toulouse – Jean Jaurès, 2015.

rapides et sont génératrices d'idées et de concepts (*entertainment*, conception de trajets, design, etc.).

### **3. Les sources de production de données**

#### **3.1 Les sites web**

Rares sont aujourd'hui les entreprises à ne pas disposer de leur propre site Internet. Pour un grand nombre d'entre elles, il constitue l'interface numéro un avec ses clients. Ces pages ont plusieurs vocations : achats, présentation des biens et des services proposés, information sur les événements, publicité, etc. L'analyse du trafic et de la navigation des internautes représentent une source d'information immense les concernant : temps passé sur la page, clics, parcours effectué (logs), taux de retour sont autant d'éléments permettant d'améliorer sa page et l'expérience client qui en découle lors de la navigation.

#### **3.2 Les applications pour *smartphones***

De plus en plus d'applications pour *smartphone* ont fait leur apparition ces dernières années. Véritables mines d'or pour leurs éditeurs, ces logiciels téléchargeables sur nos téléphones portables peuvent remplir diverses fonctions : réaliser des achats, trouver un établissement, jouer, regarder une vidéo ou un document, etc. L'ensemble de ces actions fournit des informations précieuses quant aux utilisateurs qui doivent, en contrepartie du téléchargement, accepter de fournir les données relatives à l'utilisation de ces applications (position géographique, messagerie, photos, etc.). L'analyse de ces éléments permet ainsi aux entreprises qui les développent de mieux connaître leur clientèle et mieux la segmenter, d'améliorer les services proposés et augmenter le confort de navigation.

#### **3.3 La géolocalisation**

La géolocalisation est un procédé qui consiste, *via* Internet, à localiser un lieu, un objet ou une personne sur un plan en fonction de sa position. Elle peut se faire en temps réel ou en différé par le biais de différents terminaux que sont les ordinateurs portables, les *smartphones*, ou les objets connectés équipés d'Internet. Ce procédé est de plus en plus utilisé par les internautes comme par les entreprises.

Avec plus de 7 milliards de téléphones mobiles en service dans le monde, la géolocalisation permet de cartographier et d'analyser les mouvements de populations.

### 3.4 Les réseaux sociaux

L'explosion des réseaux sociaux depuis l'apparition de Facebook en 2004 a profondément modifié les rapports entre humains et technologie pour devenir un outil incontournable de la communication moderne. Les entreprises ont depuis quelques années investi ces sites de façon à pouvoir communiquer directement avec leurs clients et de mieux les appréhender.

D'autre part, chaque page *liée*, chaque « ami », constitue une base de renseignement considérable pour les acteurs économiques qui peuvent ainsi avoir un accès direct à leurs centres d'intérêt, leurs avis concernant une marque, et proposer ainsi, directement sur le réseau social ou sur d'autres pages, des suggestions de produits ou de services correspondant au mieux aux différents profils des utilisateurs.

### 3.5 L'Internet des objets

Le cabinet Gartner définit l'Internet des objets comme « un réseau d'objets physiques qui contiennent de la technologie embarquée pour communiquer et comprendre ou interagir avec leurs statuts internes ou l'environnement externe »<sup>19</sup>. La connectivité grandissante des individus a donné lieu à la création de nouveaux objets destinés à améliorer et avoir un meilleur contrôle de leur vie. Équipés de capteurs reliés à Internet, ou de technologies RFID (Radio Frequency Identification), LAN sans fil (réseau local) ou WAN (réseau sur grande distance), ils permettent de fournir de précieuses indications concernant leur état de santé, leurs déplacements, leur consommation énergétique, etc. (DELORT, 2015, p.16-17).

D'abord utilisées dans le domaine des assurances et de la finance, les données fournies à travers ces dispositifs (voitures, réfrigérateurs, bracelets, etc.) donnent l'opportunité aux organisations d'adapter leurs propositions à leurs clients, d'ajuster la tarification en fonction de leurs comportements.

Ce marché florissant n'en est encore qu'à ses prémices. Selon de cabinet de conseil IDATE, le nombre d'objets connectés pourrait passer de 12 milliards en 2012 à 80 milliards en 2020.

---

<sup>19</sup> Gartner. *IT Glossary: Internet of things*. [en ligne]. Disponible sur <https://lc.cx/4m6y>. (Consulté le 12-01-2016).

## 4. Les grands acteurs du *Big Data*

### 4.1 Les développeurs

Les développeurs sont à l'origine de l'évolution numérique, bien qu'ils soient indissociables des techniciens mettant au point les structures informatiques physiques. Si IBM, qui commercialise à la fois du matériel informatique, des logiciels et des services, reste un des pionniers de la *business intelligence*, si bien que de nombreuses entreprises se sont créées et développées sur son modèle pour répondre aux besoins croissants liés aux nouvelles technologies. Ces entités ont pour domaine d'activité la conception de programmes destinés à être exploités par les acteurs économiques et privés, et peuvent être spécialisés dans différents domaines comme la finance, la maintenance, le marketing, etc. Criteo est un excellent exemple de cette dernière activité. Cette compagnie française créée en 2005, aujourd'hui cotée au NASDAC<sup>20</sup>, est spécialisée dans la publicité ciblée et le *retargetting* en ligne dont la Fnac, Priceminister ou *The Guardian* sont clients. Son succès est tel qu'il commence à faire de l'ombre au géant Google<sup>21</sup>.

### 4.2 Le GAFA

Les précurseurs de la collecte et la revente de données sont les multinationales leaders sur les différents marchés que représente Internet, souvent désignées par l'acronyme GAFA :

- Google pour les moteurs de recherche : Rien qu'en France, le géant américain comptabilise à lui seul 46% des parts de marché avec Chrome, le reste étant partagé entre Firefox (21%) et Safari (13%)<sup>22</sup>. Au niveau des requêtes, plus de trente milliards de pages sont indexées par le moteur de recherche qui comptabilise en moyenne trois milliards et demi de requêtes quotidiennes.
- Amazon pour le e-commerce : Cette entreprise reste pionnière en matière de traitement de la donnée, en raison de la création de son système pointu de recommandations.
- Facebook pour les réseaux sociaux : avec près de 3.5 milliards de comptes en 2014, le géant des réseaux sociaux est depuis plus de 10 ans le plus gros fournisseur de données personnelles.

---

<sup>20</sup> National Association of Securities Dealers Automated Quotation, C'est le plus grand marché électronique au monde et le deuxième marché d'action (Source : Wikipedia)

<sup>21</sup> Rushton Katherine. Online advertising firm Criteo spurns Europe for \$500m US flotation. The Telegraph, 20 juillet 2013 [en ligne]. Disponible sur <http://urlz.fr/3iLH>. (Consulté le 21-03-2016).

<sup>22</sup>Stratcounter Global Stats. 5 Desktop, Tablet & Console Browsers in France from February 2015 to February 2016. [en ligne]. Disponible sur <https://lc.cx/4muS>. (Consulté le 18-03-2015).

- Apple pour les appareils connectés (*smartphones*, ordinateurs portables, tablettes, montres connectées...)

Ces acteurs sont en quelque sorte les « grossistes » de la donnée. Ils collectent à eux-seuls la majorité des données disponibles sur la Toile et certains ont fait de leur revente une activité à part entière tel que Google avec ses 3.3 milliards de requêtes quotidiennes en 2015 ou encore Facebook (près de 3.5 milliards de comptes en 2014)<sup>23</sup>. Ces entreprises ont été les pionnières en matière de traitement de données et disposent aujourd’hui d’une influence considérable sur la scène mondiale.

#### 4.3 Les collectivités et acteurs territoriaux

Les collectivités territoriales sont de gros fournisseurs de données. Certaines d’entre elles comme l’État lui-même, met à disposition de tout un chacun des jeux de données divers issus de l’activité humaine sur le territoire (répertoire des musées, fréquentation de lieux touristiques, données sociodémographiques, etc.) et disponibles en *open data*. D’autres opérateurs privés, comme par exemple les compagnies de téléphonie mobile ou les transports (SNCF, Amadeus, etc.), les revendent ou créent des partenariats avec d’autres organismes au sein desquels les informations sont mutualisées à des fins techniques, commerciales, technologiques, etc.



L’ensemble des éléments vus ci-dessus constituent les systèmes d’information utilisés par les entreprises actuelles qui comprennent donc des éléments à la fois informationnels, technologiques et organisationnels et dont la finalité est d’atteindre des objectifs fixés par les dirigeants (REIX, FALLERY, KALIKA, 2011, p. 5-6).

---

<sup>23</sup> Blog du Modérateur. *Chiffres Google 2015*, août 2015. [en ligne]. Disponible sur <http://urlz.fr/37sE>. (Consulté de 24-02-2016).

## Chapitre 3. Approches disciplinaires

Après avoir balayé l'aspect technique et technologique du sujet, il convient à présent de s'attacher à deux autres disciplines liées au *Big Data* que sont l'économie et le Droit.

### 1. Approche économique

En passe de devenir le « pétrole du 21<sup>ème</sup> siècle », la collecte et la revente de données pèsent actuellement un poids énorme dans l'économie mondiale et pourrait, selon le Boston Consulting Group, représenter près de 1 000 milliards d'euros d'ici 2020<sup>24</sup>, offrant une multitude d'opportunités à tous les professionnels. Ce phénomène prend une importance telle qu'il fait partie des priorités du gouvernement pour redynamiser l'industrie française<sup>25</sup>. Au cours de l'année 2014, près de 3.5 zéttaoctets ont été produits mais cette masse reste insignifiante comparée aux 40 zéttaoctets qui seront produits à l'horizon 2020 grâce au développement des objets connectés.

En 2009, le Centre d'Analyse Stratégique (CAS)<sup>26</sup> prévoyait 42 variables d'expression du système de l'économie du numérique entre 2015 et 2025, réparties en cinq grandes composantes : les technologies, le marché, les usages, le contexte et les gouvernances (MUSSO, 2008, p. 18). Aujourd'hui les débats organisés autour de ce sujet ne cessent d'augmenter et impliquent de grands acteurs du territoire comme les personnalités de décisions publiques, de la recherche et le monde de l'entreprise.

Un nombre grandissant de plateformes d'intermédiation ont vu le jour ces dix dernières années (Google, Amazon, Uber, Twitter, etc.) assurant ainsi le lien entre le client et l'entreprise, qui par ces intermédiaires peut établir une relation personnalisée entre ceux-ci. Leur puissance dépend directement du nombre d'utilisateurs et du nombre de traces numériques qu'ils laissent, permettant d'enrichir leurs services mais aussi leurs profits. Deux conséquences découlent de cette position d'oligopole des principaux acteurs du web (Google, Amazon et Facebook): leur concentration géographique (dans la Silicon Valley) et donc leur capacité à contourner les législations européennes<sup>27</sup>.

---

<sup>24</sup> Cité par LeMonde.fr. *Métier de la data : les quatre profils qui vont s'arracher* [en ligne]. Disponible sur <http://urlz.fr/3iMC>. (Consulté le 24-03-2016).

<sup>25</sup> RICHARDIN Anaïs. *La data, pétrole de l'économie numérique*. Stratégies, 13 mai 2014, [en ligne]. Disponible sur <http://urlz.fr/3iVo>. (Consulté le 21-12-2015).

<sup>26</sup> Organisme de réflexion, d'expertise et de concertation placé auprès du Premier ministre.

<sup>27</sup> PREMIER MINISTRE, FRANCE STRATÉGIE. *"Big Data" : création de valeur, enjeux et stratégies de régulation* [en ligne]. Disponible sur <http://urlz.fr/3iW5>. (Consulté le 25-03-2016).

## 2. Approche éthique et juridique

En 2012, les révélations d'Edward Snowden à propos de la NSA ont profondément indigné les citoyens du monde entier, en particulier ceux résidant aux États-Unis<sup>28</sup>. Cet ancien informaticien de la CIA, ayant collaboré avec l'organisme de sécurité du pays, avait confié à la presse américaine et britannique les pratiques des deux entités en matière d'espionnage des organismes politiques mondiaux et des particuliers par le biais de PRISM, un logiciel permettant d'intercepter les flux de conversations et d'actions se déroulant la Toile. Ce scandale a eu pour effet d'alerter les individus de l'utilisation faite des traces numériques produites et soulever de nombreuses questions en matière de droit à la vie privée.

Car si la liberté et la vie privée relèvent des droits fondamentaux des citoyens, il n'en demeure pas moins que chaque mouvement, chaque clic, chaque conversation est enregistré, dupliqué et stocké de manière *quasi* définitive. Pour Céline Castets-Renard<sup>29</sup>, professeure de Droit privé et membre de l'IRDEIC<sup>30</sup>, le problème majeur relève du fait que la technologie progresse bien plus vite que le Droit. Les définitions et l'encadrement de la législation concernant l'exploitation de données sont basés sur des textes obsolètes qui peinent à évoluer en même temps que la société.

Le texte de référence, la Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, définit l'usage de l'informatique comme de suit<sup>31</sup> :

« L'informatique doit être au service de chaque citoyen. Son développement doit s'opérer dans le cadre de la coopération internationale. Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques. » (Article 1)

Ce texte, bien que modifié et enrichi en août 2004 et janvier 2016, ne concerne que la France. La CNIL (Commission nationale de l'Informatique et des Libertés) a pour mission de veiller à l'application de cette loi de sorte qu'elle ne porte atteinte ni aux droits de l'homme ni à la vie privée, ni aux libertés individuelles et publiques. Or, plusieurs facteurs remettent en question son efficacité et contribuent à la difficulté de création de lois cohérentes.

---

<sup>28</sup> SZADKOWSKI Michaël, LELOUP Damien. *Prism, Snowden, surveillance : 7 questions pour tout comprendre*. Le Monde Technologies, [en ligne]. Disponible sur <https://lc.cx/4mEm>. (Consulté le 26-02-2015).

<sup>29</sup> Lors de la conférence *Les enjeux juridiques qu'impliquent le partage et les usages des Big data* dans le cadre de la Grande Conf Sciences humaines et sociales et *Big Data*, Toulouse, le 21-02-2016.

<sup>30</sup> Institut de recherche en droit européen, international et comparé

<sup>31</sup> Legifrance. *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés*, 2013, 76 p. [en ligne]. Disponible sur <https://lc.cx/4mrB>. (Consulté le 21-11-2015).

Tout d'abord, la diversité de la nature des données qui appartiennent à différents régimes de propriété (privée, intellectuelle, industrielle) nécessitent une législation parcellaire contraignante, au cas-par-cas, ralentissant considérablement l'évolution des lois. D'autre part, la prédominance du Droit européen sur le Droit français, dont les textes et les objectifs sont parfois divergents et ne permet pas une action commune dans tous les pays de l'Union. Théoriquement, le gouvernement européen stipule l'obligation d'informer et d'obtenir l'accord des individus du recueil et l'exploitation des données qu'ils fournissent. Dans les faits, de nombreux sites et applications installent des mouchards dans les systèmes que nous utilisons et peuvent ainsi recueillir de nombreuses données comme le nom, l'adresse mail ou encore les identifiants de réseaux sociaux. En plus de cela, une grande majorité d'entre elles sont stockées dans des *data center* situés en dehors du pays et dupliquées sur différents serveurs, un peu partout dans le monde à l'image de Facebook, Google ou Amazon qui constituent des lobbyistes d'un poids déterminant et contre lesquels les gouvernements ne peuvent faire face.



Ce chapitre nous a permis de comprendre, au-delà de leur dimension technique, les répercussions de l'explosion des données disponibles. Car si l'aspect technologique reste incontournable, les pendants économiques et juridiques en sont indissociables. Nous avons également identifié les difficultés de contrôle liées à ces nouvelles pratiques, dont la banalisation a fait apparaître un marché florissant en pleine expansion suscitant de nombreuses interrogations, aussi bien pour les entreprises, que pour les citoyens que nous sommes.

## Conclusion

La première partie de ce mémoire a permis d'éclaircir et de dégager les concepts fondamentaux nécessaires à la bonne compréhension de ce sujet. Nous avons tout d'abord pu constater que la collecte et l'analyse des données ne relèvent pas de faits nouveaux, mais que l'avènement des outils technologiques a grandement contribué à leur efficacité et à leur généralisation. Le second chapitre a permis d'identifier, de manière non exhaustive, les principaux outils utilisés (supports, programmes, etc.), les biais par lesquels collecter des informations (*smartphones*, sites web, etc.), leur nature (structurées, non-structurées, opérationnelles, transactionnelle, etc.) et l'utilisation qui peut en être faite, ainsi que les principaux acteurs à l'origine de cette nouvelle approche de l'étude des populations. Enfin, nous avons constaté qu'au-delà de son aspect technique, le *Big data* englobe également des problématiques d'ordre économique et juridique.

De nombreux enjeux découlent de cette tendance, en particulier pour les entreprises qui doivent désormais repenser leur manière d'aborder un marché en mutation constante. Comment tirer profit de ces nouvelles techniques ? Quel est l'intérêt de les appliquer au sein de la stratégie d'une organisation ? Quelles sont les conditions de leur succès ? Ces interrogations sont à l'origine de la question de départ suivante : **Quel est l'intérêt pour les entreprises d'adopter une stratégie centrée sur les *datas* ?**

Nous appliquerons par la suite cette réflexion au domaine de l'hôtellerie, un secteur de poids dans l'économie française. Nous verrons notamment de quelle façon son intégration a réussi, en peu de temps, à modeler l'approche mercantile des acteurs économiques, leur façon d'aborder la stratégie marketing et les opportunités qui en découlent.

Plusieurs hypothèses seront mises en avant, visant à répondre à la problématique découlant des différentes interrogations imposées par la première partie de ce mémoire : **En quoi l'exploitation des données peut-elle permettre aux hôteliers d'affiner leur approche marketing et satisfaire au mieux les besoins et les attentes de leurs clients ?**

**Partie II.     *Big data* et marketing : quelles  
opportunités pour les entreprises ?**

## Introduction

Au même titre qu'Internet il y a de cela quelques années, le *Big data* est en passe de révolutionner une nouvelle fois notre rapport à la technologie. L'ensemble des acteurs mondiaux, privés ou publics, est aujourd'hui amené à réfléchir à de nouvelles organisations, à de nouveaux processus engendrés par ce déluge de données. De nombreux secteurs d'activité se sont déjà engouffrés dans cette tendance actuelle : santé, aéronautique, automobile, énergie, tous ont misé sur des stratégies *data driven* afin de parfaire leurs services et rester compétitifs. Et l'industrie du tourisme n'est pas en reste. Car si l'ère du *Big data* est l'ère du changement pour les entreprises, il en va de même pour la société, dans laquelle Internet et la connectivité occupent désormais une place centrale.

La seconde partie de ce mémoire vise à appréhender les opportunités, marketing notamment, que peut revêtir le *Big data*. Nous mettrons en avant les mutations engendrées par l'âge d'or de l'information et des échanges instantanés, notamment dans la stratégie des entreprises en lien avec le secteur du tourisme. Nous traiterons de son impact organisationnel de manière globale, puis de ses apports dans les différents éléments du plan marketing. Nous terminerons ensuite avec les enjeux et les limites de l'intégration des *datas*.

Ce sera également l'occasion de répondre aux hypothèses découlant des interrogations relatives aux recherches effectuées dans le cadre de la première partie :

**Hypothèse n°1 :** L'exploitation des données n'est réservée qu'aux grands groupes qui, contrairement aux hôteliers indépendants, disposent de moyens (financiers, humains et matériels) importants.

**Hypothèse n°2 :** Les hôteliers doivent développer de plus en plus de stratégies afin de pouvoir collecter des données pertinentes et utiles à la prise de décision.

**Hypothèse n°3 :** La collecte et le traitement de données émises sur Internet peut se substituer aux méthodes de collecte traditionnelles en matière de comportement des clients (questionnaires de satisfaction, entretiens consommateurs, enquêtes, etc.).

Afin de répondre à ces hypothèses, nous étudierons dans la partie suivante de ce mémoire les stratégies marketing déployées par les acteurs économiques ainsi qu'un état des lieux du parc hôtelier français. La troisième partie portera ensuite sur l'analyse des pratiques des grands groupes de notre secteur ainsi que celles des indépendants.

## Chapitre 1. Le *Big data* dans les entreprises

La numérisation de l'information engendrée par l'essor de l'informatique a bouleversé le fonctionnement de l'ensemble des opérateurs privés et publics. L'acceptation générale et la diffusion des technologies à l'ensemble de la société ont engendré de profondes mutations d'ordres économiques et sociologiques qui ont mené les entreprises à revoir leur approche pour pouvoir s'adapter à de nouveaux paradigmes. Ce premier chapitre fera un état des lieux et des changements opérés au cours de la dernière décennie.

### 1. Vers un changement des processus décisionnels

La prise de décision, quelle que soit l'organisation, est incontournable dans une optique de pérennisation et d'adaptation à l'environnement. De nombreux sociologues se sont penchés sur cette thématique afin de mieux comprendre quelles sont les facteurs influençant les choix des dirigeants. Au fil des décennies, différentes théories ont vu le jour grâce aux travaux de chercheurs.

L'une des plus célèbres reste celle de la rationalité limitée mise en avant en 1958 par James March et Herbert Simon dans leur ouvrage *Organisations*. Elle repose sur le principe selon lequel les acteurs, quels qu'ils soient, disposent d'une quantité réduite d'informations à leur connaissance et sont donc amenés à faire des choix sans pouvoir prendre en compte l'ensemble des possibilités qui s'offrent à eux. Les contraintes matérielles, humaines et financières ainsi que l'analyse séquentielle d'un problème les poussent généralement à appliquer les décisions qui sont les plus simples. Ces préceptes mettent l'accent sur les limites des capacités cognitives et de la disponibilité des informations.

Ce concept s'oppose à celui de l'anticipation rationnelle défendu par Robert LUCAS dans *Expectations and the neutrality of money*<sup>32</sup> qui, s'inspirant des travaux de ses confrères Mutt et Friedman, affirme que les individus disposent de suffisamment d'informations pour pouvoir anticiper et ajuster leur comportement dans un contexte économique donné et en déduire les conséquences qui en découlent. Il suffit pour cela de se baser sur les performances passées de l'entreprise. Cette théorie s'oppose aux principes keynésiens dominants à cette époque.

Au vue des progrès effectués en matière de TIC depuis, ces préceptes tendent à être dépassés. En effet, l'ère du tout numérique a eu pour effet de libérer l'accès aux données à tout un chacun

---

<sup>32</sup> Portail de l'économie et des finances. *Faciléco* : Robert Lucas [en ligne]. Disponible sur <http://urlz.fr/3itv>. (Consulté le 19-01-2016).

à propos d'une infinité de thématiques ainsi qu'à des outils de plus en plus perfectionnés qui se révèlent être un atout de taille pour les décideurs. D'après Bill SCHMARZO (2014, p.3) cette mutation sociotechnique tend à transformer l'acte de décision classique qui semblait jusqu'alors immuable.

Tableau 1 - Différences entre décision classique et décision *data* centrée

Prise décision classique	Prise de décision à l'ère du <i>Big data</i>
Vision rétrospective	Vision prospective, recommandations
Disponibilité limitée des données (10 %)	Sources de données multiples et en réseau
Données en lots, incomplètes, disjointes	Données en temps réel, organisées
Supervision de l'activité	Optimisation de l'activité

Source : JOUANNY Jean-Simon (2014, p. 27).

L'apport de l'informatique a permis de baser la réflexion sur des faits aisément quantifiables et non plus sur des hypothèses. En plus des données issues des résultats d'exploitation antérieurs de l'entreprise (*reporting*), il est désormais possible d'analyser les événements en cours (*monitoring*) mais également ceux à venir (prédiction) avec un degré de fiabilité et de précision bien supérieur à ceux d'antan.

## 2. Les secteurs impactés

Nombreux sont les secteurs à disposer d'un volume de données conséquent s'inscrivant de manière volontaire ou involontaire dans une démarche *Big data*. Les premières industries à s'être engouffrées dans une stratégie centrée sur les *datas* ont été, en dehors des précurseurs tels que Google ou Yahoo!, celles de la santé, de la finance, des télécoms et des transports très vite suivies par l'ensemble des domaines d'activité. Le tourisme n'est pas en reste et s'inscrit désormais dans cette tendance générale qui vise à donner la primauté à l'information dans le processus de décision. Le point commun de toutes ces entités est que la production d'informations diverses fait partie intégrante de leurs métiers.

Dans un contexte de concurrence exacerbée auquel l'ensemble les acteurs économiques doivent faire face, l'introduction des nouvelles technologies dans les organisations relève de l'incontournable. Si la mise en place d'une démarche ne concerne encore que les grands groupes, de plus en plus de petites structures ont aussi recours à des outils simples pour dynamiser leurs

ventes et les relations avec leurs clients. Les opérateurs majeurs de l'industrie du tourisme ont eux aussi intégré l'importance des métadonnées, à l'image d'Amadeus devenu la référence au niveau des GDS (*Global Distribution System*).

### 3. La mise en place d'une démarche *Big data*

#### 3.1 Constat

D'après une étude réalisée par EMC fin 2013<sup>33</sup>, 74 % des entreprises en France sont convaincues de l'intérêt du *Big data* pour leur activité. En réalité, à l'heure actuelle, 41 % d'entre elles n'ont encore engagé aucune dépense en la matière. Plusieurs raisons expliquent cette différence de proportion mais la principale reste l'imprévisibilité du retour sur investissement sur lequel 35 % des dirigeants se posent des questions.

L'outillage occupe une place de taille dans l'analyse des données, c'est pourquoi il est primordial de savoir bien s'équiper. Le problème rencontré dans six entreprises sur dix est le traitement de données non structurées, qui ne sont pas prises en compte dans les solutions de *reporting* classiques.

Selon une enquête menée par NetMediEurope en 2015<sup>34</sup>, les données devraient être un pôle majeur d'investissement au cours des 3 prochaines années. Le rapport a permis d'établir le constat suivant : seulement 14 % des organisations de plus de 1 000 salariés ont déjà fait l'usage d'une plateforme *Big data* mais son introduction est prévue dans 55 % des interrogés appartenant à cette catégorie. Ces chiffres sont cependant largement inférieurs lorsqu'il s'agit de petites et moyennes entreprises dont l'effectif comporte moins de 100 employés.

Bien que l'échantillon soit composé d'entreprises venant de divers secteurs, il permet néanmoins d'avoir une idée des pratiques des acteurs économiques. Dans la mesure où le secteur de l'hôtellerie est composé en grande partie de petites structures familiales, nous pouvons supposer que la banalisation de l'usage des *Big data* se fera de manière extrêmement lente et que seuls les grands groupes en bénéficieront.

---

33 Cité par *Big data* Paris. Guide du *Big data*: L'annuaire de référence à destination des utilisateurs, 2015, p.5 [en ligne]. Disponible sur <http://urlz.fr/3ipE>. (Consulté le 05-11-2015).

34 Cité par Fléchaux Reynald. *Big Data* : les entreprises ne sont pas au niveau. Silicon.fr, 24 mars 2016 [en ligne]. Disponible sur <http://urlz.fr/3ifi>. (Consulté le 26-03-2016).

### 3.2 L'émergence de nouveaux métiers

Les technologies nécessaires à l'intégration de la composante *Big Data* dans une organisation nécessitent un savoir-faire et des compétences particulières dont ne disposent pas forcément l'ensemble des dirigeants. Bien que certains outils, comme des tableurs ou des logiciels CRM soient accessibles à tout un chacun, la capacité à en retirer des informations pertinentes reste une question en suspens pour beaucoup.

Face à ces besoins croissants d'analyse, de nouveaux métiers ont émergé et la demande ne cesse d'augmenter du côté des professionnels qui se tournent désormais vers des experts en informatique (*data scientists*, *managers* d'entrepôts de données, *chief data officers*, *data protection manager*, etc.) plutôt que vers des individus issus du secteur concerné. Cette manœuvre présente néanmoins un frein de taille : le fossé générationnel qui sépare souvent dirigeants et spécialistes des *datas*. Dans un contexte où la plupart des comités exécutifs sont composés de membres de plus de 50 ans bénéficiant d'un long parcours professionnel et d'habitudes bien ancrées, ils font parfois preuve de réticence à l'idée de faire confiance aux *digital natives* (ou génération Y) dont l'âge ne dépasse pas les 35 ans<sup>35</sup>. Cette génération a pourtant développé une aisance naturelle, voire une expertise, vis-à-vis des nouvelles technologies dans la mesure où elle y a eu accès dès le plus jeune âge et a grandi avec.

Ces éléments ont naturellement un impact sur la question de la culture d'entreprise et de son évolution qui doit être en accord avec les mutations sociétales. Certaines entreprises ont tout de même franchi le pas suite au succès des *start-up* développées au cours des dernières années et qui, pour certaines, sont cotées en bourse après seulement quelques années d'activité (Vente-privee, Blablacar, VeryLastRoom, etc.). C'est le cas d'Accor qui depuis a misé sur une collaboration avec une toute jeune entreprise (Shadow Comex) dont les salariés ont tous moins de 35 ans. Pour son patron Sébastien Bazin, le pari s'avère être une réussite puisque qu'il a récemment déclaré que « *Plus aucune décision ne sera prise sans leur demander leur avis* »<sup>36</sup>.



Ce chapitre a fait état du niveau d'implantation du *Big data* au sein des firmes françaises toutes catégories confondues. Il en ressort que si une grande majorité pressent que cette nouvelle révolution numérique sera incontournable dans les années à venir, beaucoup d'entre elles n'ont pas encore franchi le pas de les instaurer comme pilier central de l'acte de décision.

---

<sup>35</sup> Les Echos Business. *Portrait-robot du manager digital* [en ligne]. Disponible sur <http://urlz.fr/3ipx>. (Consulté le 24-03-2016).

<sup>36</sup> *Ibid.*

## Chapitre 2. Les effets du *Big Data* dans la stratégie marketing

Face à une civilisation de plus en plus connectée grâce à des technologies en progrès constant, la rapidité et l'efficacité sont de rigueur pour l'ensemble des acteurs économiques quel que soit leur secteur. Le marketing est la fonction dans laquelle le *Big data* a pris une place prépondérante au cours des dernières années. Il est vrai que les nombreuses possibilités qu'il offre constituent une arme de poids et une chance à chacun de pouvoir se démarquer de la concurrence grâce à une meilleure compréhension du client. Plusieurs composantes et dimensions du plan marketing sont ainsi abordées sous un nouvel angle, dans le but d'augmenter les ventes et la fidélité du client.

### 1. La stratégie marketing

#### 1.1 Définition

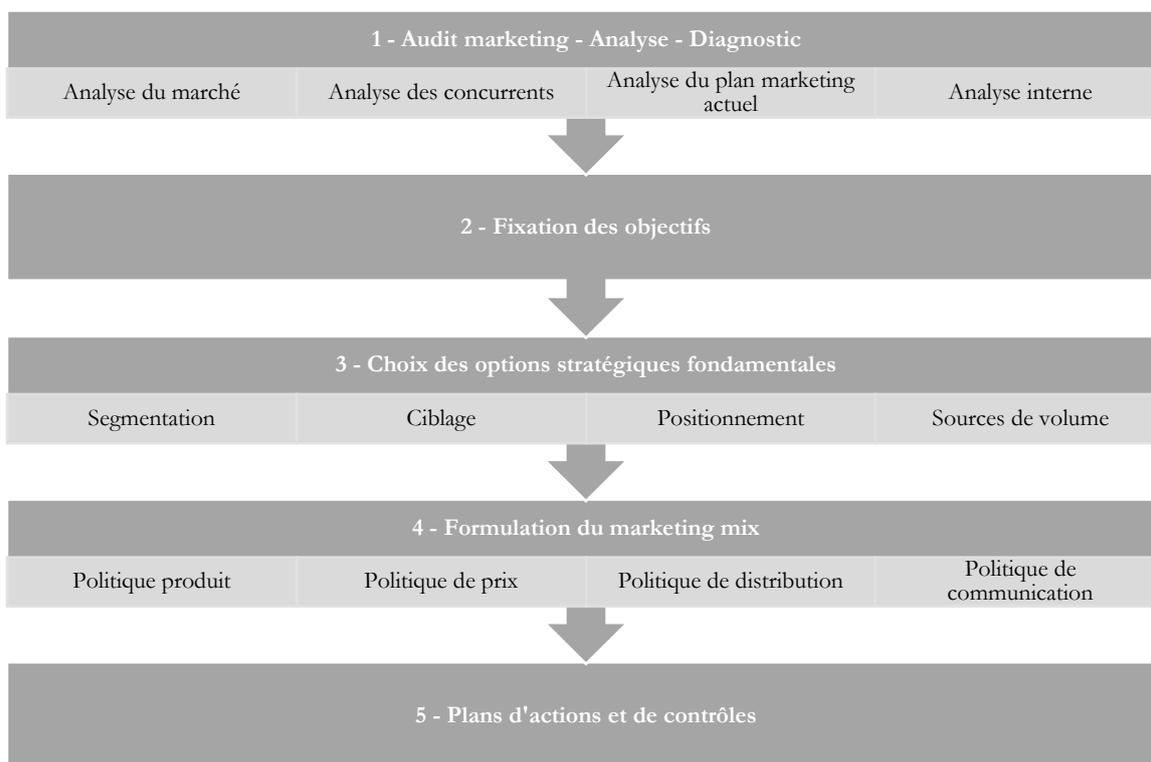
La stratégie marketing est depuis plus d'un demi-siècle au cœur de la réflexion des entreprises de tous horizons. Apparu il y a plus d'un demi-siècle, il a évolué en accord avec la société et les lois du marché pour devenir un élément primordial dans la stratégie globale des entreprises de biens et de services.

D'après le Mercator (2013, p.3) « Le marketing est l'effort d'adaptation des organisations à des marchés concurrentiels, pour influencer en leur faveur le comportement des publics dont elles dépendent, par une offre dont la valeur perçue est durablement supérieure à celle des concurrents. Dans le secteur marchand, le rôle du marketing est de créer de la valeur économique pour l'entreprise en créant de la valeur perçue par les clients ». P. KOTLER (2009, p. 5) ajoute à cette définition la notion de « mécanisme social », qui prend tout son sens à l'heure de l'hyper-connectivité de la société actuelle. Car les *Big Data* ne sont-elles pas, au fond, le reflet de ce que nous sommes et des questions que nous nous posons ?

En hôtellerie, ces définitions prennent un véritable sens dans la mesure où l'objet de la vente est un service, par définition intangible. Sa production et sa consommation se faisant de manière concomitante, la qualité de la prestation se doit d'être homogène quelle que soit la personne servie.

Le déroulement du plan marketing se fait de la manière suivante<sup>37</sup> :

Figure C - Déroulement du plan marketing



Dans un contexte d'importance primordiale accordée aux données, l'ensemble des éléments ci-dessus peuvent être impactés de manière plus ou moins prononcée et tendent à changer les façons classiques d'aborder chaque problématique. Nous nous concentrerons par la suite sur les deux étapes pour lesquelles l'implantation du *Big data* relève d'un enjeu de taille, à savoir le choix des options stratégiques et le mix marketing.

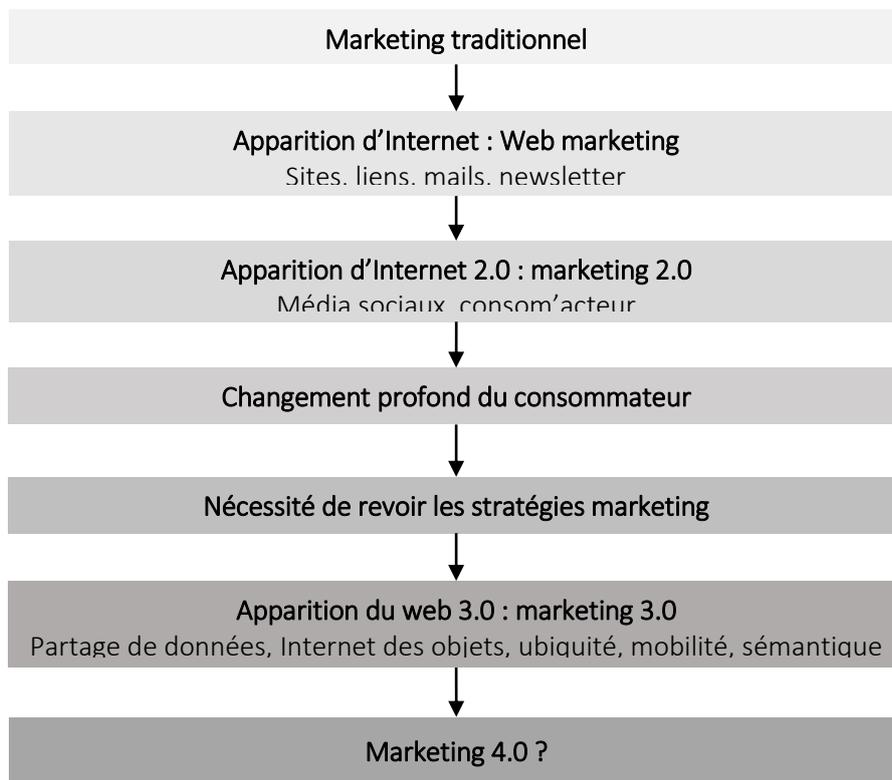
## 1.2 L'arrivée d'Internet et son impact sur le marketing

Comme vu précédemment, Internet et toutes les actions qui en découlent ont évolué pour devenir le puissant outil que nous connaissons aujourd'hui. Pour Carole MARSELLA (2011, p.115-116), enseignante chercheuse à l'ICD<sup>38</sup> de Paris, nous avons assisté au cours des deux décennies à un échange de jeux de pouvoir entre entreprise et client :

<sup>37</sup> PICHON Paul-Emmanuel. *Marketing stratégique*. Cours de Master 1 MHR, ISTHIA, Université Toulouse - Jean Jaurès, 2015

<sup>38</sup> Institut International du Commerce et du Développement

Figure D - Évolution du marketing sous l'impulsion d'Internet



Source : Carole MARSELA, (2011, p. 116)

À ses débuts, Internet présentait un caractère unilatéral. Les internautes pouvaient effectuer une recherche et accéder à des contenus mis en ligne par un nombre limité d'acteurs. Quelques offreurs étaient alors présents, à l'image d'Amazon, et la concurrence dans le domaine de l'e-commerce encore très limitée, offrant ainsi un avantage concurrentiel de taille à ces organismes. Avec la propagation de l'informatique dans le foyer de milliers, puis de millions de personnes, les contributions à l'enrichissement de la Toile ont explosé, revêtant désormais un aspect communautaire indéniable faisant évoluer l'Internet classique vers l'Internet 2.0. Cette nouvelle forme de partage qui a considérablement facilité l'accès à l'information fait la part belle aux interactions entre les individus de tout bord grâce à l'avènement des sites communautaires comme Tripadvisor, permettant à tout un chacun de donner son avis sur une prestation de service consommée. Ces sites ont grandement contribué au renversement des rapports de pouvoir entre clients et entreprises.

## 2. Les apports dans les options stratégiques fondamentales

### 2.1 La segmentation

La segmentation est à la base de toute stratégie car elle permet de déterminer des groupes homogènes présentant des caractéristiques similaires comme le lieu d'habitation, l'âge, le sexe, les centres d'intérêt, les comportements d'achat, etc. et adapter ainsi son offre et son message à chaque destinataire. Les critères Récence-Fréquence-Montant (RFM), socio-démographiques, comportementaux, style de vie étaient jusqu'ici les plus utilisés pour dégager des profils-type sans pouvoir toutefois avoir une connaissance plus poussée des consommateurs. C'est en ce sens que les *Big data* revêtent une importance décisive. Pour M. TEBOUL (2015, p. 53), Les critères rationnels classiques, appuyés sur des faits objectifs permettent difficilement de prendre en compte l'influence de facteurs tiers comme les désirs, les valeurs, les préférences. Les progrès réalisés en matière d'analyse permettent désormais de constituer des profils de plus en plus précis en fonctions de différents points de données auxquels les entreprises peuvent avoir accès comme le montre la figure ci-dessous :

Figure E - Tableau des variables clients et leur définition

Exemple d'étiquette de variable	Valeur de la variable (et définition)
@fields.arg_ipClient	10.30.50.78 (adresse IP du visiteur)
@timestamp	2014-04-11T07:29:54.119Z (date et heure de l'évènement)
@fields.geoip.ip	109.10.125.56 (coordonnées GPS de lu visiteur)
@fields.arg_url	<a href="http://www.1001pharmacies.com/dodie-sucettes-0-6-mois-anatomique-silicone-message-maman-et-transparente-x2-p6651?utm_source=google-shopping&amp;utm_medium=comparateur-google&amp;utm_campaign=dodie-sucettes-0-6-mois-anatomique-silicone-message-maman-et-transparente-x2&amp;gclid=CNjfh571170CFQoIwwodELIAZQ">http://www.1001pharmacies.com/dodie-sucettes-0-6-mois-anatomique-silicone-message-maman-et-transparente-x2-p6651?utm_source=google-shopping&amp;utm_medium=comparateur-google&amp;utm_campaign=dodie-sucettes-0-6-mois-anatomique-silicone-message-maman-et-transparente-x2&amp;gclid=CNjfh571170CFQoIwwodELIAZQ</a> (URL de la page visitée)
@fields.arg_os	Windows (système d'exploitation mobilisé par le visiteur)
@fields.nb_visit	1 (nombre de visite du site réalisé par le visiteur)
@fields.page_type	["product"] (type de page consultée par le visiteur)
@fields.productView.price_c	2.96 (prix du produit consulté par le visiteur)
@fields.arg_pigReco	click (indique si la page consultée est issue d'un clic recommandation)
@fields.arg_reco_id	["6650"] (identité de la recommandation cliquée)
@fields.arg_reco_position	["0"] (position de la recommandation cliquée)
Etc.	

Source : VAYRE,

L'ensemble de ces données est issu de la navigation sur une page web et permet de cumuler un certain nombre d'informations concernant les requêtes de la personne en question. Certaines entreprises vont bien plus loin dans l'identification de leur clientèle. En 2014, IBM annonçait la sortie de Bluemix Watson, un programme disponible en *open source*, capable de mettre en avant plus de 40 variables explicatives du comportement *via* l'analyse de données non structurées<sup>39</sup> qui

<sup>39</sup> Annexe C : Programme Watson développé par IBM

constituent aujourd'hui la majeure partie du web<sup>40</sup>, lui permettant de comprendre le comportement humain et l'environnement dans lequel il évolue. Ce type d'outil permet, outre une segmentation plus fine, de découvrir de nouveaux marchés de niche sur lesquels s'aventurer, mais surtout d'identifier des opportunités en temps réel qui pourraient constituer un avantage concurrentiel de taille. D'autres utilisations peuvent en être faites à l'image de la collaboration récente d'IBM avec le groupe Marriott, dont les grandes lignes seront exposées ultérieurement.

## 2.2 Le ciblage

Alors qu'il était auparavant difficile pour les entreprises de pouvoir cibler un segment de clientèle particulier à travers les media classiques (Radio, télévision, presse, affichage), Internet a considérablement changé la donne à ce sujet en offrant la possibilité de s'adresser à un public particulier et d'en mesurer l'impact. Plusieurs politiques sont donc applicables, qu'elles soient différenciées (adaptation de l'offre à différents segments à travers un produit ou une marque), concentrées (ciblage d'un segment en particulier), ou encore individualisées (adaptation à un micro segment ou des clients individuels) (LENDREVIE, LÉVY, LINDON, 2009, p. 826).

Comme pour la segmentation, l'apport des *datas* donne lieu à l'étude des comportements en temps réel à travers, par exemple, l'examen de la position géographique qui déclenche l'envoi d'un message promotionnel à un individu situé à proximité de l'enseigne ou d'un lieu précis<sup>41</sup>. Pour Monsieur X, de nombreuses opportunités résident dans l'exploitation des réseaux sociaux. En important, comme dans le cas du groupe dont il est le directeur marketing, le nom des clients présents dans leur base de données propres dans celle de Facebook. Cette opération permet dans un premier temps de déterminer les profils des clients effectifs de la marque (*custom audience*), puis dans un second temps, de faire apparaître les utilisateurs n'ayant jamais consommé de prestations dans un des établissements de l'enseigne mais présentant des caractéristiques similaires aux consommateurs effectifs (*similar audience*). Une fois cette opération effectuée, un message promotionnel est alors envoyé sur la page d'accueil de ces prospects. D'autres outils, comme le *Data-as-a-Service* (DaaS) se basent sur l'observation des comportements en temps réel à partir de l'étude de *posts*, ayant par exemple pour objet un départ en vacances, auquel l'entreprise répondra par une offre répondant aux attentes de l'auteur.

---

<sup>40</sup> IBM Bluemix. *Getting started with the Personality Insights service* [en ligne]. Disponible sur <http://urlz.fr/3hB0>. (Consulté le 13-01-2016).

<sup>41</sup> DataFloq. *Using Big Data to Target In-Market Shoppers* [en ligne], Disponible sur <http://urlz.fr/3h7r>. (Consulté le 20-03-2016).

Deux stratégies peuvent ainsi être adoptées : extensive, qui vise à conquérir de nouveaux clients pour augmenter les ventes, et intensive qui s'applique plutôt à augmenter les quantités ou la valeur des produits achetés par les clients en fonction de ses centres d'intérêt (par exemple EasyJet qui propose sur son site l'achat de chambres d'hôtel ou la location de voitures en plus de billets d'avion) (LENDREVIE, LÉVY, LINDON, 2009, p. 828). Des cibles prioritaires peuvent aussi être dégagées *via* des méthodes de *scoring* effectuées à partir des données clients récoltées en interne (postes de dépenses les plus élevés, fréquence de consommation, etc).

Bien que les possibilités offertes par les technologies soient *quasi* infinies, une étude menée par Razorfish et Adobe<sup>42</sup> a néanmoins démontré que 76 % des marketeurs échouaient dans l'exploitation de ces données comportementales à des fins de ciblage<sup>43</sup>.

### 2.3 Les sources de volume

La définition des sources de volume donnée par le Mercator (2009, p. 829) est la suivante : « Origine des ventes d'un produit : nouveaux clients de la catégorie de produits, clients de la concurrence, augmentation des quantités consommées par les clients actuels, etc. ». Cette composante des options stratégiques fondamentales étroitement liée au ciblage peut être de trois sortes : la cannibalisation volontaire (d'autres produits de l'entreprise), la concurrence directe (produits de même catégorie vendus par la concurrence) et la concurrence élargie diffuse (produits appartenant à d'autres catégories) (*Ibid.*). Ces différentes stratégies sont généralement combinées par les entreprises. On peut par exemple citer l'exemple de Sofitel dont la clientèle est composée de gens qui n'y séjournent pas, de ceux qui y séjournent occasionnellement en alternance avec d'autres enseignes et des clients qui séjournent exclusivement dans les hôtels de la chaîne. L'objectif de la marque est alors de fidéliser les clients qui ont tendance à fréquenter d'autres établissements au lieu de concentrer ses efforts sur le captage de nouveaux clients. Étant données les possibilités de personnalisation offertes par le *Big data*, cette manœuvre semble être la plus adaptée.

### 2.4 Le positionnement

Le positionnement a pour finalité d'influencer la perception qu'un client a d'une marque ou d'une entreprise à travers ses choix. Cette démarche s'inscrit sur le long terme et doit être cohérent avec l'ensemble de la stratégie marketing, en particulier au niveau du mix marketing.

---

<sup>42</sup> Une agence de marketing et un éditeur de logiciels (photo, vidéo, lecture de PDF, etc.) américains.

<sup>43</sup> Cité par Datafloq. *Using Big Data to Target In-Market Shoppers* [en ligne]. Disponible sur <http://urlz.fr/3h7r>. (Consulté le 21-03-2016).

La compétitivité accrue de l'ensemble des secteurs mène ces derniers à devoir se distinguer de leurs concurrents pour attirer de nouveaux clients et fidéliser ceux ayant déjà consommé un bien ou un service. La différenciation repose sur quatre grands axes : les performances du produit, l'imaginaire construit autour de la marque, les publics ciblés et le mode d'utilisation (LENDREVIE, LÉVY, 2013, p. 767-784).

L'ensemble des technologies aujourd'hui à la disposition des individus permet de faire passer des messages, des actions visibles par le biais de différents supports (vidéo, photo, texte, etc.) et canaux (réseaux sociaux, page web, emailing, etc.). Bien entendu, l'ensemble de ces manœuvres ne remplacera pas l'avis personnel que se fera un consommateur d'un produit lors de son utilisation, mais peut contribuer à entretenir la représentation qu'il s'en fait. Les données non-structurées recueillies par le biais de retours d'expériences clients sur les réseaux sociaux par exemple, constituent des indicateurs utiles et permettent de très rapidement travailler sur le repositionnement en cas de dissonances entre image voulue et image perçue.

### **3. Apports dans le mix marketing**

#### **3.1 La politique de produit**

Le produit, qu'il soit matériel ou immatériel, est à la base des revenus de toute entreprise. Quelles que soient les qualités (objectives ou subjectives) qui lui sont attribuées, il permet d'identifier une marque et de la différencier de ses concurrents (téléphonie, automobile, assurance, compagnies aériennes, etc.). Face au contexte actuel, chaque entité se doit d'innover afin de séduire des consommateurs versatiles et extrêmement informés qui n'hésitent pas à faire jouer la concurrence pour dénicher la meilleure affaire possible. Face à ces nouveaux comportements, la différenciation se fait, outre sur le prix qui reste l'élément de décision numéro un dans acte d'achat, sur le produit et la valeur ajoutée qui lui est incombée. Les PDG de tous bords en sont conscients et 82 % d'entre eux affirment aujourd'hui prendre en compte les attentes des clients dans la définition d'un nouveau produit ou service et 90 % dans les cinq ans à venir (COINTOT et EYCHENNE, 2014, p. 14)

Au sein d'un hôtel, c'est la chambre qui est l'élément principal de la prestation. Cependant, à produit égal, c'est l'ensemble des prestations et des composantes annexes qui feront la différence auprès des clients (restauration, spa, accueil, etc.). Pour les attirer et les fidéliser, les industries collaborent désormais avec ceux qui seront amenés à consommer le produit à travers le *crowdsourcing*. On peut par exemple citer l'exemple de Danone avec le vote organisé autour des nouvelles saveurs de Danette que les consommateurs souhaiteraient retrouver dans les rayons

de leur supermarché. Ce genre d'opération présente plusieurs avantages : Tout d'abord il permet d'être au plus près des attentes des consommateurs en limitant les coûts liés aux études de marché et enquête sur le terrain, mais il implique aussi le futur client dans la démarche de conception d'un nouveau produit et parvient de manière détournée à attirer son attention et à créer un lien avec lui.

### 3.2 La politique de prix

Le *Big data* occupe une place prépondérante dans la politique de fixation de prix. Si plusieurs méthodes existent pour y parvenir, l'avantage premier des nouvelles technologies est la possibilité d'en appliquer plusieurs de manière simultanée en fonction de différents facteurs.

Airbnb et Uber illustrent à eux deux les différentes possibilités offertes par la compilation de données en vue de fixer le prix d'un service. Chez le premier, lors de la mise en ligne d'une annonce, un prix est recommandé par le site, qui se base sur l'analyse des tarifs pratiqués par les locations présentant les mêmes caractéristiques (type de logement, type de quartier, etc.), mais laisse néanmoins à l'hébergeur le choix de la tarification finale<sup>44</sup>. Dans le second cas, c'est l'entreprise elle-même qui fixe le montant de la course. Pour y parvenir, elle identifie la date, l'heure et le lieu de réservation et en déduit ainsi un prix qui pourra être accepté ou refusé par le conducteur et le passager lors de la commande<sup>45</sup>.

Ces pratiques sont comparables à celles observées en hôtellerie et en tourisme avec le *yield management* qui « consiste à proposer des tarifs différents pour gérer au mieux les capacités (réduction des prix pour remplir les capacités ou augmentation des prix quand elles vont être saturées) et maximiser la contribution à la marge. Cette méthode est surtout employée dans les services (quand les capacités de production sont fixes et que les produits ne peuvent pas être stockés). » (LENDREVIE, 2014, p. 286).

### 3.3 La politique de distribution

L'arrivée d'Internet a profondément bouleversé les manières pour les entreprises de distribuer leurs biens et leurs services, quitte à devenir l'unique moyen de distribution pour certaines d'entre elles (Cdiscount, Ecotour, etc.). Cette mutation du commerce a eu plusieurs effets sur les organisations et leur mode de fonctionnement : la dématérialisation de la monnaie (paiement en ligne avec carte bancaire, comptes Paypal, *bitcoins*, etc.) qui facilite et accélère les transactions,

---

<sup>44</sup> Airbnb. *Quel prix fixer pour mon logement ?* [en ligne]. Disponible sur <http://urlz.fr/3hyn>. (Consulté le 22-03-2016).

<sup>45</sup> Pricing Solutions. *Lessons in pricing from Uber* [en ligne]. Disponible sur <http://urlz.fr/3hyj>. (Consulté le 22-03-2016).

avec une possibilité de vente 24 heures sur 24, 7 jours sur 7, quel que soit l'endroit, grâce au don d'ubiquité offert par les TIC et leur miniaturisation.

Les groupes comme Accor, Marriott, IHG disposent de plateformes de réservation en ligne qui permettent d'acheter des nuitées dans n'importe quel hôtel de la chaîne et qui figurent sur la première page des moteurs de recherche. Cet emplacement de choix sur la toile est dû à l'achat de mots-clés qui optimisent le référencement et dont de paiement se fait généralement au coût par clic sur le modèle des enchères<sup>46</sup>, pouvant atteindre des sommes colossales comparées aux moyens financiers dont disposent les petites structures. Adwords, l'outil principalement utilisé et développé par Google (dont 90 % des revenus sont générés par la publicité en ligne<sup>47</sup>), reste donc à la fois un élément incontournable pour pouvoir mettre un site en ligne mais peut aussi s'avérer onéreux si l'on veut figurer en bonne place sur le moteur de recherche.

### 3.4 La communication

« Comment communiquer avec eux ? Comment instaurer une relation durable et de confiance ? », sont autant de questions que se posent aujourd'hui les professionnels. Le changement profond des codes de communication entraîné par l'avènement des réseaux sociaux a grandement bouleversé notre manière d'échanger ; il est par conséquent crucial de savoir comment s'adresser à sa clientèle de manière efficace.

La multiplication des supports de connexion et l'évolution des mœurs constituent autant de possibilités de contact avec un public multi-facettes, connecté en permanence mais dont le parcours d'achat est de plus en plus complexe et non séquentiel. Ils sont aussi constamment sollicités par les marques à travers les réseaux sociaux, les emails et la publicité en ligne. Selon une enquête de KR Médias<sup>48</sup>, nous serions exposés à un nombre entre 1 200 et 2 200 de publicités quotidiennement tous supports confondus. Cette saturation impose une réflexion visant à rendre sa communication efficace et intelligente sans pour autant étouffer le client.

Sans pour autant révolutionner les techniques de communication classiques, l'exploitation de données permet de s'adresser à son public de manière intelligente, et surtout en temps réel. En amont, il est possible de cibler un public particulier, trié en fonction de ses aspirations, ses groupes d'appartenance, ses habitudes. En aval, l'impact des campagnes peut être mesuré sans

---

<sup>46</sup> Google Support. *Fonctionnement des mots-clés* [en ligne]. Disponible sur <http://urlz.fr/3imE>. (Consulté le 23-03-2016).

<sup>47</sup> Le Monde Économie. *Pour Google, la publicité sur mobile est plus que jamais stratégique* [en ligne]. Disponible sur <http://urlz.fr/3imR>. (Consulté le 24-03-2016).

<sup>48</sup> Cité par Les Échos. *« Big Ads » ou le déluge publicitaire...* [en ligne]. Disponible sur <http://urlz.fr/3hd3>. (Consulté le 14-03-2016).

difficultés et réajusté selon les besoins. Ces techniques sont massivement utilisées par les grands groupes comme en atteste Monsieur X : « *Le moyen c'est d'avoir des groupes de contrôle. Tu vas avoir une petite partie de ta cible que tu vas garder et que tu [ne] vas pas cibler avec ce message, tu [ne] vas pas leur parler de cette promotion et ça va nous permettre de voir les différences de comportement entre ceux à qui on n'en parle pas et ceux à qui on en parle. [...] Ça te permet d'adapter beaucoup plus de choses, pas seulement modifier ton message client, mais également ce qui est créatif.* ».



Nous avons pu voir au cours de ce chapitre, le déroulement d'un plan marketing ainsi que les principales composantes sur lesquelles l'analyse de données peut avoir un rôle important à jouer, en particulier au niveau des options stratégiques fondamentales et du mix marketing. Bien qu'elles aient été traitées séparément, elles sont toutes en lien les unes avec les autres et les différents outils passés en revue sont généralement applicables à différentes actions.

## Chapitre 3. Enjeux et limites

L'ensemble des composantes évoquées précédemment s'inscrit dans une optique de pérennisation des entreprises. Bien qu'elles soient exploitées depuis plusieurs décennies par le marketing classique, les nouvelles technologies ont indéniablement permis de les affiner en y incorporant un ensemble de données jusqu'alors peu pris en compte. Nous verrons à présent quels sont les enjeux et les limites liés à leur incorporation dans les organisations.

### 1. Les enjeux

#### 1.1 L'analyse des données comme facteur d'innovation

« L'innovation est un processus qui permet à une nouveauté de devenir une pratique courante et induit une démarche collective, intégrant des contraintes économiques et de temps. Elle remet également en cause en interne ou en externe des positions acquises, des prérogatives et des intérêts. De ce fait, elle est traversée d'oppositions, de débats contradictoires et génère des confrontations »<sup>49</sup>. Comme vu précédemment, l'intégration de la composante *Big data* comme élément central du processus décisionnel reste, à l'heure actuelle, l'apanage des plus grandes organisations. Néanmoins, avoir un grand nombre d'information vis-à-vis de ses clients et du contexte dans lequel ils évoluent reste un atout majeur pour les entreprises disposant ainsi de moyens pour dégager un avantage compétitif de taille.

L'hôtellerie étant un secteur ultra concurrentiel évoluant dans un marché mature, l'objectif est d'arriver à proposer des services différents et au plus proche des attentes de chaque client. Car le cœur d'activité de ce métier de services est de délivrer une expérience unique au client, depuis son intention de séjourner dans un endroit, jusqu'à quitter l'hôtel et bien après.

L'intégration des NTIC dans ce domaine s'est faite de manière progressive avec, par exemple, le remplacement des clés par des cartes magnétiques, puis l'équipement des serrures avec des dispositifs RFID permettant l'ouverture de la porte sans contact et dernièrement la biométrie (reconnaissance de l'iris ou des empreintes digitales). Ces technologies se développent également dans les solutions de paiement visant à remplacer les cartes bancaires (puces intégrées dans des objets comme les montres, factures envoyées directement sur le téléphone, etc.)<sup>50</sup>. La chambre aussi est un espace d'étude des comportements des clients avec des objets connectés comme la

---

<sup>49</sup> ZANCANARO Frédéric. *Organisation des systèmes de restauration*. Cours de Master 1 MHR, ISTHIA, Université Toulouse – Jean Jaurès, 2016.

<sup>50</sup> Coach Omnium. Livret de l'innovation et des tendances en hôtellerie, novembre 2014, 41 p. [en ligne]. Disponible sur <http://urlz.fr/3iZl>. (Consulté de 11-03-2016).

télévision, le minibar, la climatisation, etc. qui permettent d'optimiser l'expérience client en personnalisant l'offre au maximum (température de la chambre, recommandation de chaînes télévisées ou de films à la demande, boissons favorites dans le minibar, etc.).

Toutefois, l'instauration de ces outils reste anecdotique comparée au nombre d'hôtels disséminés dans le monde, en raison notamment de leur coût élevé, mais également de la méfiance des utilisateurs, de plus en plus conscients que ces informations sont collectées en vue d'une exploitation ultérieure<sup>51</sup>.

## 1.2 Se départir de l'influence des OTA

La prédominance des OTA comme Booking, ou encore Expedia ont grandement contribué à la diversification des canaux de distribution en concentrant, en 2014, près de 45 % des réservations du parc hôtelier français. Car même si la visibilité des hôtels (60 % d'entre eux ont conclu un accord avec des OTA) et leur taux de remplissage s'en sont trouvés considérablement améliorés, le montant des commissions imputées peut parfois atteindre des sommes importantes, particulièrement chez les petits hôteliers indépendants, diminuant donc la rentabilité du service vendu<sup>52</sup>.

Véritable intermédiaire entre client et hôtelier au milieu des années 2000, ces sites ont été, à leurs débuts, un formidable moyen de promotion pour les petites entreprises qui, pour la plupart, ne disposaient alors d'aucune page Internet. Forts d'un tel engouement pour leur service, ces agences de voyage en ligne se sont imposées comme un distributeur à part entière pour devenir des acteurs incontournables du secteur touristique mondial.

En réponse à leur montée en puissance et à leurs commissions parfois exorbitantes, les opérateurs hôteliers les plus importants ont créé leurs propres plateformes de réservation, à grand renfort d'achat de mots-clés assurant un référencement en première page. Seules les enseignes célèbres ont pu accéder à cette indépendance, contrairement aux petits établissements dont la rentabilité économique ne permet pas un tel investissement et pour qui ce canal de distribution reste vital. Malgré la verticalité de cette relation et son coût, les hébergements les plus humbles bénéficient d'une interface bien plus ergonomique et exhaustive que leur propre site, à défaut de pouvoir construire des alternatives au moins aussi efficaces.

---

<sup>51</sup> *ibid.*

<sup>52</sup> DELL'ORO Jean-Louis. Pourquoi les hôteliers en veulent tant à Booking et Expedia. *Challenges*, 28 mai 2014 [en ligne]. Disponible sur <http://urlz.fr/3iA3>. (Consulté le 24-03-2016).

Les protestations de l'ensemble de la profession ont mené les OTA à assouplir les règles imposées depuis plus d'une décennie. Fin 2015, Expedia a concédé aux hôteliers l'accès à certaines des données essentielles (comme l'adresse email) laissées par les clients ayant séjourné dans leur établissement, ce qui n'était jusqu'alors pas le cas<sup>53</sup>. Ce partage des informations devrait permettre aux hôteliers qui le souhaitent d'enrichir leurs bases de données et optimiser l'usage qui peut en être fait. Car si le pouvoir des agences de voyage en ligne tend à diminuer progressivement, celui des *start-up* se fait de plus en plus imposant, constituant une nouvelle menace pour le secteur tout entier.

### 1.3 Le client avant tout

Les efforts déployés dans une stratégie *Big Data* ne doivent pas se faire au détriment du service en lui-même. Dans une entreprise du secteur tertiaire, l'expérience client est un processus constitué de cinq étapes : l'accueil, l'attente, le contact transactionnel dynamique, l'offre de service et la sortie du processus (TOCQUER et LANGLOIS, 1992, p. 124). L'ensemble des innovations matérielles citées précédemment peuvent s'avérer précieux pour observer le comportement des individus au sein de l'hôtel (heure de sortie/entrée dans la chambre, montant des consommations dans les bars, les restaurants, etc.) pour pouvoir leur proposer une offre réellement adaptée et améliorer la communication en ne lui proposant que des produits susceptibles de l'intéresser.

Les factures sont également un excellent indicateur des préférences des hôtes, comme nous affirme Monsieur X qui en a lui-même fait l'expérience au sein de la chaîne pour laquelle il travaille : après avoir beaucoup voyagé, puis passé le nouvel an dans l'un des établissements du groupe avec une consommation importante en F&B, il a reçu lors de la même période de l'année suivante une promotion lui proposant de gagner 15 000 points sur sa carte de fidélité si il consommait plus de 400 euros en F&B lors de son prochain séjour. Ce système mis en place il y a plus de trois ans connaît un franc succès et a permis de considérablement augmenter les revenus générés et le taux d'*up-selling*.

Dans un registre différent, une collaboration récente entre Hilton et IBM ont permis de créer Connie, un robot faisant office de concierge capable d'informer les clients sur les attractions touristiques et les restaurants situés aux alentours ainsi que sur les différents services proposés

---

<sup>53</sup> DENISELLE Guilain. 2016 une année comme les autres, en attendant 2017. *TendanceHotellerie.fr*, 1<sup>er</sup> février 2016, [en ligne]. Disponible sur <http://urlz.fr/3jmy>. (Consulté le 26-03-2016).

par l'hôtel<sup>54</sup>. Cet automate fonctionne selon les principes des technologies cognitives et de *machine learning*, ce qui signifie que plus la machine entre en interaction avec les visiteurs plus elle apprend à communiquer avec eux de manière pertinente pour leur fournir une expérience hors du commun.

Outre les dispositifs physiques, l'engouement pour les « concierges numériques » est tel que la plupart des groupes hôteliers s'y intéresse activement en intégrant cette fonction à leurs applications mobiles. Cette tendance répond à l'augmentation considérable de l'usage de *smartphones* et de tablettes dans le cadre de séjours hors domicile, que ce soit au moment de la recherche d'informations, de la réservation ou encore sur place (+16 % de requêtes entre le premier semestre 2014 et le premier semestre 2015<sup>55</sup>). Les groupes américains Choice Hotels et Best Western proposent ainsi à leurs clients d'accéder à du contenu informatif sur les alentours de l'hôtel, aux services de restauration, à l'adhésion aux programmes de fidélité, à un kiosque numérique comprenant journaux, musique et prévisions météorologiques<sup>56</sup>.

L'ensemble de ces dispositifs ont bien entendu pour but de fidéliser une clientèle avide de nouveauté à travers leur utilisation d'une part, mais également de pouvoir enrichir les bases de données des hôteliers pour atteindre une offre dont la personnalisation est de plus en plus poussée. Selon Bertrand Destailleur, ancien directeur marketing client du Club Med, les clients ciblés grâce aux *datas* reviennent deux fois et demi plus que ceux n'étant pas connectés, ce qui constitue des perspectives prometteuses pour l'ensemble du secteur du tourisme<sup>57</sup>.

#### 1.4 Créer des partenariats

Si les données internes restent les plus importantes dans l'exploitation d'un hôtel, elles ne fournissent que peu d'indications quant à son environnement.

C'est en ce sens que les partenariats entre entités, parfois de secteurs différents, semblent s'imposer comme une solution bénéfique à tous, comme en témoigne le directeur marketing interrogé : « *On est en train de développer un partenariat stratégique avec Amadeus, et ce sera essentiellement en Cloud. Ce sera le premier programme de réservation hôtelier à être cloudé, en tout cas à cette échelle, [...] mais pour le moment tout est encore en interne donc tout est basé dans nos bureaux. [...] On ne fait pas payer pour,*

---

<sup>54</sup> Luxury Hospitality Daily. *Hilton and IBM Pilot « Connie », the world's first Watson-enabled hotel concierge*, 10 mars 2016, [en ligne]. Disponible sur <http://urlz.fr/3j5t>. (Consulté le 24-03-2016).

<sup>55</sup> LAINÉ Linda. Les requêtes Google à la loupe. *L'Échos touristique*, juin 2015, n°3147, p. 35.

<sup>56</sup> KHLAT Mathilde. Fidélisation clients : Choice Hotels se dote d'une application d'e-conciergerie. *Tourmag*, 25 novembre 2014, [en ligne]. Disponible sur <http://urlz.fr/3j6e>. (Consulté le 21-03-2016).

<sup>57</sup> BOURGINE Jérôme. Comment fidéliser un client grâce à la *data*. *L'Écho touristique*, décembre 2014 - janvier 2015, n°3132, p. 13.

*c'est plus une question d'échanges en nature on va dire. [...] Du coup, on va faire un partenariat basé sur nos systèmes de fidélité. On donne la possibilité à leurs clients de gagner des points sur notre programme de fidélité et vice-versa. [...] Pour nous c'est intéressant parce que ça nous permet d'ajouter de nouveaux clients à notre carte de fidélité, mais également de les faire utiliser nos points encore plus et on va aussi pouvoir leur donner des placements dans nos communications, comme à nos clients.»*

Outre les GDS, des alliances peuvent être établies avec le secteur bancaire (Mastercard, Visa), afin de connaître le montant des dépenses des individus, et plus précisément des clients enregistrés, les produits les plus vendus, etc., mais également avec les opérateurs téléphoniques comme Orange, dont le programme Flux Vision (créé en 2013) offre la possibilité d'étudier les déplacements sur le territoire et les comportements des touristes (fidèles à un lieu fixe tout au long de leur séjour ou itinérant, sites les plus fréquentés, etc.) Une autre option visant à augmenter le volume des BDD est le recours à des systèmes à des Property Management System (PMS) comme par exemple Opera développé par Oracle Hospitality, qui permettent de centraliser les données provenant d'établissements différents appartenant à une même chaîne. C'est le cas d'IHG qui centralise les données de ses hôtels dans un système unique permettant aux hôteliers franchisés d'accéder à des informations provenant d'autres hôtels de la chaîne.

La création d'un vaste réseau regroupant l'ensemble des opérateurs semble être une solution efficace au premier abord. Pourtant, sa mise en place comporte de nombreuses contraintes difficiles à contourner et qui feront l'objet de la prochaine section.

## **2. Les limites**

### **2.1 Assurer la protection des données et leur bonne utilisation**

La sécurité des données constitue un des points cruciaux de la pénétration de métadonnées dans les entreprises et qui nécessitent d'être manipulées avec la plus grande attention. Comme vu précédemment, l'exploitation des données est très réglementée par les autorités européennes, aussi, pour palier toute « fuite » et poursuites éventuelles, les sièges des chaînes limitent l'accès aux informations de leurs collaborateurs, comme nous l'explique Monsieur X :

*« On a notre base de données, les commandes de toutes les réservations au niveau de IHG en tant que franchiseur, mais des franchisés vont avoir leurs propres bases de données qui suivent leur PMS de leur côté et ça ne va pas être mis en commun avec d'autres hôtels nécessairement. On ne peut pas forcément donner toutes les données client à nos franchisés parce que c'est nous qui avons pris la réservation donc le traitement des données nous incombe et leur sécurisation également. On va pouvoir donner des données essentielles à nos franchisés bien évidemment, ils*

*doivent savoir des choses sur les clients pour pouvoir les accueillir, mais pas leur donner toutes les données de contact par exemple, des adresses ou ce genre de choses parce que ça nous ferait un souci au niveau légal. ».*

Le partage de données repose donc sur la confiance mutuelle entre les acteurs ainsi que sur les contraintes légales imposées par chaque pays dans lesquels les hôtels sont implantés, rendant sa mise en place souvent compliquée. Il faut également ajouter à cela l'évolution régulière de la législation, qui impose de revoir régulièrement les systèmes utilisés et la traçabilité des opérations effectuées. En ce qui concerne l'employeur du directeur marketing, l'entreprise préfère se conformer à la législation la plus stricte (celle de l'Allemagne en l'occurrence), de façon à être protégée dans l'ensemble des pays de la zone euro, même si les lois évoluent constamment.

## 2.1 L'ensemble de ces outils ne constituent pas une fin en soi

L'adoption de solution *Big data* ne constitue pas un outil magique en elle-même. Bien que les logiciels aient été considérablement simplifiés au cours des dernières années, ils demeurent des outils difficiles à manier pour les hôteliers. D'autre part, la collecte de *datas* doit s'établir de manière pertinente. Les accumuler sans pouvoir en tirer des indicateurs décisionnels relève d'une pure perte de temps et d'argent en raison des contraintes liées au stockage et à la puissance des machines utilisées. Pour Monsieur X, le point de départ réside dans l'analyse du contenu des PMS, notamment les factures, avant d'envisager l'exploitation d'autres sources.

Il nous indique, d'autre part, que la collecte ne doit constituer une intrusion dans la vie privée au risque d'éveiller la méfiance des clients et rendre une campagne de communication inefficace. En définitive, la fidélisation par une prestation finale irréprochable importe bien plus qu'une prospection massive dont les résultats seront incertains. D'autant plus qu'un « *renouveau naturel* » s'opère par le biais des clients décidant d'eux-même de pousser la porte de l'établissement pour faire l'expérience du produit sans y avoir été incité par la marque.

Même si l'analyse des comportements *via* des données récoltées sur les réseaux sociaux, sites partenaires, site propre, s'avère utile dans l'amélioration de la satisfaction client, les méthodes de collecte traditionnelles ne doivent pas être délaissées. Les questionnaires de satisfaction établis par les hôtels constituent de précieux indicateurs, car ils mettent en valeur des points précis des variables qui pourraient être améliorées (propreté de la salle de bain, confort du lit, désagréments rencontrés lors du séjour, etc.) qui ne seront peut-être pas détaillés dans un commentaire sur Tripadvisor. L'ensemble de ces données permettront, lors d'un nouveau passage du client dans l'hôtel, de prendre en compte les défauts relevés lors de son précédent séjour et de faire en sorte de les corriger (attribution d'une chambre au calme, d'une chambre

facile d'accès pour une personne à mobilité réduite, etc.) et d'améliorer sa satisfaction, et donc sa fidélité. Il en va de même pour les études de marché, qui restent incontournables dans la définition d'un plan marketing.

Nous pouvons donc répondre à la troisième hypothèse qui stipule que la collecte et le traitement de données provenant d'Internet peut se substituer aux méthodes de collecte traditionnelles en matière de comportement des clients. Même si l'exploitation des traces numériques laissées par les internautes peut s'avérer extrêmement utile, elle constitue un apport complémentaire venant enrichir les outils marketing d'origine déjà utilisés par les entreprises hôtelières, comme les questionnaires de satisfaction, qui sont un excellent moyen de connaître l'avis des clients sur des points précis et un bon prétexte pour établir un contact avec lui.



Ce chapitre a démontré que, même si l'usage des *Big data* peut s'avérer bénéfique dans la stratégie marketing, il présente toutefois certaines limites de taille. Vecteur d'innovation, il peut contribuer à s'émanciper de l'influence des OTA en optimisant la communication et la distribution des services hôteliers. La création de partenariats est également un procédé auquel il est important de réfléchir pour pouvoir enrichir les bases de données de chacun et affiner la prise de décision, de façon à toujours garantir une qualité de prestation optimale au client. Cependant, la puissance de ces outils impose une utilisation et une maîtrise rigoureuse, sous peine de voir l'ensemble des efforts fournis réduits à néant.

## Conclusion

La seconde partie de ce mémoire a permis d'établir un lien entre plan marketing et *Big data* en étudiant tout d'abord le plan marketing dans sa globalité. Cette approche théorique semblait nécessaire afin de bien comprendre l'intérêt de l'introduction de la composante *data* à celle-ci. Nous avons pu constater que l'arrivée d'Internet a considérablement changé la donne dans les relations de pouvoir existant entre client et entreprise.

En définitive, cette nouvelle ère numérique offre de nombreuses opportunités à ceux qui savent les saisir, en l'occurrence les grandes firmes. Car si le *Big data* repose *a priori* sur une idée simple, sa mise en action requiert tout de même un minimum de compétences, tant sur le plan technique que sur l'aspect marketing dont seule une poignée d'entreprises dispose. Elles embauchent pour cela de nouveaux profils de salariés spécialisés dans les domaines de l'informatique et de l'analyse statistique, faisant apparaître de nouveaux métiers.

Ces mutations du secteur ont également pour effet de transformer le processus décisionnel dans son intégralité, ce qui risque de poser problème au secteur de l'hôtellerie, connu pour ses difficultés d'adaptation et l'abandon d'habitudes bien ancrées.

Nous avons aussi pu infirmer la troisième hypothèse, qui partait du postulat que l'exploitation des traces numériques pouvait se substituer aux méthodes de collecte traditionnelles. Certes l'apport de données externes constitue un complément d'information très utile aux entreprises, mais les outils internes comme les questionnaires de satisfaction, les études de marché ou les entretiens individuels restent incontournables car ils permettent de soulever des points précis n'apparaissant pas forcément dans l'analyse de données provenant des multiples sources d'information disponibles.

Plusieurs questions demeurent pourtant en suspens en ce qui concerne l'hôtellerie : Quel est le rapport actuel que les entreprises du secteur entretiennent avec ces nouvelles technologies ? Et de quelles manières les abordent-elles ? Les outils de collecte et d'analyse ne sont-ils réellement réservés qu'aux grands groupes ? Toutes ces interrogations constituent autant de réponses qui feront l'objet de la prochaine partie.

## **Partie III. Méthodologie et terrain d'application**

## Introduction

Cette dernière partie portera sur une approche concrète de l'utilisation des *datas* dans le secteur de l'hôtellerie. L'étude théorique effectuée en amont sera à la base des éléments de réponse fournis dans les prochaines pages et nous permettront de confirmer ou d'infirmer les hypothèses émises précédemment.

Le premier chapitre portera sur la présentation de la méthodologie suivie tout au long de l'année avec les outils d'analyse utilisés afin de bien délimiter dans quel cadre l'étude a été effectuée.

Le deuxième sera consacré à l'analyse du parc hôtelier français : sa composition, ses acteurs majeurs ainsi que les difficultés auxquelles il est exposé. L'ensemble de ces données nous permettront de mieux comprendre ce marché en pleine évolution.

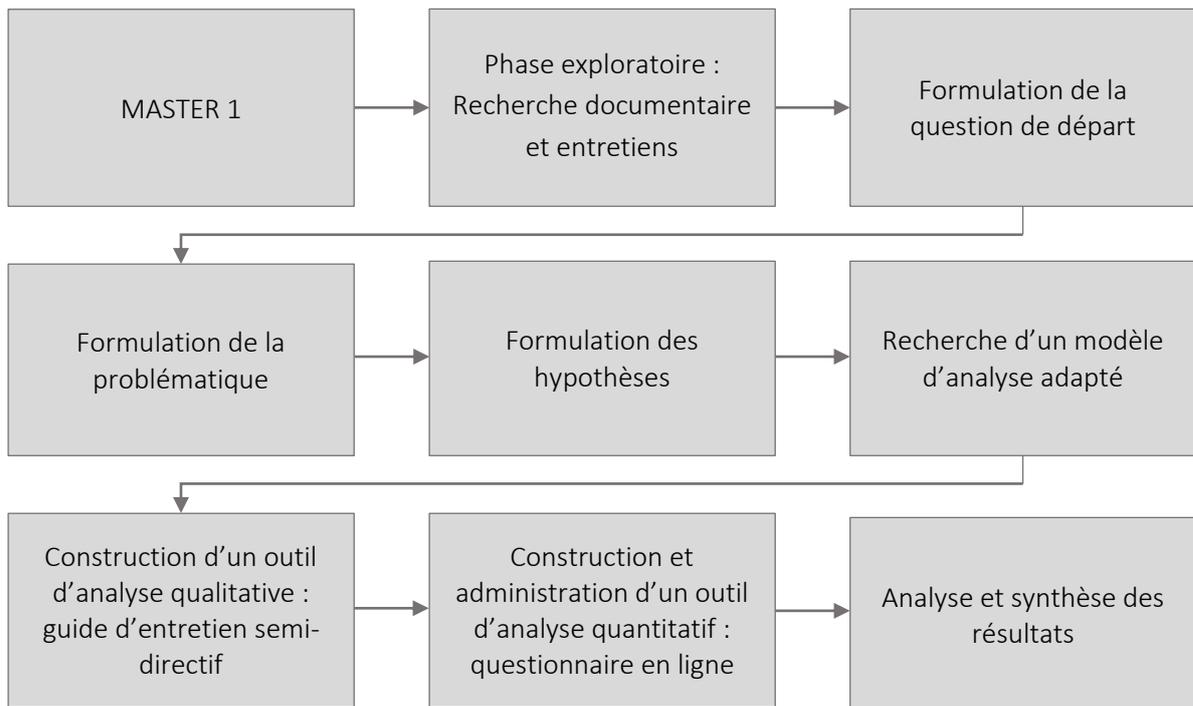
Nous étudierons ensuite le comportement des professionnels face à ce sujet et les démarches mises en œuvre jusqu'à présent. Tous les types d'hôtels seront pris en compte : indépendants, affiliés à une chaîne volontaire ou intégrée, économiques ou de gamme supérieure.

Le dernier chapitre fera état des perspectives d'avenir et de la poursuite des recherches envisagée au cours de l'année de Master 2.

# Chapitre 1. Méthodologie

La création de ce mémoire découle d'une méthodologie englobant différents outils d'analyse qui ont permis de cadrer l'ensemble du travail effectué au cours de l'année. Elle permet également de suivre le cheminement de la réflexion, allant du choix du sujet jusqu'à la finalisation de l'écrit.

## 1. Déroulement de la recherche



## 2. Rappel des objectifs

L'objectif majeur de ce mémoire est d'introduire de la manière la plus simple possible la question du *Big Data* dans les entreprises hôtelières. Préoccupation de plus en plus importante pour les entreprises comme pour les particuliers, la collecte et le traitement de données s'impose comme une question primordiale de ce début du 21<sup>ème</sup> siècle.

Le travail réalisé tout au long de l'année a permis de dégager la problématique ainsi que les hypothèses suivantes :

**Comment le *Big data* peut-il permettre aux hôteliers d'affiner leur approche marketing pour cerner au mieux leurs clients et devancer leurs attentes ?**

**Hypothèse n°1 :** L'exploitation des données n'est réservée qu'aux grands groupes qui, contrairement aux hôteliers indépendants, disposent de moyens (financiers, humains et matériels) importants.

**Hypothèse n°2 :** Les hôteliers doivent développer de plus en plus de stratégies afin de pouvoir collecter des données pertinentes utiles à la prise de décision.

**Hypothèse n°3 :** La collecte et le traitement de données émises sur Internet peut se substituer aux méthodes de collecte traditionnelles en matière de comportement des clients (questionnaires de satisfaction, entretiens consommateurs, enquêtes, etc.).

### **3. Les outils utilisés**

#### **3.1 L'étude qualitative**

##### **a) Objectif**

L'étude qualitative a pour objectif de recueillir des informations auprès d'acteurs œuvrant sur le terrain par le biais d'entretiens semi-directifs qui d'abord ont une approche concrète du sujet à travers l'analyse de la teneur des propos recueillis et du langage para-verbal des individus interrogés. Un guide d'entretien a été construit afin de n'oublier aucun point à évoquer et recentrer le répondant en cas de débordement du sujet<sup>58</sup>.

##### **b) Profil de la personne interrogée**

La personne interrogée est un directeur marketing d'un groupe hôtelier international ayant d'ores et déjà mis en place une politique *Big data* dans la stratégie globale et marketing de l'entreprise.

##### **c) Conduite de l'entretien**

L'entretien a eu lieu par téléphone durant 1h30, permettant d'obtenir les réponses les plus sincères possibles par la suite. Afin de faciliter l'analyse des réponses, l'interview a été enregistrée après avoir obtenu l'accord de l'interrogé puis retranscrite<sup>59</sup>.

L'entretien se décompose en trois phases : la phase introductive, destinée à mettre à l'aise l'interviewé et présenter l'individu, son parcours, ses activités. La phase de recentrage a pour

---

<sup>58</sup> Annexe D : Guide d'entretien

<sup>59</sup> Annexe E : Retranscription de l'entretien avec Monsieur X

vocation de faire basculer la conversation sur le cœur du sujet et récolter des réponses pertinentes pour pouvoir confirmer ou infirmer les hypothèses émises. Enfin, la phase de conclusion qui laisse la parole à l'interrogé, à ses impressions et durant laquelle il est remercié de sa participation.

La dernière étape de l'étude qualitative consiste à analyser les résultats obtenus lors des entretiens et se pratique en deux temps : tout d'abord la retranscription de chaque entretien afin de dégager les idées et les thèmes principaux évoqués, puis procéder dans un second temps à des regroupements entre les propos tenus, qu'ils soient similaires ou divergents. Dans la mesure où un seul entretien a été mené dans le cadre de cette étude, cette seconde étape se fera de façon plus sommaire.

### 3.2 Analyse quantitative

#### a) Objectif

L'analyse quantitative a quant à elle pour but d'avoir une vision globale et chiffrée de l'utilisation du *Big data* chez les hôteliers, quel que soit leur classement, leur zone géographique ou leur statut (indépendant ou appartenant à une chaîne) et ainsi pouvoir déterminer des profils présentant des similitudes.

#### b) Échantillon et mode administration

L'administration de ce questionnaire quantitatif s'est faite en ligne par le biais de l'outil Google Forms, sur la forme de l'auto-administration<sup>60</sup>. Il est majoritairement constitué de questions fermées pour faciliter le traitement des réponses, mais également de questions ouvertes et semi ouvertes pour laisser aux hôteliers l'occasion d'exprimer leur ressenti par rapport au sujet. L'échantillon est composé de l'ensemble des hôtels français classés à ce jour, allant de zéro à cinq étoiles Palace, dont les adresses email ont été collectées par le biais du site *data.gouv*, plateforme *open Data* du gouvernement français (plus de 12 000 hôtels recensés<sup>61</sup>). Le taux de réponse étant assez faible avec ce type d'enquête, il est essentiel de l'envoyer au plus grand nombre d'acteurs possible pour avoir des résultats exploitables et représentatifs d'une tendance générale.

---

<sup>60</sup> Annexe F - Résultats bruts du questionnaire en ligne

<sup>61</sup> Data.gouv. *Hôtels classés en France* [en ligne]. Disponible sur <http://urlz.fr/3hBL>. (Consulté le 7-02-2016).

Les résultats obtenus seront ensuite analysés à l'aide d'un tableur. L'objectif majeur de cette étape consiste à croiser les données entre elles de façon à en faire ressortir du sens et des corrélations. Pour ce faire, il faut trier les données, créer des tableaux croisés dynamiques et appliquer des filtres. Tout comme pour les dirigeants d'entreprises, ce processus s'est avéré délicat.

#### 4. Les limites

Plusieurs limites d'ordre matériel et technique se sont imposées lors de la conception de ce mémoire : tout d'abord les contraintes de temps, tant au niveau des recherches qu'au niveau de la planification d'entretiens qui n'ont permis de rencontrer qu'un seul professionnel. Ces derniers étant peu disponibles, il a été difficile d'obtenir des réponses de leur part. De plus, les résultats du questionnaire en ligne dont le taux de retour très faible (à peine 1,5 %) est propre à ce mode d'administration, mais également à l'envoi d'une telle quantité d'emails pour laquelle Google impose des limites<sup>62</sup>.

Au niveau de la recherche documentaire, la difficulté première a été de pouvoir aborder le *Big data* de manière compréhensible par tout un chacun, car la majorité des ouvrages et articles universitaires évoquent ce sujet de manière très scientifique, demandant des compétences pointues en statistiques. De plus, rares sont les entreprises qui exposent de manière explicite les techniques auxquelles elles ont recours pour récolter des données sur leurs clients, probablement pour des raisons de confiance et de confidentialité.



Bien que cette analyse n'ait pas la précision d'une étude scientifique, elle constitue tout de même un ensemble de pistes de réflexion autour des mutations économiques et sociétales qu'engendre l'avancement des technologies.

---

<sup>62</sup> Sur les 12 000 envoyés, seule la moitié a atteint les destinataires. Cela est aussi dû au fait que beaucoup d'adresses n'ont pas été actualisées par l'administrateur

## Chapitre 2. Le parc hôtelier français

Avant de passer à l'analyse de l'enquête et de l'entretien exploratoires, il convient tout d'abord d'établir un panorama complet de l'hôtellerie française à ce jour.

### 1. Panorama du parc hôtelier français

Véritable référence en matière de tourisme mondial, l'hôtellerie française a subi de nombreux changements au cours de la seconde moitié de cette décennie. En cause, une conjoncture économique et sociale incertaine dont de nombreux acteurs ont souffert, suite à la crise financière de 2008 et aux attentats perpétrés dans plusieurs destinations phares du globe lors des deux précédentes années. D'autre part, la réforme du classement des établissements de tourisme mis en place depuis 2012 a également fortement contribué à l'évolution de la composition du parc hôtelier français comprenant à ce jour plus de 17 000 hôtels.

Tableau 2 - Parc hôtelier français depuis 2010

Gamme	2010		2015		Variation depuis 2010
<b>1 étoile</b>	3 072	18 %	636	5 %	- 2 436
<b>2 étoiles</b>	9 127	54 %	5 031	39 %	- 4 096
<b>3 étoiles</b>	3 876	23 %	5 376	42 %	+ 1 500
<b>4 &amp; 5 étoiles</b>	898	5 %	1 755	14 %	+ 857
<b>Total des hôtels classés</b>	16 973	100 %	<b>12 798</b>	100 %	- 4 175
<b>Part des hôtels classés</b>		90 %		75 %	
<b>Hôtels non classés</b>	1 911	10 %	4 345	25 %	+ 2 434
<b>Total des hôtels</b>	18 884	100 %	<b>17 143</b>	100 %	- 1 741

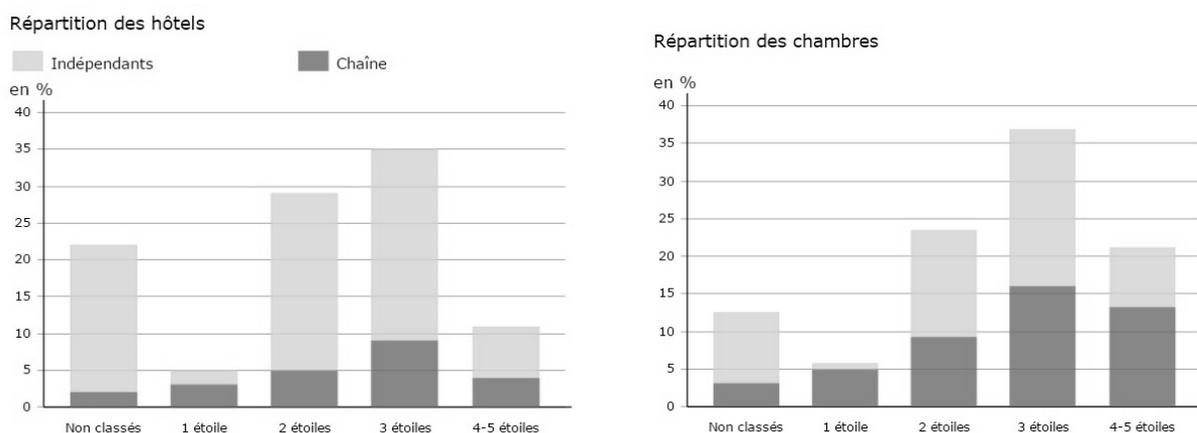
Source : Coach Omnium, Panorama de l'hôtellerie française, 2015

La disparition de l'ancien système de notation a engendré un recul considérable du nombre d'étoiles comme critère de choix d'un établissement (14 % en 2015 contre 64 % en 2008<sup>63</sup>). Le prix reste bien entendu le premier élément déterminant et tend à devenir une donnée bien plus fiable pour se faire une idée du standing d'un hôtel.

De plus, le marché de l'hôtellerie française est considéré comme mature par les acteurs qui le composent avec un chiffre d'affaires global en progression de 2.1 % an depuis 2009<sup>64</sup> mais qui tend à stagner dans les années à venir. Afin de pallier cette baisse potentielle, il est important que le secteur tout entier entreprenne des démarches pour redynamiser leurs ventes et redorer une profession dont l'image se fait de plus en plus vieillissante.

La répartition entre indépendants et hôtels de chaîne est très hétérogène, quelle que soit la catégorie, comme le montrent les graphiques ci-dessous :

Figure F - Répartition des hôtels et des chambres selon le type d'hôtel et leur classement en 2014



Source : Insee<sup>65</sup>.

Nous pouvons observer que la répartition se fait de manière très inégale en termes de nombre de chambres ou en termes d'établissements. Les indépendants (chaînes volontaires comprises) dominent largement le marché en matière de structures mais ce sont les chaînes intégrées qui l'emportent au niveau du volume de chambres. Cette différence s'explique par le fait que ces derniers disposent d'établissements comportant trois fois plus de chambres que dans ceux des indépendants.

<sup>63</sup> Coach Omnium. *Panorama de l'hôtellerie en France*, 2015, p. 4-5 [en ligne]. Disponible sur <http://urlz.fr/3hUQ>. (Consulté le 18-03-2016).

<sup>64</sup> Étude Xerfi citée par Ernst & Young. *L'hôtellerie en France : Les grands défis du secteur vus par ses principaux acteurs*, octobre 2014 [en ligne]. Disponible sur <http://urlz.fr/3haD>. (Consulté le 21-03-2015).

<sup>65</sup> Insee. *En 2014, une chambre d'hôtel sur deux appartient à une chaîne* [en ligne]. Disponible sur <http://urlz.fr/3hTA>. (Consulté le 20-03-2016).

## 2. L'hôtellerie indépendante

Cette forme d'hôtellerie reste très largement majoritaire sur le territoire (83 % des hôtels<sup>66</sup>). Ce sont généralement des établissements familiaux de type économique, jouissant d'une capacité restreinte de 26 chambres en moyenne. Le problème principal auquel leurs dirigeants sont confrontés reste la faible rentabilité de leurs affaires qui ne leur permet pas d'investir et de les maintenir adaptées à une demande qui varie constamment. Selon le Comité pour la Modernisation de l'Hôtellerie et du Tourisme Français<sup>67</sup>, près d'un quart des hôtels de France sont considérés comme vieillissants, voire vétustes. Plusieurs causes, outre les influences négatives externes à la profession, sont à l'origine de cette crise<sup>68</sup> :

- Des taux d'occupation faibles dont la plupart ne dépasse pas les 50 % ;
- Une méconnaissance des attentes des clients ;
- Une adaptation difficile à Internet ;
- Des prix en baisse en raison de la pression exercée par les OTA (Online Travel Agency ou agences de voyages en ligne) ;
- Une méconnaissance notable des bonnes pratiques commerciales et de marketing hôtelier ;
- Des outils informatiques de gestion peu exploités ;
- Un conservatisme indéniable des habitudes en matière de gestion des établissements.

L'ensemble de ces raisons font que beaucoup de dirigeants se tournent désormais vers l'adhésion à des chaînes volontaires (pour 30 % d'entre eux<sup>69</sup>) ou intégrées, quand ils ne souhaitent tout simplement pas vendre leurs affaires.

## 3. L'hôtellerie de chaîne

### 3.1 Les chaînes intégrées

Malgré un contexte instable, ce sont ces hôtels qui résistent le mieux face aux difficultés. Bien que représentant seulement 18 % du parc hôtelier français, ils comptabilisent à eux-seuls 47 %

---

<sup>66</sup> Étude Xerfi citée par Ernst & Young. *L'hôtellerie en France : Les grands défis du secteur vus par ses principaux acteurs*, octobre 2014 [en ligne]. Disponible sur <http://urlz.fr/3haD>. (Consulté le 21-03-2015).

<sup>67</sup> Comité pour la modernisation de l'hôtellerie et du tourisme français. *La petite hôtellerie en difficulté*, 2013, p. 4-5 [en ligne]. Disponible sur <http://urlz.fr/3kTq>. (Consulté le 19-03-2016).

<sup>68</sup> *Ibid.* p.27-28.

<sup>69</sup> Insee. *En 2014, une chambre d'hôtel sur deux appartient à une chaîne* [en ligne]. Disponible sur <http://urlz.fr/3hTA>. (Consulté le 23-03-2016).

des nuitées, soit prêt d'une chambre sur deux<sup>70</sup>. Ils bénéficient en effet de capacités bien supérieures à celle des hôteliers indépendants et disposent d'outils de gestion et de communication performants (centrales de réservation, applications mobiles, programmes de fidélité, etc.). Quelques opérateurs, sous 60 enseignes différentes, ont la main mise sur ce marché, dont Accor en tête sur le territoire français (47 % des hôtels de chaîne), suivi de Louvre Hotel (racheté par le groupe Jin Jang en 2016).

Tableau 3 - Classement des 10 chaînes intégrées présentes en France (en nombre d'hôtels)

Enseigne	Hôtels en France 2016	Chambres en France 2016	Hôtels monde, dont France
<b>Ibis</b>	383	33 690	1 031
<b>Campanile</b>	324	20 164	397
<b>Ibis Budget</b>	324	25 074	537
<b>Première classe</b>	246	17 641	252
<b>Kyriad + Kyriad Prestige</b>	242	14 506	242
<b>Hotel F1</b>	238	17 906	246
<b>B &amp; B hotels</b>	232	17 324	322
<b>Mercure Hotels</b>	226	22 080	775
<b>Ibis Style</b>	141	9 920	333
<b>Novotel</b>	113	15 412	414
<b>Balladins</b>	106	5 598	107

Source : Coach Omnium

Les deux groupes français se partagent la grande majorité du classement : Accor, en tête des groupes les plus représentés sur le territoire au travers de cinq enseignes (Ibis, Ibis Budget, Ibis

<sup>70</sup>Coach Omnium. Étude chaînes hôtelières intégrées 2016 [en ligne]. Disponible sur <http://urlz.fr/3hS5>. (Consulté le 20-03-2016)

Style, Mercure, et F1) et dont le nombre ne cesse d'augmenter année après année ainsi que Louvre Hotels avec ses marques Première Classe, Campanile, Kyriad et Kyriad Prestige.

Il est également à noter que 79 % de l'offre totale, toutes chaînes confondues, porte sur les hôtels économiques, 16 % sur les hôtels milieu de gamme et 5 % sur les haut de gamme.

### 3.2 Les chaînes volontaires

Avec 23 enseignes recensées par Coach Omnium en 2016<sup>71</sup>, l'hôtellerie de chaîne volontaire représente près de 28 % des établissements classés et non classés en France.

Tableau 4 - Classement des chaînes volontaires en France

Enseigne	Hôtels en 2016	Chambres en 2016	Hôtels dans le monde (hors France)	Capacité moyenne par hôtel
Logis	2 094	41 113	150	20
Châteaux & Hôtels Collection	351	9 188	47	26
Hotusa Hoteles	302	17 819	2 336	59
Best Western	299	15 799	3 632	53
Contact Hotel	270	7 360	0	27
Inter-Hotel	269	11 939	1	44
Citotel	251	7 596	31	30
Relais et Châteaux	132	3 420	331	26
Relais du silence	113	2 497	60	22
Hotels & Preference	113	6 948	19	61

<sup>71</sup> Coach Omnium. *Panorama113 2016 sur les chaînes hôtelières volontaires*, [en ligne]. Disponible sur <http://urlz.fr/3hRY>. (Consulté le 20-03-2016).104

Bien qu'occupant une place non négligeable sur la scène du tourisme, une baisse significative du nombre d'adhérents a été constatée au cours des dernières années (cinq points entre 2014 et 2016). La raison principale pour laquelle les hôteliers s'inscrivent dans cette démarche est l'accès à des outils de communication et de marketing qu'ils ne pourraient pas mettre en place par eux-mêmes. Cette forme d'hôtellerie souffre néanmoins elle aussi du poids des OTA dont la visibilité écrase celle des enseignes volontaires entraînant de plus en plus d'hôteliers à se détacher de ces dernières.



Ce chapitre a planté les bases de la composition du parc hôtelier français ainsi que son évolution au cours des cinq dernières années. Nous nous attacherons à présent à vérifier si ces tendances sont confirmées sur le terrain et correspondent aux résultats du questionnaire quantitatif.

## Chapitre 3. Analyse des actions des acteurs

Après avoir posé les bases qui permettent de mieux appréhender les spécificités du parc hôtelier français, nous allons à présent nous attacher aux comportements des hôteliers et au niveau de l'intégration des *Big datas* au sein de leur entreprise.

### 1. Profil des répondants

110 individus ont répondu au questionnaire administré par email. L'ensemble des catégories sont représentées mais ne sont pas toujours en adéquation avec les proportions observées dans la réalité. Les participants sont répartis de la façon suivante :

Tableau 5 - Répartition des répondants en fonction de leur catégorie et de leur statut

	1*		2*		3*		4*		5* & Palaces		Total	
Indépendants	3	50 %	18	55 %	16	41 %	13	44 %	0	0	51	47 %
Volontaires	2	33 %	13	39 %	14	36 %	9	32 %	2	66 %	40	36 %
Intégrés	1	17 %	2	6 %	9	23 %	6	21 %	1	33 %	19	17 %
Total	6	5 %	33	30 %	39	35 %	29	26 %	3	3 %	110	100 %

- La majorité des répondants (65 %) appartient à des catégories situées entre deux et trois étoiles ;
- Près de la moitié 47 % sont des établissements indépendants, 36 % adhérents à une (ou plusieurs) chaîne volontaire et 19 % à une chaîne intégrée ;
- 80 % des établissements interrogés existent depuis plus de 15 ans.
- Toutes les régions sont représentées mais la région Auvergne-Rhône-Alpes comptabilise à elle seule 23 % des réponses.

Ces résultats permettent de mieux définir les caractéristiques de l'échantillon ainsi que le contexte dans lequel les hôteliers évoluent.

### 2. Constat face à l'utilisation des données

Tous les acteurs du tourisme ne sont pas égaux face à la récolte, au traitement, au stockage et à l'exploitation des données. Cette affirmation était d'ailleurs le point de départ de la première hypothèse affirmant que l'intégration du *Big data* n'était réservée qu'aux grands groupes et aux

établissements haut de gamme qui, contrairement aux hôteliers indépendants, disposent de moyens (financiers, humains et matériels) importants.

## 2.1 Niveau de connaissance du sujet

Le questionnaire a permis de mettre en avant le fait que, pour la moitié d'entre eux, ce sujet reste encore inconnu. À la question « Avez-vous déjà entendu parler du *Big data* ? », 51 % des participants ont répondu de manière négative. Pour les autres, 37 % n'ont pas la moindre idée de ce que ce terme signifie (« Je ne sais pas », « aucune idée ») et nous pouvons remarquer que les définitions qui en sont données sont très évasives voire énigmatiques.

La plupart des participants est consciente que le sujet est lié à l'informatique et aux données mais peu arrivent à réellement donner une explication précise. Les mots-clés les plus cités sont « base de données », « collecte » et « analyse », ce qui prouve qu'ils connaissent plus ou moins les procédés de traitement de l'information même s'ils ne les utilisent pas au quotidien dans l'exploitation de leur hôtel. La notion de volume est également citée à plusieurs reprises (« recoupage de plusieurs informations sur un sujet », « grand ensemble d'informations », « stockage de gros volumes de données », « ensemble des infos stockées sur le web », etc.) sans pour autant être l'élément de réponse principal.

## 2.2 Utilisation des données au sein de l'entreprise

Après leur avoir exposé la définition de manière complète, incluant plusieurs exemples de données disponibles (historiques de réservations, géolocalisation, fichiers clients, réseaux sociaux, etc.), la question suivante visait à savoir si les interrogés utilisaient certaines d'entre-elles. Les réponses obtenues ont été très tranchées :

- La totalité des hôtels une et deux étoiles ainsi que quelques hôtels trois étoiles ont répondu « non » (42 % des répondants). Le manque de compétences est de loin la raison la plus évoquée (64 %), suivie par l'absence d'intérêt pour le sujet (25 %) et une dépense trop lourde (11 %).
- L'ensemble des hôtels allant de trois à cinq étoiles a répondu de manière affirmative et 48 % sont indépendants.

Nous pouvons remarquer que le rapport entre indépendants et affiliés à une chaîne volontaire ou intégrée est assez équilibré. Ces données permettent de répondre à la première hypothèse affirmant que seuls les établissements de chaîne exploitent les données. Les chiffres ci-dessus

prouvent que les hôtels de gamme moyenne et supérieure sont bel et bien les seuls à réellement afficher à réel intérêt pour les traces numériques. L'intégration des *datas* dans le processus de décision ne repose donc pas sur l'affiliation à une ou plusieurs chaînes, mais sur le standing de l'hébergement. Nous pouvons supposer que cet écart tient au fait que plus l'on monte en gamme, plus il est possible de créer des postes spécialisés dans le traitement de données. En ce qui concerne l'affiliation à un type de chaîne, il est difficile de déterminer son influence sur les pratiques car les différences de proportions sont moins nettes.

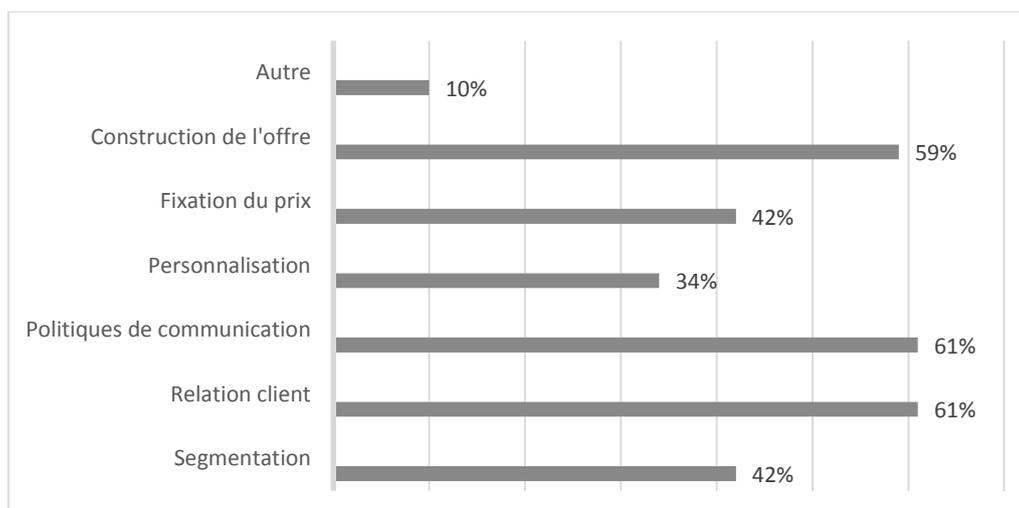
### 2.3 Outils utilisés et données recueillies

Un fait est cependant notable : à peine 26 % des une et deux étoiles affirment disposer d'un logiciel CRM, ce qui montre des lacunes en matière d'outils de gestion tant pour les indépendants que pour ceux appartenant à une chaîne. Selon Monsieur X, *« la plus grosse barrière à l'entrée des hôteliers indépendants c'est vraiment le savoir, le savoir-faire. Plus que le coût en soi. Dans le sens où [l'on pourrait] considérer que le savoir-faire est un coût, effectivement, ça pousse à engager des gens qui savent le faire, mais pour moi c'est vraiment un problème de savoir-faire parce qu'il y a plein de plateformes qui permettent de faire plein plein de choses en ligne et peu chères. Et puis [il n'y a pas] pas à tous les utiliser en tant qu'hôtelier indépendant. [S'il arrive] à avoir un programme d'emailing assez customisable et qui marche assez bien c'est déjà un gros atout et très peu d'hôtels indépendants qui le font correctement. »*.

Pour les autres catégories, les données majoritairement récoltées sont issues de l'analyse du trafic du site de l'entreprise (38 %) notamment *via* Google Analytics, qui permet de connaître la nationalité des internautes, le support utilisé (*smartphone*, tablette, ordinateur), les avis postés sur les réseaux sociaux (17 %). Les coordonnées et données relatives à l'âge, au motif de voyage sont quant à elles enregistrées lors de la réservation ou de l'adhésion à un programme de fidélité.

Concernant les outils de traitement utilisés, certains sont plus répandus que d'autres. Les tableurs et les logiciels CRM arrivent tête (47 % et 42 %), mais beaucoup les combinent avec des logiciels de traitement spécifiques (37 %) ou font appel à une société tierce spécialisée dans le traitement des données (27 %).

Figure G - Usage des données récoltées par les hôteliers

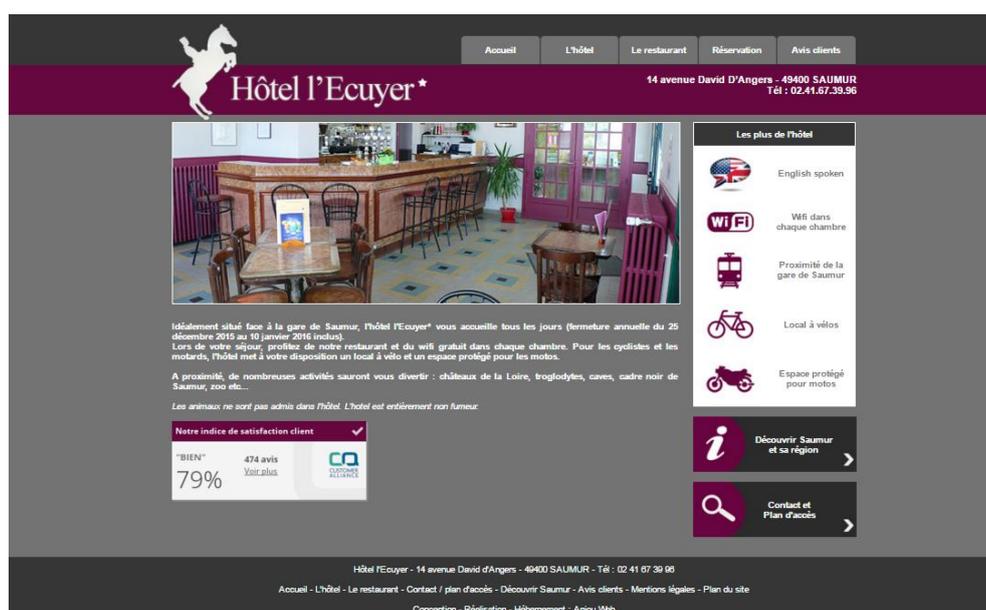


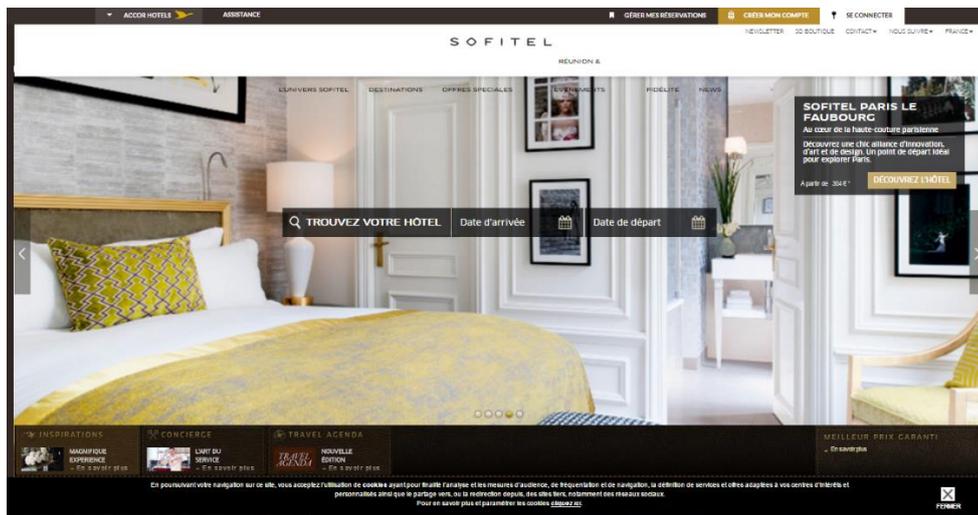
Trois usages principaux se détachent lorsqu'il s'agit de l'utilisation des données récoltées : la relation client, les politiques de communication et la construction de l'offre. La personnalisation ne vient, quant à elle, qu'en avant-dernière position, derrière la fixation du prix et la segmentation. Ces chiffres vont quelque peu à l'encontre de ceux ressortis de la revue de littérature qui cite ces trois derniers éléments comme étant les axes centraux de l'exploitation des *Big data*.

### a) Les sites Internet

Contrairement à ce que nous aurions pu penser, la totalité des répondants dispose d'un site Internet, de l'hôtel indépendant une étoile au cinq étoiles de chaîne. Nous pouvons toutefois remarquer quelques différences de taille en comparant la page de deux établissements :

Figure H - Sites d'un hôtelier 1\* indépendant et 5\* de chaîne intégrée





Il existe tout d'abord un écart notable entre les deux chartes graphiques. La première est simple et composée de trois couleurs, avec de photographies des parties de l'établissement (lobby, restaurant, chambre). La seconde page est plus agréable à l'œil avec des images de qualité supérieure et un *design* épuré.

Outre l'aspect esthétique, nous constatons que deux éléments diffèrent d'un site à l'autre : alors que la page d'accueil de l'enseigne Sofitel dispose d'un calendrier permettant de vérifier les disponibilités d'un hôtel avec la possibilité de réserver directement en ligne, celle de l'hôtel Ecuyer affiche une brève description de l'établissement avec un onglet réservation donnant uniquement accès à une messagerie en ligne avec laquelle il est possible d'envoyer un email pour retenir une chambre. D'autre part, la présence de cookies sur le site de la marque d'Accor visant à enregistrer le passage et le parcours d'un internaute pour améliorer la qualité d'une navigation ultérieure sur ce même site en personnalisant par exemple le bandeau avec le nom de la personne ou en ajustant le contenu. La plupart des sites commerciaux ont aujourd'hui adopté ce procédé pour recueillir des informations sur leurs clients et enrichir l'expérience faite de la marque, en amont de la prestation. Globalement, le but est de garder le visiteur le plus longtemps sur le domaine, de façon à avoir le plus de données possible sur son parcours (onglets cliqués, temps passé par sections, produits les plus consultés, dates de séjour envisagées, etc.) pour pouvoir adapter l'offre en aval (services, tarification, recommandations).

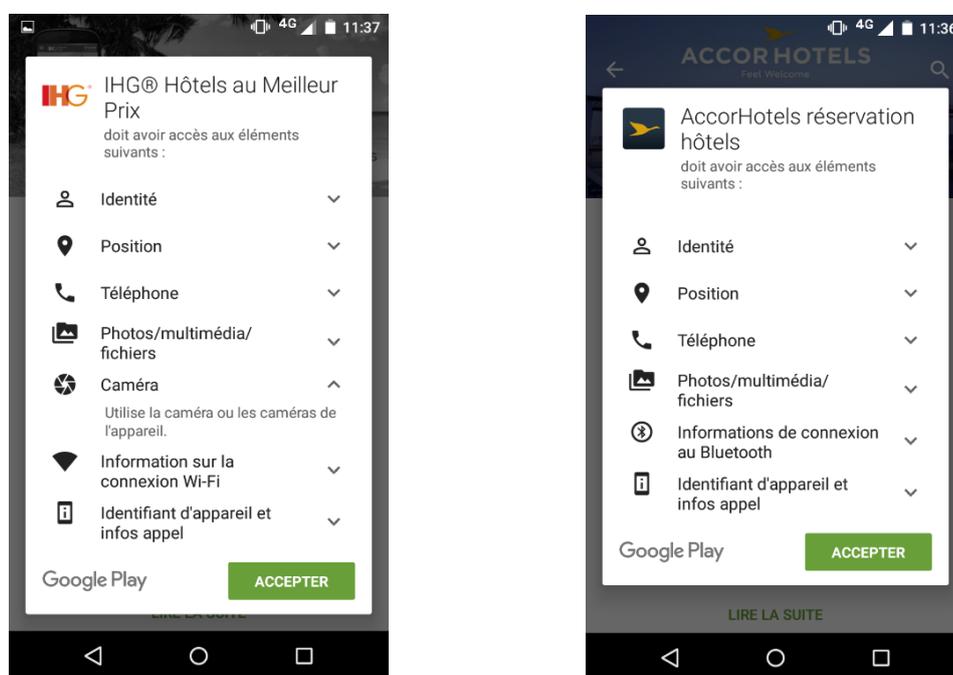
Parallèlement, 73 % des répondants affirment analyser le trafic et le contenu de leur site web. Encore une fois, ce sont des établissements une et deux étoiles qui ne procèdent à aucun traitement. La différence de proportion avec les questions précédentes montre bien que, même s'ils ne l'identifient pas toujours comme tel, les hôteliers manipulent les *datas* avec plus ou moins d'attention et de conscience. Cette idée est confirmée par le directeur marketing interrogé qui

affirme que « beaucoup d'entre eux mènent un politique *Big data* sans forcément savoir qu'il y a un nom pour ça ».

## b) Les applications mobiles

Les applications mobiles ont beau être moins répandues, elles n'en demeurent pas moins très présentes avec 40 % des établissements signalant en posséder une (quatre et cinq étoiles principalement.). Près de 70 % d'entre eux appartiennent à une chaîne volontaire ou intégrée qui incluent souvent cet outil au « *package* » destinés à leurs collaborateurs. L'analyse de l'utilisation des applications repose sur les mêmes principes que celle des sites web, à la différence près qu'elles accèdent également au contenu du téléphone de l'utilisateur comme en témoigne la capture d'écran ci-dessous, réalisée lors de leur téléchargement.

Figure I - Éléments requis lors de l'installation d'une application



Ces deux images montrent que de nombreuses informations personnelles, comme les photos, l'identité ou encore les appels, sont ainsi récoltées en plus de celles laissées lors de l'utilisation des applications. Nous pouvons légitimement nous demander si l'ensemble de ces éléments ne constituent pas une intrusion dans la vie privée, bien que le mobinaute ait été averti des contreparties liées à leur téléchargement.

Outre leur utilité dans le cadre d'une recherche d'hôtel, les applications constituent également un support d'information pour le client lors de son séjour, comme nous l'affirme Monsieur X :

« [...] ils arrivent sur le réseau wifi de l'hôtel et donc on sait [qu'ils sont] dans cet hôtel et on peut [leur] donner par exemple les menus du restaurant pour ce soir, ou des choses un peu utiles. Ça nous permet de changer le contenu de l'application, du coup on peut passer d'une application qui permet de réserver à une application qui permet par exemple de gérer [leur] séjour dans l'hôtel. ».

### c) Les réseaux sociaux

Devenue incontournable, la présence sur les réseaux sociaux est un minimum à effectuer pour les hôteliers pour pouvoir être visibles sur la toile. Le gros avantage réside dans le fait que contrairement aux sites web propres aux hôteliers, l'interface reste identique et gratuite dans le cadre d'un usage basique. Certaines fonctionnalités peuvent être ajoutées, comme la possibilité de réserver une chambre dans le cas des hôtels Accor. De plus, ces plateformes revêtent un aspect dynamique que n'ont pas les autres supports, par l'ajout d'articles, de publications ou de commentaires clients, mais également par le ciblage des personnes à qui la page va être suggérée.

#### 2.4 Accès à d'autres données complémentaires

Nous souhaitons savoir si en dehors des données auxquelles ils ont déjà accès, les hôteliers aimeraient obtenir d'autres compléments d'informations internes ou externes. Quasiment la moitié (46 %) nous a indiqué déjà bénéficier de toutes les données nécessaires au pilotage de leur établissement. Pour les autres, nous retrouvons des éléments relatifs à la concurrence, aux coordonnées des clients et aux critères sociodémographiques, mais aussi les *datas* détenues par les OTA (Booking principalement). Ce dernier point nous montre que, malgré un assouplissement des règles entre agences de voyage en ligne et hôteliers, la rétention d'information des intermédiaires est toujours de mise.

## 3. Le stockage des données

Comme évoqué dans la première partie de ce mémoire, l'exploitation des *datas* dépend directement des supports de stockage de celles-ci, il semblait donc intéressant de savoir à quels matériels les dirigeants ont recours dans leurs hôtels. Les structures personnelles sont privilégiées à 68 % et le *cloud computing* à 11 %, mais 21 % répartissent leurs informations entre ces deux formes de stockage.



Plusieurs conclusions sont à tirer de ce chapitre. Tout d'abord, l'écart notable de l'utilisation des données entre les différentes catégories d'hôtels : les établissements économiques (une et deux

étoiles) sont largement moins enclins à analyser les différentes informations à leur disposition que ceux appartenant aux catégories supérieures. Cette différence pourrait être expliquée par la taille des hôtels indépendants (26 chambres en moyenne) qui ne permettent qu'une hausse limitée de leur remplissage. Le manque de connaissance est aussi très présent, même si plusieurs des répondants analysent plusieurs variables sans pour autant en prendre conscience. Ces difficultés d'adaptation des décideurs pourraient résider dans le fait que cette évolution, contrairement à l'avènement de l'informatique dans les années 90, ne requiert pas uniquement apprivoisement d'un nouveau support physique, mais l'apprentissage d'un savoir technique et théorique souvent difficile à aborder. Toutefois, concernant le développement des techniques destinées à récolter des informations, nous pouvons constater que des outils simples comme des logiciels *open source* ou des tableurs sont suffisants pour traiter les données. Les réseaux sociaux et la page web de l'établissement constituent également d'excellentes sources externes. Le recours à des partenariats avec des entités tierces reste encore une pratique marginale et réservée aux grands groupes. Toutefois, la plupart des outils de bases sont aisément accessibles à tout un chacun et permettent de récolter suffisamment d'informations (trop pour certains hôteliers qui ne savent pas comment les gérer). Nous pouvons donc en partie infirmer les trois hypothèses mises en avant lors du début de cette partie.

## Chapitre 4. Perspectives futures

Suite aux travaux effectués au cours de la première année de Master, plusieurs perspectives de poursuite sont envisageables lors du mémoire de Master 2.

### 1. Possibilités de stage en Master 2

La compréhension des concepts relatifs au *Big data*, a permis d'acquérir des connaissances théoriques potentiellement bénéfiques pour le secteur de l'hébergement. En ce sens, il serait judicieux d'envisager un stage dans un cabinet de conseil aux entreprises du secteur de l'hôtellerie et de la restauration. La maîtrise de l'exploitation des données étant une compétence de plus en plus recherchée, plusieurs opportunités pourraient se présenter.

Le but ultime serait, à terme, de pouvoir créer un guide destiné aux hôteliers indépendants et surtout économiques, afin qu'ils puissent à leur tour bénéficier des avantages liés à l'exploitation des données

### 2. Travail de recherche envisagé

Pour atteindre cet objectif, différentes pistes et éléments de recherche pourraient être étudiés comme la conduite d'entretiens semi-directifs avec des professionnels de plusieurs natures :

- D'autres responsables marketing de grandes chaînes hôtelières comme ACCOR, qui possèdent d'ores et déjà des services consacrés au *Big Data* ;
- Des hôteliers indépendants afin d'établir un parallèle entre leurs stratégies et celle des hôtels de chaîne concernant l'utilisation des *datas* ;
- Des « grossistes » de données tel qu'Etalab (administrateur de data.gouv) pour savoir quels sont les organismes partenaires, fournisseurs et exploitants, des jeux de données mis à disposition, ainsi que les procédés de récolte et de traitement ;
- Des entreprises spécialisées dans le traitement et la commercialisation de données (sans pour autant être dédiée au secteur HR) afin de comprendre leurs procédés, les données sélectionnées, leur mise en relation (Tourinflux, Orange Flux, SoToulouse, Portail Bretagne Open *data*, Amadeus *data*) ;
- Des clients, pour connaître leur avis sur l'usage commercial qui est fait des données les concernant.

L'apprentissage et la maîtrise d'outils comme Google Analytics est également prévu pour pouvoir avoir une approche réellement professionnelle de ces concepts.

## Conclusion

La dernière partie de ce mémoire nous a donné la possibilité d'avoir une vision concrète de l'élaboration de ce travail d'une part, mais également des comportements observables sur le terrain d'autre part.

La démarche respectée ainsi que les outils d'analyse qualitative et quantitative ont été détaillés pour justifier de la pertinence des résultats exposés par la suite. Les limites, d'ordres technique, temporel et financier sont aussi exposées.

Les résultats obtenus par l'administration des questionnaires quantitatifs nous ont donné la possibilité de constater quelles étaient les comportements des hôteliers face aux nouvelles évolutions technologiques engendrées par la massification des données disponibles et le développement des moyens d'analyse. Il en ressort globalement que l'appréhension de ce vaste sujet reste très hétérogène suivant la catégorie des hôtels.

Les perspectives de poursuites de cette étude en Master 2 ont aussi été évoquées afin d'approfondir et d'affiner les investigations menées lors de cette année.

## Conclusion générale

L'ensemble du travail effectué tout au long de ce mémoire nous a permis d'apporter plusieurs éléments de réponses en ce qui concerne l'intégration des *Big data* dans le secteur de l'hôtellerie. Si les recherches réalisées en amont ont démontré un aspect technique indéniable du sujet, nous avons également pu élaborer un questionnaire s'intéressant à l'utilisation qui peut en être faite dans le cadre du marketing hôtelier.

Afin de répondre à cette problématique, plusieurs hypothèses ont été émises sur l'aptitude des hôteliers à mettre en place une telle démarche. Toutes sont basées sur la revue de littérature et l'entretien exploratoire réalisés qui ont permis de délimiter un cadrage pour la suite de ce mémoire. La méthodologie et le terrain d'application retenus pour la troisième partie ont permis de confronter les affirmations directement auprès des acteurs concernés, à savoir le parc hôtelier français.

Même si de nombreux éléments de réponse sont apparus, il est encore difficile de pouvoir pleinement affirmer qu'elles répondent de manière précise et définitive aux hypothèses et demandent à être davantage approfondies lors de l'année de Master 2.

Nous avons tout d'abord pu constater que l'exploitation des données n'est pas uniquement réservée aux hôtels appartenant à des chaînes. De nombreux hôteliers indépendants ont d'ores et déjà entamé une démarche tournée vers les *datas*, en utilisant des outils simples comme des tableurs ou des logiciels CRM, ce qui montre bien que les dispositifs ne sont pas forcément aussi complexe que l'on pourrait le croire. L'écart le plus flagrant se situe davantage au niveau du classement des hôtels. De manière globale, plus l'établissement est haut de gamme, plus il sera enclin à collecter, traiter, stocker et exploiter les informations. Ceci pourrait être expliqué par le fait que la plupart des hôtels indépendants ont un nombre de chambre bien inférieur à ceux affiliés à une chaîne, mais cette variable mériterait d'être examinée plus en profondeur ultérieurement. Le manque de connaissance du sujet est aussi un point crucial, particulièrement pour les une et deux étoiles.

Ensuite, les moyens mis en œuvre par les organisations sont, tout comme pour le traitement des données, constitués de dispositifs simples et peu onéreux comme le référencement sur les réseaux sociaux ou la création de programmes de fidélité. Bien sûr, les chaînes hôtelières disposent de techniques plus complexes comme la création d'applications mobiles et de « packages » de gestion dont bénéficient leurs collaborateurs, qui leur permettent, de mener des

politiques plus complètes et aisément vérifiables. Le souci majeur dans ce cas est de savoir trier les informations pour ne garder que celle présentant un véritable intérêt pour l'entreprise, ce qui semble encore difficile pour une majorité d'entre eux. Nous avons néanmoins pu remarquer que les données les plus recherchées étaient les coordonnées et les données sociodémographiques comme l'âge, le sexe ou la nationalité.

Enfin, concernant le fait que les traces numériques peuvent se substituer aux méthodes de collecte classiques, il a été montré que les deux sont indissociables et complémentaires. En d'autres termes, les informations tirées de questionnaires de satisfaction ou d'entretiens menés auprès des consommateurs par exemple, constituent l'élément de base auquel viennent s'agréger celles provenant de traces numériques.

L'ensemble de ces éléments de réponse ont donné l'opportunité d'établir un constat des pratiques des hôteliers du territoire. Le succès d'une politique marketing centrée sur les *datas* repose essentiellement sur une combinaison subtile d'objectifs prédéfinis, d'outils basiques mais accessibles et d'un minimum de connaissance du sujet.

Pour conclure, nous pouvons nous demander si le phénomène *Big data* relève d'un simple effet de mode qui s'essouffera dans un avenir plus ou moins proche, ou bel et bien d'un nouveau virage technologique incontournable pour les entreprises. Présenté aujourd'hui comme un élément de domination de la concurrence, sa diffusion à l'ensemble des acteurs économiques pourrait le transformer en norme, lui faisant perdre ce côté magique que beaucoup lui confèrent.

# Bibliographie

## Ouvrages

BRASSEUR Christophe. *Enjeux et usages du Big Data : technologies, méthodes et mise en œuvre*. Paris : Éditions Hermes Science Publications, 2013, 214 p.

COINTOT Jean-Charles, EYCHENNE Yves. *La révolution Bid Data : les données au cœur de la transformation de l'entreprise*. Paris : Éditions Dunod, 2014, 240 p.

DE BAYNAST Arnaud, LENDREVIE Jacques. *Publicitor 8<sup>ème</sup> édition*. Paris : Éditions Dunod, 2014, 570 p.

DELORT Pierre. *Le Big Data*. Paris : Éditions Presses Universitaires de France, collection Que-sais-je ?, 2015, 128 p.

GORUNESCU Florin. *Data mining: Concepts, Models, and Techniques*. New York: Éditions Springer, collection Intelligent systems Reference Library (volume 12), 2011, 360 p.

KOTLER Philip, KELLER Kevin. *Marketing Management 13<sup>ème</sup> édition*. Paris : Éditions Pearson, collection Pearson Education, 2009, 960 p.

LENDREVIE Jacques, LÉVY Julien, LINDON Denis. *Mercator 9<sup>e</sup> édition : Théories et nouvelles pratiques du marketing*. Paris : Éditions Dunod, 2009, 1230 p.

LENDREVIE Jacques, LÉVY Julien. *Mercator 11<sup>e</sup> édition - Tout le marketing à l'ère numérique*. Paris : Éditions Dunod, 2014, 1040 p.

MAYER-SCHONBERGER Viktor, CUKIER Kenneth. *Big Data: A Revolution that will transform how we live, work and think*. Londres : Éditions John Murray, 2013, 256 p.

MAYOL Samuel. *Le Marketing 3.0*. Paris : Éditions Dunod, 2011, 200 p.

MUSSO Pierre. Délégation interministérielle à l'aménagement et à la compétitivité des territoires (DIACT). *Territoires et cyberspaces en 2030*, Paris : La Documentation française, 2008, 148 p.

REIX Robert, FALLERY Bernard, KALIKA Michel. *Systèmes d'information et management des organisations*. Paris : Éditions Vuibert, 2011, 480 p.

SCHMARZO Bill. *Big data : Tirer parti des données massives ou développer l'entreprise*. Paris : Éditions First Interactive, Paris, 2014, 214 p.

SCOLE Robert, SHEL Israel. *L'ère du contexte : Ces nouvelles technologies qui bouleversent notre environnement*. Paris : Éditions Diatino, 2014, 286 p.

TEBOUL Bruno, BOUCHER Jean-Marie. *L'absolu Marketing: Web 3.0, Big Data, neuromarketing...*, Paris : Éditions KAWA, collection Tout savoir sur..., 2013, 298 p.

TUFFERY Stéphane. *Data mining et statistique décisionnelle. L'intelligence des données*. Paris : Éditions tecnip, 2012, 826 p.

TOCQUER Gérard, LANGLOIS Michel. *Marketing des services : Le défi relationnel*. Paris : Éditions Dunod, 1992, 188 p.

### **Articles académiques**

VAYRE Jean-Sébastien. *Comment coproduisons-nous notre environnement numérique marchand ? Les objets numériques création et consommation*, 2014, Ax-les-Thermes, p. 1-21. [en ligne]. Disponible sur <https://hal.archives-ouvertes.fr/hal-01056152> (Consulté le 23-12-2015).

### **Rapports officiels**

Legifrance. *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés*, 2013, 76 p. [en ligne]. Disponible sur <https://lc.cx/4mrB>. (Consulté le 21-11-2015).

### **Mémoires :**

JOUANNY Jean-Simon. *Big data et marketing : Les opportunités pour l'entreprise*. Mémoire de Master 1 Web Éditorial, Poitier : Université de Poitier, 2014, 99 p.

## **Annexes**

# Table des annexes

Annexe A : Tableau de conversion des unités de mesure de mémoire informatique...	87
Annexe B : Classification des informations individuelles.....	88
Annexe C : Programme Watson développé par IBM .....	89
Annexe D : Guide d'entretien .....	90
Annexe E : Retranscription de l'entretien avec Monsieur X.....	91
Annexe F - Résultats bruts du questionnaire en ligne .....	104

## Annexe A : Tableau de conversion des unités de mesure de mémoire informatique

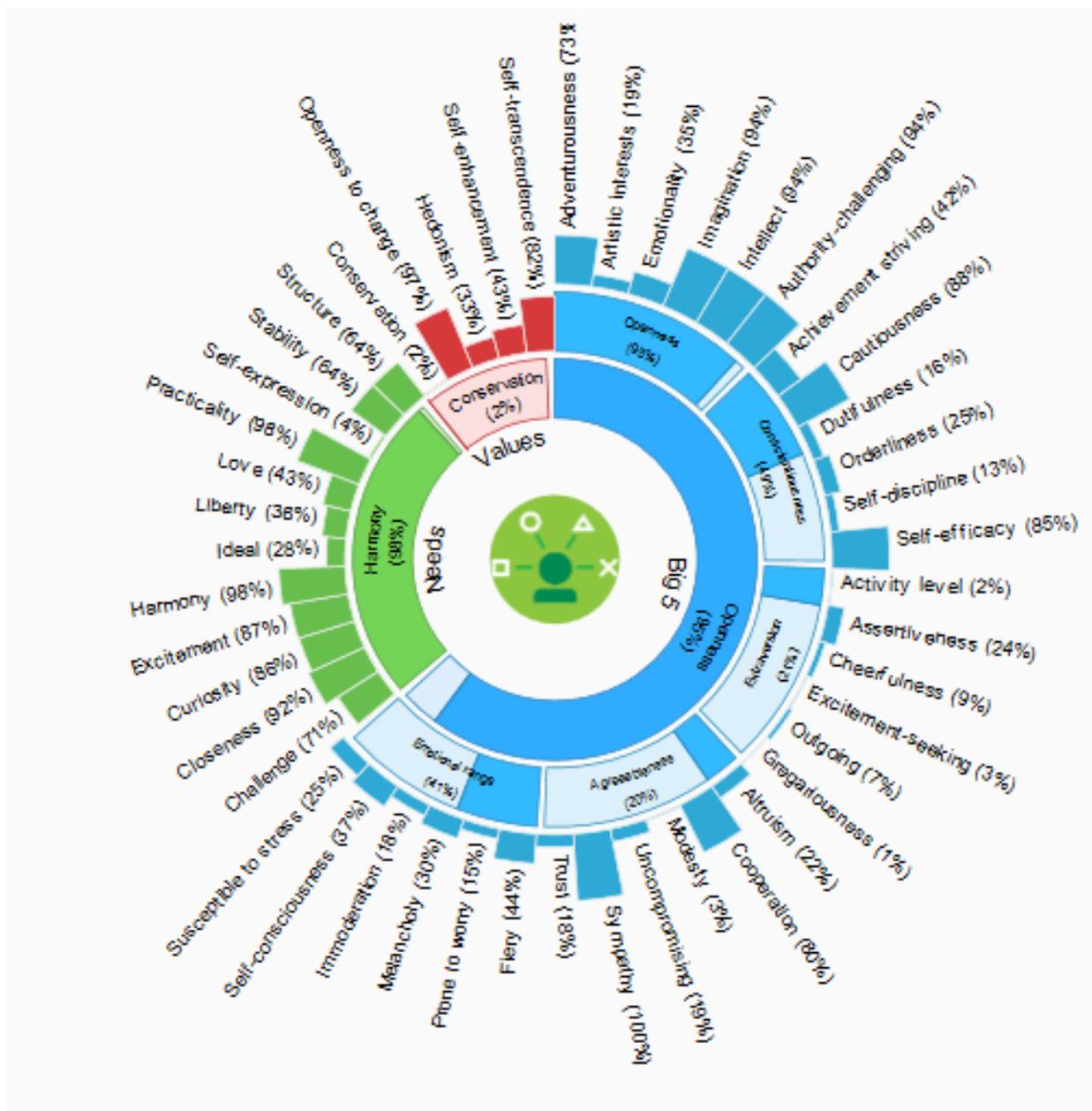
1 octet	8 bit
1 Kilo-octet (Ko)	1 000 octets
1 Méga-octet (Mo)	1 000 000 octets
1 Giga-octet (Go)	1 000 000 000 octets
1 Tera-octet (To)	1 000 000 000 000 octets
1 Pétaoctet (Po)	$10^{15}$
1 Exaoctet (Eo)	$10^{18}$
1 Zettaoctet	$10^{21}$
1 Yottaoctet	$10^{24}$

## Annexe B : Classification des informations individuelles

<b>LES DONNÉES OBJECTIVES</b>	
<b>Données d'identification et de contact</b>	Identité civile, adresse postale, numéro de téléphone, numéro d'immatriculation d'un véhicule, numéro du permis de conduire, adresse électronique, références bancaires, numéro de carte bancaire, numéro de Sécurité sociale, pseudonyme sur Internet...
<b>Données physiques et de signalement</b>	Taille, poids, couleur et écart entre les yeux, état de santé, maladies contagieuses, vaccins, empreintes digitales, photographie, voix...
<b>Données sociodémographiques</b>	Date de naissance, âge, genre, statut marital, nombre d'enfants, niveau d'études, profession...
<b>Données juridiques</b>	Capacité juridique, casier judiciaire, amendes...
<b>Données financières et foncières</b>	Solde du compte bancaire, crédits et dettes, fréquence des impayés, interdit bancaire, allocations sociales, pensions alimentaires versées/reçues, propriété foncière, hypothèques...
<b>DONNEES SUBJECTIVES</b>	
<b>Préférences et centres d'intérêt</b>	Orientation sexuelle, préférences alimentaires, vestimentaires et culturelles, destinations touristiques préférées...
<b>Opinions et activités politiques, religieuses et syndicales</b>	Opinions politiques émises publiquement, adhésion à un syndicat, croyances intimes déclarées ou manifestées publiquement, signature d'une pétition...
<b>Données comportementales</b>	Apparence, élocution et style de langage, types d'achats et d'établissements fréquentés, pratique et fréquence des loisirs, consommation d'électricité ou d'eau, utilisation d'un service public, données de connexion et navigation sur Internet, consommation de drogues (alcool, tabac, stupéfiants)...
<b>Données géographiques</b>	Lieux de retrait d'argent et de paiement par carte bancaire (péages...), bornes d'enregistrement par puces sans contact, indentation Bluetooth...
<b>Données relationnelles</b>	Taille du réseau social, fréquence des rencontres, nombre d'amis, de partenaires sexuels, participation à un club ou à une association, appels téléphoniques, emails, SMS envoyés et reçus...

Source : JAUREGUIBERRY Francis & PROULX Serges. *Usages et enjeux des technologies de communication*, Paris : Éditions Erès, collection Erès Poche-société 2011, 144 p.

Annexe C : Programme Watson développé par IBM



## Annexe D : Guide d'entretien

Monsieur X, Responsable marketing Europe d'un groupe hôtelier international

### Phase introductive

Pouvez-vous me décrire votre parcours et me parler de votre poste ?

### Phase de recentrage

1. Utilisez-vous les *Big Data* dans votre entreprise ? Au niveau de quel(s) service(s) ?
2. Depuis quand ?
3. Quelles sont les données que vous récoltez ?
4. Comment procédez-vous (site web, cookies, achat de jeux de données, applications mobiles, géolocalisation, *open Data*, réseaux sociaux, transactions bancaires...) ?
5. Dans quel but les utilisez-vous ?
6. Avez-vous un service dédié dans votre entreprise ou confiez-vous cette tâche à un organisme tiers ? Pourquoi ?
7. Quels sont les outils auxquels vous avez recours ?
8. Quelles sont les données auxquelles vous souhaiteriez avoir accès en plus de celles que vous avez déjà ?
9. Avez-vous rencontré des difficultés lors de la mise en place du *Big Data* ? De quel ordre ?
10. Avez-vous remarqué une amélioration de vos performances suite à l'intégration des *Big Data* dans votre stratégie globale et marketing ?
11. En termes de marketing, sur quels points les données permettent-elles d'améliorer vos services (relation client, construction de l'offre, tarification, personnalisation du service...) ? De quelle manière ?
12. Les *Big Data* vous aident-elles dans votre prise de décision ?
13. **Informez-vous vos clients de la collecte de leurs données ? Pourquoi ?**
14. Quel est votre avis concernant l'aspect éthique que revêt le recours à l'utilisation de données ?

### Perspectives générales et limites

1. Quelles sont les perspectives d'avenir ? Vos projets ?
2. Pensez-vous que l'exploitation des *Big Data* est aujourd'hui incontournable pour les hôteliers ?
3. Quels sont selon vous les facteurs de succès de l'intégration des *Big Data* dans les hôtels ?
4. Quelles sont selon vous, les limites de l'utilisation des données ?
5. Pensez-vous que cette pratique n'est accessible qu'aux hôteliers disposant d'une assise financière importante ? Pourquoi ?

### Conclusion de l'entretien

1. Il y a-t-il des choses sur lesquelles vous voudriez revenir ?
2. Remerciements

## Annexe E : Retranscription de l'entretien avec Monsieur X

Profil : Directeur marketing Europe d'un groupe hôtelier

Effectué le 18/02/2016

Durée : 1h35

### Phase d'introduction

**Étudiant : Pouvez-vous vous présenter ? Votre métier, votre parcours, ce que vous faites tous les jours.**

Professionnel : Bien, pas de soucis, on commence par ça. Donc du coup j'ai commencé chez IHG il y a 4 ans et demi maintenant, quelque chose comme ça. J'ai commencé en, tout ce qui était *social marketing*, donc tout ce qui était lié aux média sociaux et à l'époque c'était tout ce qui était, donc heu... comment transposer une marque. Moi mon focus était sur les marques de luxe, donc Intercontinental et Crown Plaza pour nous, et donc j'ai commencé à... travailler sur la transposition de ces marques en question sur l'univers des média sociaux, donc savoir en gros, comment est-ce-que tu arrives à transposer les grandes marques, les standards d'une marque depuis ce qu'on sait au niveau du produit, c'est-à-dire, voilà les chambres d'hôtel ressemblent à ça, le bien ressemble à ça etc., sur les caractères d'un *tweet* par exemple ou sur Facebook, comment est-ce-que tu arrives à transposer cette expérience client auxquels nos clients sont bien évidemment en droit d'attendre sur des réseaux qui ont leurs propres codes, leurs propres façons de travailler. Donc ça a été mon premier, mon premier... le premier objectif pour mon... , donc j'ai pas mal bougé dans cette équipe, mais en gros pendant deux ans c'est ce que j'ai fait et après je me suis heu... réorienté vers ce qu'on appelle chez nous le marketing commercial qui est tout ce qui est, comment dire... marketing ayant pour but de générer des revenus directs c'est-à-dire qu'à la différence du marketing de marque par exemple, on va plus avoir tendance à jouer sur la notoriété, la préférence donc sur comment générer cette notoriété spontanée, notoriété aidée et les préférences de marques, donc je vais plutôt m'attarder sur tout ce qui va générer des revenus directement, c'est-à-dire des offres spécifiques, du marketing de contenu etc... donc tout ce qui est en-dessous, un petit peu plus bas et donc c'est de ça que je m'occupe heu... à l'heure actuelle, partout dans à peu près une centaine d'hôtels sur les 650 que nous avons et donc c'est sur la France, Genève, pour moi puis tous les hôtels et *resorts* pour notre marque de luxe.

### Phase de recentrage

**Du coup vous recueillez toutes les impressions des clients qui vous servent comme vous l'avez dit, à faire des offres personnalisées. Depuis quand IHG a-t-il commencé à véritablement investir dans cette politique de recueil de données afin d'améliorer le service, l'expérience client à travers la personnalisation ?**

J'ai un peu envie de dire depuis le tout début. En fait ce qui est difficile à savoir, c'est de savoir à partir de quand tu estimes qu'une personnalisation commence. Si tu veux on a des hôtels qui ont plus de 150 ans d'histoire et qui... si tu veux... ont un fonctionnement très *old school* de l'hôtellerie. Jusqu'en 1950-52, le Grand Hôtel à Paris a ouvert par exemple, la manière très *old school* de faire de la personnalisation, des offres personnalisées, c'était à l'époque où tu étais ton bureau derrière ta réception et d'apprendre la liste de tous les VIP qui étaient à l'époque dessinés parce qu'il n'y avait pas de photos donc heu... il y avait donc les portrait de tous les VIP et tout nouveau réceptionniste se devait de les apprendre par cœur pour pouvoir leur heu... comment dire... les reconnaître, leur dire bonjour de manière personnalisée, savoir quelles étaient leurs préférences, c'est-à-dire avoir un fichier client à chaque fois en disant « donc voilà, le prince de Machin son *drink* préféré c'est l'absinthe, Monsieur Detruc, je sais pas, sa chambre préférées c'est la 363 etc... ». Donc tout ça c'était déjà de la datalisation avant l'heure, d'une certaine manière. Tu allais par exemple envoyer des choses par courrier qui allaient être bien évidemment nominatives donc il y avait pas mal de choses comme ça, donc si on veut quelque part, de nos jours la seule différence c'est que c'est beaucoup plus facile de le faire à très grande échelle en fait. C'est ce qui a beaucoup évolué sur les 10-20 dernières années avec le traitement automatique des données etc... Après quant à savoir quels sont les dates liées à des requêtes

spécifiques ça dépend du moment où tu commences.... On a eu notre... On a fait... On a été la première entreprise hôtelière à passer à un système informatique de réservation qui s'appelle Holidex qui est notre petit bébé qui est toujours là par ailleurs que nous changeons l'année prochaine. Mais ça fait plus de 40 ans qu'il fonctionne et c'était la première fois, c'était révolutionnaire à l'époque que des réservations étaient traitées de manière informatisée. C'est-à-dire qu'avant ça tu fonctionnais encore avec ce qu'on appelait LEQUEDOC qui étaient en fait d'énormes recueils, peut-être je sais pas, de format A1 ou quelque chose comme ça avec plein plein de techniques différentes pour le faire donc c'était ça qui était intéressant, donc tu faisais pas la même chose en Suisse que ce que tu faisais en France, que tu faisais en Allemagne, que tu faisais aux États-Unis par exemple, mais effectivement c'était un énorme cardex avec plein de techniques différentes pour réussir à retrouver, lier les réservation entre elles etc... donc ce que Holidex a rendu possible c'est quelque chose de beaucoup plus simple, mais à l'époque ça a commencé avec des choses comme Amadeus, des choses comme ça et donc il y a eu pas mal de choses de ce côté-là et effectivement, ce que ça a permis après c'est pour retrouver très facilement des réservations en recherchant dans la base de données etc, déjà ça a permis d'avoir une base de données un peu plus facilement accessible et après potentiellement de pouvoir dire, de pouvoir faire des choses très funs, ce que tu pouvais pas faire du tout avant, ou de manière très limitée, avec un cardex ou alors de manière très compliquée. C'est inventer une nouvelle façon d'interroger ta base de données, c'est-à-dire que par exemple, tu pouvais dire « j'aimerais toutes les entrées dont l'anniversaire tombe un 5 décembre par exemple et hop, d'arriver à ressortir toutes les personnes dont l'anniversaire tombait un 5 décembre et imaginons que tu étais la dernière semaine de novembre, donc voilà, tu pouvais leur envoyer un petit courrier en disant « Joyeux anniversaire, voici une petite offre pour vous pour célébrer cette occasion » dont quelque part on peut considérer que ça c'était l'une des premières façon de faire, de notre côté dans tous les cas. Mais bon après ça a accéléré, puisqu'à l'époque ça prenait très très longtemps pour interroger ces données. Pour te donner un ordre d'idée des premières façons d'interroger des bases de données, avec Holidex c'était avec des cartes perforées tu vois. Donc il fallait les coder, donc il fallait des développeurs pour t'aider à coder les cartes et les perforateurs qui allaient pouvoir transmettre les données à l'ordinateur qui allait pouvoir te rendre une liste imprimée sur telex etc. Enfin tu vois le genre, c'était très *oldschool* comme truc. (Silence) Ce qui est intéressant, c'est que de nos jours on parle beaucoup de *Big Data* et de trucs comme ça, mais au final ce sont pas vraiment des choses nouvelles. Utiliser des données clients, ça a toujours existé, c'est quelque chose qui a toujours été là. C'est juste que ça a été rendu beaucoup plus facile de le faire pour beaucoup plus de gens en même temps. C'est-à-dire qu'à l'époque tu le faisais beaucoup pour tes VIP, tu le faisais beaucoup pour ton top 10, top 20, top 30, parce que ça prenait du temps, parce qu'il fallait le faire de manière très manuelle etc... Quand les premiers ordinateurs ont commencés, ça a permis d'accélérer une partie de ce traitement de données et petit à petit c'est devenu de plus en plus rapide. Par exemple chez nous, donc heu..., c'est HolidayInn international qui a lancé Holidex et si tu voulais y avoir accès, vu que c'était dans les années 60, il fallait que tu envoies ta carte perforée parce que la seule manière d'interroger ta base de données c'était d'être physiquement dans la même pièce. Donc du coup voilà. Après, le seul boom de l'autre côté ça a été dans les années 90 où tout a commencé à être relié par Internet et là ça a commencé à être vraiment différent parce que tu pouvais avoir petit à petit et à distance, accès à distance à des données, donc ça prenait encore du temps mais tu pouvais le faire depuis chez toi, enfin entre guillemets, et après vraiment dans les années 2000 où c'était l'accès instantané aux données et aux bases de données, ce qui permettait de faire des interrogations directes et instantanées. Maintenant si je veux savoir justement quels sont les gens qui sont nés le 5 décembre dans notre base de données, la différence c'est que je peux le savoir dans à peu près 20 minutes et avoir la liste dans un fichier Excel que je peux utiliser immédiatement. C'est très différent de à l'époque où il fallait sûrement disons, un mois pour avoir le résultat.

### **Quelles sont les données qui vous intéressent ? Sur vos clients, vos clients potentiels ?**

Alors c'est... (rires) c'est plein de choses ! C'est toujours intéressant parce que c'est le gros, du moins une des grosses questions de l'utilisation de la donnée de nos jours. C'est à dire qu'est-ce que tu récoltes, qu'est-ce qui est intéressant. C'est même plus profond que ça. L'approche, la première approche il y a une dizaine d'années, c'était vraiment de se dire « on récolte tout et on verra plus tard si on en a besoin, mais au pire comme ça on l'a déjà ». Ça a marché pendant un certain temps mais le problème c'est qu'il y a tellement de points de données que tu peux récupérer sur quelqu'un. Le problème c'est que non seulement ça prend énormément de place à stocker, beaucoup de place, beaucoup de puissance, d'ordinateurs à interroger évidemment et ça, ce sont des choses à prendre en ligne de compte parce que potentiellement quand t'es une boîte comme IHG où, comment dire, dans laquelle on a beaucoup, beaucoup de clients à prendre en ligne de compte, si tu comptes le nombre de comptes fidélité et le

*Reward Club* donc on a 92,3 millions de membres. Si je prends que ces gens-là et les données qu'on a sur les membres de notre programme de fidélité c'est déjà un nombre de points de données. De mémoire, on est à plus de 240 points de données différents par personne ça ce ne sont que nos membres, c'est qu'une toute petite partie de tous nos clients qui viennent chez nous et je ne te parle même pas des clients potentiels. Du coup, pour répondre à ta question, quand t'as plus de 200 pour nos membres qui sont les personnes pour lesquelles on va essayer de récolter le plus de choses, ce sont les personnes pour lesquelles on a les données de contact les plus précises. Ce sont les plus faciles pour nous, ce sont les personnes avec qui il est plus facile de communiquer. De fait, celles qui pourraient nous intéresser le plus, on va avoir envie de pouvoir motiver notre message de manière à le rendre le plus efficace possible. Par exemple ça va être de voir que l'anniversaire d'un client potentiel ça va pas être très intéressant, parce qu'il si je lui envoie un message, alors qu'il est jamais venu chez nous et que je lui dis « Votre anniversaire arrive, venez faire l'expérience notre nouvelles offre... ». C'est plus inquiétant qu'autre chose. Je me dis « Comment ils savent ça, je suis jamais venu chez eux, je suis allé sur leur site web une seule fois et ils connaissent ma date de naissance, c'est un peu étrange. ». Même si c'est des choses qu'on peut avoir très facilement, néanmoins c'est pas forcément le genre de choses qu'on va vouloir mettre en avant parce que c'est trop intrusif. C'est pas forcément le bon moyen non plus. La personne va pas agir là-dessus parce qu'elle va le prendre mal. Aucun intérêt pour nous.

### **Du coup il s'agit plus de jouer sur la fidélisation que sur la prospection ?**

Je pense que la grosse priorité pour nous c'est la fidélisation pour la bonne et simple raison que la prospection elle est déjà faite de manière assez naturelle. Notre *curse*, notre *blessing* en même temps, dans l'industrie hôtelière c'est que des choses comme Tripadvisor ont paru très très vite depuis les années 2000 et en fait c'est que, comment dire, c'est que ça a rendu le secteur très transparent. La recherche très transparente. Déjà depuis, de mémoire, 2004, dès ce moment-là, 94 % de toutes les recherches d'hôtels qui se font au moins en partie en ligne. Tout le monde va à un moment ou à l'autre au moins en ligne pour vérifier quelque chose sur l'hôtel. En général ça peut à peut être le lieu, ça peut être les photos, ça peut être une adresse, ça peut être n'importe quoi. Mais si tu veux voir, cette partie-là fait que tu as cette dynamique naturelle de procession d'une certaine manière. Donc l'intérêt quelque part pour nous c'est de fidéliser les gens qui passent nos portes parce que de toute façon tu as un renouveau naturel qui est énorme en matière de ces clients eux-mêmes. Tu as bien sûr une partie de clients fidèles hein, c'est le but quand même, mais tu as toujours une grosse partie de gens qui viennent chez toi pour la première fois de toute façon. L'idée c'est vraiment de potentiellement arriver à fidéliser ces clients qui ont de toute façon choisi de faire l'expérience de ton produit à la base. Ça coûte beaucoup moins cher de faire ça que d'aller les chercher des nouveaux clients chez les autres puisque t'as tellement d'autres facteurs qui rentrent en ligne de compte comme le prix ou la localisation. Du coup il y a des choses qui sont pas forcément contrôlables dirons-nous. Je dis ça, mais bien évidemment on fait aussi de la prospection, on cherche de nouveaux clients, c'est pas ça que je veux dire, mais la grosse priorité ça va être parce que c'est l'une des meilleures armes que nous avons pour prospecter et fidéliser en même temps quelque part. Pour revenir à ta question d'avant sur comment on utilise les données etc. La plus grosse manière pour nous de le faire, ce sont des offres trimestrielles seulement pour nos membres fidèles qui s'appelle, la promotion actuelle s'appelle *accelerate* mais on change de nom assez régulièrement. Effectivement c'est une batterie d'offres pour gagner des points bonus pour notre programme et toutes les offres sont basées sur des habitudes de séjour précédentes. Pour te donner un exemple très très concret, l'année dernière en 2014-2015, sur ce semestre là j'avais énormément voyagé pour le travail et j'avais passé le nouvel an dans un de nos hôtels à Londres et parce qu'on était avec 3- 4 potes et qu'on avait deux chambres et que on a rasé le F&B de l'hôtel coté restauration et bar (rires). Du coup l'offre que j'ai reçu en 2015-2016 sur ce même semestre de l'année 2015, l'un de nos objectifs c'était que je pouvais je gagnais 15 000 points bonus si je dépensais plus de 400 euros en F&B dans un de nos hôtels. Donc du coup c'était vraiment l'idée de regarder comment tu as dépensé l'année dernière, ce qui s'est passé, quelles sont tes marques préférées, où tu vas d'habitude etc... puis on va en extrapoler des offres spécifiques. Ceux qui dépensent beaucoup en F&B du coup on va te proposer ça. On sait par exemple que tu as tendance à voyager avec Holiday Express dans des villes secondaires et bien peut-être que cette fois on va te donner un objectif là-dessus comme un séjour en plus. Par exemple passer 13 nuits dans des Holiday Express dans des villes secondaires à la place des 12 nuits que tu as passé le semestre d'avant. Ça peut être une tentative d'*up sale* par exemple « Pourquoi ne pas tenter une nuit dans un Holiday Inn dans une capitale, puisque d'habitude vous restez plus dans des Holiday Express de villes secondaire. Pourquoi ne pas vous faire plaisir en essayant d'*up seller* dans un Holiday Inn ? ». Il y a pas mal de choses comme ça, de petits *push* et c'est vrai que c'est l'une des manières les plus efficaces pour faire ça. On a introduit ça depuis 3 ans maintenant et ça marche très très très bien. Non

seulement en terme de revenus générés, c'est quand même le but, mais également en terme de nombres de membres qui vont s'inscrire dans cette promotion. C'est quelque chose qui fonctionne vraiment très bien.

### **Avez-vous constaté une réelle amélioration que vous pouvez mesurer ?**

Ha oui oui oui, absolument ! Quoi qu'on fasse tout est mesuré au millimètre et on prend pas uniquement en compte le chiffre d'affaires, on prend en compte toute l'incrémentalité de ce qu'on fait. Sur ce type d'offre en général, le moyen c'est d'avoir des groupes de contrôle. Tu vas avoir une petite partie de ta cible que tu vas garder et que tu vas pas cibler avec ce message, tu vas pas leur parler de cette promotion et ça va nous permettre de voir les différences de comportement entre ceux à qui on n'en parle pas et ceux à qui on en parle. On voit qu'en moyenne, je te dis n'importe quoi, on a un *uplift* de 2,5 % alors que dans l'autre groupe on a un *uplift* de 2 % du coup on voit que l'incrémentalité n'est que de 0,5 % alors que si on avait uniquement regardé l'augmentation du CA on se serait resté sur la campagne à 2 %. Donc voilà (rires).

### **Comment est-ce que vous procédez-vous pour récolter les données sur vos clients ? Est-ce que vous les récoltez quand ils sont sur place, grâce à l'analyse du trafic de votre site, cookies, achats de jeux de données auprès de sociétés tierces, la géolocalisation... ?**

C'est une vaste question. Il y a énormément, énormément de choses. J'imagine que le point de départ doit être la facture du client puisqu'en hôtellerie il y a tout ce qui est PMS. Donc du coup, comme c'est relativement précis puisqu'en général tout est mis sur la facture finale, c'est facile à décortiquer et analyser très facilement. Par exemple savoir quelle est la dépense moyenne du client en F&B. C'est relativement facile puisque sur les factures qui sortent du PMS on peut voir les dépenses moyennes, ça va être ça du coup. Ou savoir quelles sont les dépenses au spa, savoir quels types de massages sont les plus utilisés ou savoir quels types de chambres sont les plus utilisées, c'est quelque chose qu'on a de toute façon sur les factures.

### **Ce sont donc des données majoritairement produites en interne.**

Absolument. Donc ça ce sont des choses que l'on a et qui sont accessibles de manière très simple pour nous. Ça va être ce que les anglais appellent le *hanging fruit*, c'est-à-dire le fruit le plus facile à attraper parce qu'on est au plus près de la branche. Mais t'as également plein d'autres choses qu'on va utiliser, notamment pour tout ce qui est publicité. Je sais pas si tu t'y connais forcément en la matière, mais les besoins de l'année dernière et de l'année d'avant c'était la publicité programmatique ou l'achat programmatique de publicité, je sais pas si ça te dit forcément quelque chose.

### **L'achat d'encarts publicitaires sur des sites tiers ?**

Ouais mais ça va un peu plus loin que ça. Typiquement, l'achat classique de bannières en ligne c'était « Tiens, j'ai envie d'acheter des bannières sur le monde.fr », je vais leur dire « Hey, ce serait bien d'avoir une bannière de tête sur votre site ! » et ils vont te dire « Super, ça fera 2000 euros pour un jour », ce qui est super nickel, on se serre la main et voici. Ça c'était comme ça se faisait à l'époque, c'était très difficile à mesurer, très difficile à récupérer on va dire, pour tout plein de raisons parce que les technologies ont beaucoup évolué depuis. Il a plein de raisons pour lesquelles ça marchait pas forcément super bien pour faire. C'est passé à la trappe comme type de publicité en ligne. C'était davantage l'achat de mots-clés sur AdWord pour tout ce qui est Google et autres moteurs de recherche puisque ça te permettait d'avoir un meilleur ciblage en fait. Donc le moteur de recherche reste très très important, particulièrement en hôtellerie et dans le domaine du voyage. Pour te donner un ordre d'idées, le plus gros client de Google à l'heure actuelle c'est le groupe Priceline, leur deuxième plus gros client c'est le groupe Expedia et en troisième c'est un de nos proches concurrents.

### **ACCOR pour ne pas le citer.**

Voilà. Donc le voyage ça leur réussit. Pourquoi ? Parce que ça marche très très bien, c'est très facile, puisque la plupart des clients vont chercher un vol pour aller là-bas. Donc tu tapes *flight* machin machin et tu vas réussir, tu vas chercher un hôtel, un hôtel à Paris par exemple et c'est ça rentre de manière très naturelle dans le parcours client typique. Donc ça c'est ce qui s'est fait pendant très longtemps. La révolution, entre guillemets, il y a 3-4 ans qui a commencé à s'opérer, c'est qu'il est devenu possible, grâce à plein plein de choses de réussir à acheter les

bannières de manière beaucoup plus intelligente. Au lieu de te dire « Je vais acheter la mienne sur un site en particulier » il est devenu possible de dire « Je vais acheter des bannières pour un public en particulier ». Je vais acheter des bannières et les montrer à des gens qui ont par exemple plus de 35 ans, qui ont dépensé tant en voyage sur leur Mastercard depuis l'année dernière, qui par ailleurs, qui ne sont par ailleurs pas encore mes clients et qui, pour une raison inconnue, aiment les tutus roses. Je peux montrer de manière spécifique à ces gens-là, à ne pas utiliser les données, dans ce cas présent, les données qui me sont propres puisque ce ne sont pas encore mes clients et qu'ils ne sont pas dans ma base de données de clients fidèles par exemple ni dans celle des clients passés. Donc je peux exclure ou inclure ce type de données. Il a des choses qui sont propres que tu as à l'intérieur de ta compagnie, après tu as des données qu'on appelle *second party* qui viennent de tes partenaires directs. Typiquement pour nous, ça va être nos fournisseurs divers donc par exemple Opera ou ce genre de choses, ça peut être si on utilise des fournisseurs particuliers en matière de technologies etc. Ça peut aussi être nos partenaires Visa, Mastercard et autres ordonnateurs de crédits, ce qui nous permet par exemple d'avoir cette idée de « Ok, je vais mettre des gens qui ont dépensé tant en hôtels » c'est-à-dire que toi, par exemple, tu dépensais individuellement dans l'hôtel, mais je veux des gens, un groupe de gens, qui en général dépensent tant dans les hôtels au cours de l'année passée par exemple. Ça me permet de qualifier mon groupe beaucoup plus facilement. Et enfin du peux utiliser ce qu'on appelle des *third party* qui sont des choses ni ne sont pas forcément reliées à toi mais qui permettent d'enrichir encore ta recherche. Typiquement ça va être tout ce qui *face brief change*, ce genre de choses qui permet à l'heure actuelle de loader une partie de tes données sur la plateforme Facebook, de chercher qui sont les gens, qui dans ta base de données à toi, sont aussi sur Facebook et de les cibler particulièrement. Ça te permet de te servir d'un post sponsorisé sur Facebook, en sachant que es déjà un de mes clients et que tu préfères tel produit de telle marque par exemple.

**Vous faites quand même appel à des tiers comme Facebook, Mastercard et Opéra en plus de votre propre base de données. C'est donc une combinaison de tous ces éléments qui vous permettent de mieux cibler votre publicité.**

Absolument. Au final t'as plein d'autres manières de le faire. Par exemple quelque chose qui est apparu il n'y a pas très très longtemps, les *high pickers* ce genre de technologies qui te permettent de tracer, comment dire, de détecter des Iphone qui arrivent dans une certaine zone. Apple a sorti ça il y a deux ans et demi ou trois ans il me semble à leur conférence. En gros, la révolution c'est que c'est en gros une sorte de routeur wifi qui pouvait reconnaître les Iphones qui arrivent dans un certain périmètre autour de lui et te permettait du coup d'envoyer des notifications sur le téléphone, si tu avais l'autorisation du client de le faire bien sûr. Nous on l'a fait de manière beaucoup beaucoup plus simple, c'est-à-dire que si tu as notre application IHG par exemple, et que tu arrives dans un hôtel IHG, on le sait car tu vas te connecter au réseau wifi de l'hôtel. Tout simplement, il faut des fois savoir utiliser des choses qui existent déjà tu vois. En fait ils arrivent sur le réseau wifi de l'hôtel et donc on sait que tu es dans cet hôtel et on peut te donner par exemple les menus du restaurant pour ce soir, ou des choses un peu utiles. Ça nous permet de changer le contenu de l'application, du coup on peut passer d'une application qui te permet de réserver à une application qui te permet par exemple de gérer ton séjour dans l'hôtel.

**Du coup, quels sont les outils auxquels vous avez recours dans votre service marketing ? Est-ce que c'est un logiciel spécifique, un tableau Excel ? Conservez-vous vos données en interne, ou ailleurs avec le Cloud ?**

Ça dépend lesquelles. Les données de réservation on les garde entièrement en interne, les données de tiers en général dans le *Cloud*, ça dépend. Ce qu'on est en train de développer, la nouvelle génération de ce qu'on appelle GRS pour *Global Reservation System*, donc la nouvelle version de Holidex qui arrive l'année prochaine. On est train de développer un partenariat stratégique avec Amadeus, et ce sera essentiellement en Cloud. Ce sera le premier programme de réservation hôtelier à être *cloudé*, en tout cas à cette échelle. On va vraiment passer à ça, mais pour le moment tout est encore en interne donc tout est basé dans nos bureaux même si c'est Atlanta, Singapour, Shanghai etc... pour les sites de réservation en soi, pour plein plein plein plein de raisons, notamment de sécurité, parce que réussir à gérer ces données physiques soi-même est en général, a été pendant très très très longtemps beaucoup plus *safe* que de le faire gérer par des sociétés tierces. Après, à l'époque, il fallait des développeurs marchant avec IBM et je sais qu'il y avait des copies ou quelque chose comme ça qui étaient dispos chez IBM, mais on a arrêté ça il y a assez longtemps. Après pour tout ce qui est Web, (hésitation) c'est pas particulièrement confidentiel, on utilise tout ce qui est *omni-tree* qui a été racheté par Hadopi et qui est maintenant Hadopi Analytics,

c'est ce qui fait partie de la suite. Ça c'est pour tout ce qui est données web et digital et c'est quelque chose qu'utilise également pour nos achats de publicité en ligne. La grande problématique maintenant c'est de réussir à pouvoir actionner ce que tu récoltes en matière de données, très rapidement et de manière très simple et si possible de manière instantanée. C'est-à-dire que je veux pouvoir réussir à cibler quelqu'un avec une bannière publicitaire ou quelque chose un email, une notification sur une appli. C'est pour cibler quelqu'un dès le moment où il rentre ou dès que je réussis à avoir ce pan de données qui permet de mieux le cibler et je vais pouvoir le faire de manière instantanée. C'est là qu'est le vrai challenge à l'heure actuelle, c'est plus vraiment sur la récolte, même si c'est bien sûr super important l'actionnabilité des données que tu récupères. Donc on revient un peu au point originel qui est « Est-ce-que tout est intéressant ? Et surtout comment est-ce-que tu réussis après à en faire quelque chose. C'est pour ça que tu as vu apparaître des nouveaux jobs d'une certaine manière, qui en soi ne sont pas nouveaux, comme *data scientist*, comme *data engineer*, ce genre de spécialistes dont le job est justement de trouver des nouvelles manières de faire parler tes données. Les corrélations c'est facile à trouver. Par exemple je peux trouver une corrélation entre le niveau d'éducation moyen des personnes et le taux de salinité de l'eau dans l'Essonne et je suis sûr que j'arrive à trouver une corrélation là-dessus, mais ça veut pas dire que les deux sont liés.

### **Le but du jeu est donc de trouver des corrélations pertinentes pour pouvoir analyser et prévoir des comportements.**

Je te donne un exemple qui est super intéressant. je sais pas si tu connais INQ qui est une sorte de Castorama anglais, on avait pas mal discuté avec eux notamment parce que ça a été un des premiers à faire des choses intéressantes, pas forcément en matière de *data mining*, mais d'utilisation des données qu'ils récoltaient. Une des choses qu'ils ont fait et qui étaient super fun, c'était qu'il y a quelques années ils ont commencé à essayer de regarder et de mettre en place des corrélations et de se dire « ok, par exemple si on prend des données comme le temps qu'il fait, on arrive à trouver des corrélations entre le temps qu'il fait et le type d'objet qu'on vend ». Ils sont comme Castorama ils vendent des trucs à l'extérieur donc par exemple, si il pleut dehors, est-ce-que du coup ça affecte nos ventes de mobilier de jardin ou de clôtures de jardins ? Du coup, quand est-ce-qu'on doit parler à nos clients. Si on a une prévision de météo qui te dit qu'il va faire beau ce week end, il est jeudi, et si on s'adresse aux gens et qu'on leur dit « Hey, du coup venez chez nous, c'est l'occasion de faire du jardinage ! ». Est-ce-que ça marche ? Ils ont testé plein de choses et ils ont vu que plein de choses marchaient. Mais l'un des trucs les plus étonnants auquel ils avaient pas forcément pensé à la base en testant tout ça, c'est qu'ils ont réussi à augmenter leurs ventes de clôtures de jardins de manière massive en l'associant à la force du vent. Ils prévoyaient des bourrasques de vent 24 heures en avance et ils mettaient des bannières en ligne par rapport à ça dans les zones où les bourrasques de vent allaient apparaître, 50km/h quelque chose comme ça, et ils ont envoyé un truc en disant « Alerte météo, grosse rafales de vent, n'oubliez pas de protéger votre jardin grâce à cette clôture ». Efficace. Et ça leur a permis de doubler leurs ventes de clôtures de jardins l'année dernière. C'est rigolo parce que c'est quelque chose qu'ils avaient pas vraiment prévu. t'associes pas forcément les deux et t'y penses après. C'est vrai qu'à l'époque, au moment où ça a été fait les gens se sont dits « Ha ouais c'est vrai que c'est malin, mais il fallait le voir, le tester et voir si ça marche. C'est là que ça devient intéressant et qu'il faut pas tomber dans le piège de tester nécessairement n'importe quoi. On peut tout tester très rapidement maintenant via bannières en ligne d'abord ce qui coûte cher mais te donne des résultats instantanés du coup tu peux te dire très facilement « ok, ça marche et maintenant on va faire ça.. ». Typiquement niveau météo t'as des marques bière ou de cidre qui le font très bien. Soit ils le font avec des données météo macro, soit via le *billboard* en lui-même qui arrive à voir la température, et si température passe au-dessus d'un certain seuil, ils *switchent* la publicité qu'il y a sur le *billboard* et ça va être « Se réchauffer autour d'une bière à la maison » à « Profitez d'une bière au soleil autour d'un pique-nique ce week end ». Ça te permet d'adapter beaucoup plus de choses, pas seulement modifier ton message client, mais également ce qui est créatif, ce qu'est la campagne en elle-même de manière assez simple.

### **Je souhaiterais revenir sur les bases de données. Avez-vous mis en place une base de données commune pour tous les hôtels IHG ou est-ce-que chaque territoire a sa propre base de données ? Avez-vous par exemple accès aux données des hôtels de Shanghai ?**

Alors, nous fonctionnons majoritairement avec un système de franchise pour le management de nos hôtels. Dans l'hôtellerie très classique tu as le propriétaire des murs, le gérant de l'hôtel, et qui appelé son hôtel L'Hôtel de la Mairie par exemple. L'hôtel et l'enseigne, tout est géré par une seule personne qui est propriétaire de l'hôtel. Aux US HolidayInn était la première marque d'hôtellerie franchisée à l'époque, dans les années 50. C'était de la franchise

pure, c'est-à-dire que tu allais voir un propriétaire qui veut faire un hôtel dans ses murs et au lieu de mettre un panneau hôtel de la mairie, c'est très bien mais concrètement ça va pas mener à grand-chose, on va leur dire au contraire de mettre un panneau HolidayInn et d'utiliser notre système de réservation, profiter de nos techniques et de canaux de distribution etc. Ça peut être beaucoup plus efficace pour eux de profiter de la force de notre marque en échange d'une *smallfee* qui va bien évidemment nous rémunérer. Du coup c'est le propriétaire qui manage son hôtel mais qui va le faire selon les standards d'une marque. je pense que c'est plus intéressant pour eux parce que ça leur permet d'attirer une clientèle qui par exemple peuvent être de nos clients fidèles. Entre les deux t'as le *contract management* où c'est le franchiseur qui va aussi être le manager de l'hôtel pour un propriétaire tiers qui détient les murs. Comme la très grande majorité de nos hôtels sont en franchise pure, c'est quelqu'un d'autre qui les gère et qui les détient, parfois deux personnes différentes par ailleurs. Pour te donner un ordre d'idée, sur 5 054 hôtels on doit en avoir 600 managés et le reste en franchisé. Tout ça pour dire que ce qui est intéressant pour nous c'est tout ce qui est dans la protection des données privées et l'utilisation des bases de données entre différentes entités juridiques dans différents pays, si bien qu'on peut pas forcément donner toutes les données client à nos franchisés parce que c'est nous qui avons pris la réservation donc le traitement des données nous incombe et leur sécurisation également. On va pouvoir donner des données essentielles à nos franchisés bien évidemment, ils doivent savoir des choses sur les clients pour pouvoir les accueillir, mais pas leur donner toutes les données de contact par exemple, des adresses ou ce genre de choses parce que ça nous ferait un soucis au niveau légal, nous sommes responsables des données en question. Au niveau politique global, on prend la politique de protection la plus stricte, à l'heure actuelle en Europe c'est l'Allemagne qui a une politique de protection la plus stricte. Du coup on prend ça pour l'appliquer à tout le monde parce que c'est plus dangereux d'utiliser par exemple la politique états-unienne avec laquelle nous serions dans l'illégalité dans la plupart des pays européens, ce qui ne serait pas souhaitable. Du coup on part sur les standards les plus stricts comme ça c'est plus simple. On le fait d'une certaine manière parce que dans des pays on va pouvoir savoir certaines choses et du coup on va vraiment s'adapter. Si c'est vraiment spécifique au sein d'un pays, on essaie de voir dans quelle région on peut trouver le plus petit dénominateur commun et de se lancer là-dessus. Ça fait des économies d'échelles beaucoup plus importantes plutôt que d'essayer d'avoir des politiques différentes dans chaque pays. Surtout que ça change très très régulièrement. Il est plus facile de se tenir au courant des 4 plus gros pays et on est en lien avec tout ce qui est union européenne, ce genre de choses pour se tenir au courant de ce qui se passe au niveau de la législation au niveau européen puisque cela nous impacte directement par la suite. Du coup je ne me souviens plus de la question originelle. Ha oui, les données communes ! Du coup oui et non, puisqu'on a notre base de données, les commandes des toutes les réservations au niveau de IHG en tant que franchiseur, mais des franchisés vont avoir leurs propres bases de données qui suivent leur PMS de leur côté et ça va pas être mis en commun avec d'autres hôtels nécessairement. Ça répond à ta question ?

**Tout à fait. J'imagine que mettre tout ça en place ça a quand même un coût très important entre l'achat de programmes, le recours à des personnes spécialisées. Du coup pensez-vous que seules les grandes chaînes ont la possibilité de le faire et non pas les indépendants ?**

(Silence) Oui et non. Je pense que la plus grosse barrière à l'entrée des hôteliers indépendants c'est vraiment le savoir, le savoir-faire. Plus que le coût en soi. Dans le sens où tu pourrais considérer que le savoir-faire est un coût, effectivement, ça pousse à engager des gens qui savent le faire mais pour moi c'est vraiment un problème de savoir-faire parce que tu as plein de plateformes qui te permettent de faire plein plein de choses en ligne et peu chères. Et puis t'as pas à tous les utiliser en tant qu'hôtelier indépendant. Si t'arrives à avoir un programme d'emailing assez customisable et qui marche assez bien c'est déjà un gros atout et t'as très peu d'hôtels indépendants qui le font correctement. Mais t'as plein de solutions pour le faire de manière assez simple. Mais ça reste un problème de savoir-faire plus qu'autre chose. Après, ne serait-ce que des logiciels que tout hôtelier veut utiliser, indépendant ou pas, type Opera ou ce genre de choses, ce sont des logiciels de bases de données à la base. On leur donne d'autres noms mais au final ce sont des logiciels de bases de données. Tu pourrais le faire avec Excel ce serait la même chose. Du coup c'est plus le savoir-faire pour utiliser, extraire les données, d'un Opera par exemple, et savoir les utiliser dans un plan marketing ou dans une communication plus qu'autre chose. Après tu serais surprise ; on travaille avec beaucoup d'hôteliers indépendants bien évidemment, qui ne peuvent avoir qu'un seul hôtel et qui se trouvent être avec nous mais qui arrivent néanmoins à mettre en place des politiques d'utilisation de données assez poussées sans pour autant avoir des moyens pharamineux pour le faire. C'est tout bête, mais si t'as Excel et un PMS tu peux faire un publipostage sans trop de soucis. C'est un peu *oldschool* quelque part, par rapport à toutes les

techniques avancées dont on vient de parler, mais ça peut être tout aussi efficace si c'est bien fait. Ce qui va manquer à un hôtelier indépendant c'est l'échelle et l'adaptation aux tendances du marché à l'heure actuelle qui font que les séjours sont de plus en plus courts, qui font que les gens ne restent pas fidèles à une destination en particulier d'une année sur l'autre mais c'est plutôt de découvrir des nouvelles choses. Là où ça va devenir compliqué c'est comment fidéliser un client en étant indépendant, comment le faire revenir sur cette destination. C'est là que c'est très compliqué et c'est pour ça que les groupes marchent fort. C'est comme aller dire à quelqu'un « On a vu sur notre site que vous avez fait Paris, Londres et Berlin, pourquoi ne pas faire un petit tour à Porto ?

### **Du coup est ce que vous pensez que toutes ces techniques-là se substituent aux techniques de collecte de données traditionnelles comme les questionnaires de satisfaction, les sondages... ?**

Pout nous tout est lié. C'est-à-dire que notre questionnaire de satisfaction est de toute façon lié à notre base de données à la base des réservations. On envoie notre questionnaire de satisfaction interne, c'est-à-dire pas Tripadvisor, mais tout ce qui est technologies propres qui nous permettent de faire ça et du coup on va l'envoyer à un échantillon statistiquement significatif de nos clients par email du coup tout est lié à notre base de données de réservations. T'as pas besoin d'avoir des choses séparées, néanmoins ça va être très important d'avoir des choses comme des enquêtes de satisfaction puisque ça te permet de pousser encore plus en avant ta customisation derrière. L'un des outils qu'on fournit à nos hôtels ce sont les PMS. C'est-à-dire tout hôtel va normalement le matin ou le soir, un rapport avec la liste des arrivées du jour d'après. Tu vas regarder et tu vas voir qu'il y a 64 personnes qui arrivent, les demandes spéciales, on voit qu'il y a tel ou tel VIP, ce type de choses. Ce que nous on leur fournit en plus c'est une liste qui est enrichie avec des données qu'on a. C'est-à-dire que déjà tu vas pouvoir les ordonner par ordre d'importance et mettre en haut les gens qui ont bien sur le plus de points de fidélité, mais aussi qui sont *Elite* parce qu'on a plusieurs niveaux de statuts donc les *Spire*, les *Platine*, et les classiques. Bien évidemment on va mettre les plus hautes statuts en premier mais on va également mettre en avant d'autres choses : par exemple quelqu'un qui a eu une enquête de satisfaction au-dessous d'un 8 sur 10 dans son hôtel précédent, c'est quelque chose qu'on va mettre en avant puisque ça permet à l'hôtel qui le reçoit, qui n'a pas forcément été l'objet de l'enquête de satisfaction en question, de faire en sorte que l'expérience se passe bien puisqu'on sait que cette personne a eu une expérience peu satisfaisante ou moins que super satisfaisante lors d'un séjour précédent. Donc on va essayer de se rattraper sur les points sur lesquels il a été déçu précédemment. C'est encore une fois une manière de customiser ton service grâce à l'historique que tu as sur tes clients et grâce aux données que tu récupères sur tes clients et c'est ça qu'il est important de savoir-faire de manière pertinente au final. T'as toujours plein de technologies, plein de nouvelles choses qui apparaissent, mais ce dont il est important de se souvenir c'est qu'au final notre cœur de métier c'est le service, c'est créer de l'émotion, c'est de réussir à accueillir un client au mieux et à faire en sorte qu'il vienne ou revienne chez nous et c'est ça qu'il faut réussir à bien faire. Réussir à envoyer un email intéressant comme INQ qui a réussi à trouver une corrélation entre clôture de jardin et vitesse du vent, c'est très bien, c'est très beau, ça marche, c'est top, mais ce qui, à mon sens est encore plus important, c'est de réussir à donner des armes à nos collègues pour faire leur travail au mieux. Le meilleur marketing que tu puisses avoir c'est quand l'expérience de ton produit est satisfaisante. C'est là où les données prennent tout leur sens au final et on revient quelque part aux tout débuts avec les gravures des VIP derrière. La différence c'est que maintenant tu as les gravures en question sous forme de fiches clients beaucoup plus importantes sur tous les clients. Après il faut réussir à comprendre quels sont les points de données qui sont les plus importants pour pouvoir mettre des priorités dans ta liste des arrivées et comme tu arrives à adapter ton discours par rapport à ça. Typiquement, on va voir que ce membre évite de télécharger notre application et on va lui demander de le faire pendant le *check-in* parce que ça lui permet de gérer ses séjours de manière beaucoup plus simple. T'as plein d'actions que tu peux faire mais faut savoir reconnaître quelles sont les plus intéressantes pour que tes clients partent de ton hôtel et se disent « Wow ! C'était top ! ».

### **Vous arrivez à avoir les visages de vos clients ?**

Non, ça on le fait pas parce que c'est interdit dans la plupart des pays européens d'avoir de la donnée comme ça. Mais après tu as d'autres moyens de le faire dans d'autres pays comme les États-Unis ou le Japon et dans lesquels ça pose aucun problème. Tu as la manière très manuelle de prendre ton top 5 de ta liste d'arrivée et les chercher sur LinkedIn, c'est bête mais ça marche bien. Mais la question c'est de savoir si ça t'apporte vraiment quelque chose. Mais c'est une bonne question parce que tu commences à voir arriver des trucs intéressants sur le visage qui peut par exemple te servir comme clé de chambre, donc tu as la possibilité de manière assez simple en plus, de mettre un mini webcam au-dessus de chacune de tes chambres et vu qu'il n'y a pas de serrure sur ta porte de

chambre, ton visage est enregistré à la réception lors de ton arrivée et quand tu arrives à ta chambre la webcam reconnaît que c'est toi et t'ouvre la porte de ta chambre. Après c'était un hôtel au Japon qui avait tenté de faire ça et c'est pas quelque chose que l'on ferait parce qu'encore une fois il ne faut pas traverser la barrière entre savoir ce qui est intéressant et ce qui *creepy* et là pour le coup on est dans un truc assez inquiétant. Mais potentiellement, peut-être que dans 10-15 ans c'est ce qui pourrait marcher. Qui aurait pu te dire il y a une dizaine d'années que tu allais donner toutes tes données et partager activement tous tes trucs avec un type sur une messagerie instantanée ? Au final aujourd'hui tout le monde est sur Facebook et partage ses données sans même s'en rendre compte, malheureusement. C'est un trésor pour Facebook qui vent ces données à prix d'or, mais faut bien comprendre que néanmoins tu payes pour le service avec tes données.

### **Par curiosité, est ce que vous savez combien ça coûte d'acheter des données sur Facebook ?**

La manière dont ils le font c'est pas exactement acheter des données comme ça. C'est-à-dire qu'ils vont pas te donner des packs de données que tu peux payer 10 000 d'un coup, c'est plus utiliser leur plateforme et leurs données d'un point de vue publicitaire. Typiquement, le moyen le plus simple de faire ça, tu peux le faire toi-même, c'est ouvert à tout le monde. Imaginons que pour ton mémoire tu voudrais avoir des réponses à ton questionnaire et tu veux que ton questionnaire touche particulièrement les gens qui sont dans une position similaire à la mienne, tu pourrais très facilement utiliser Facebook pour faire ce qu'on appelle un *custom audience* dont le but est d'amplifier l'audience de ton post, vu que ton poste ne va être visible que par tes amis et qui va le montrer à une toute petite partie de tes amis, parce que Facebook pense savoir mieux que toi lesquels de tes amis vont être les plus intéressés par le post en question, mais tu peux amplifier l'audience en payant pour. Tu vas aller sur leur plateforme et tu vas leur dire que tu voudrais bien toucher des gens qui ont comme lieu de travail Accor, Marriott, Starwood et consort, avec un titre de poste en lien avec le marketing et qui vit en France ou en Angleterre par exemple. Le modèle de Facebook est très simple, ils n'ont pas vraiment inventé la poudre là-dessus, c'est soit en CPM, soit en CPC donc coût par clic ou coût par mille. En publicité en ligne le cout par milliers d'impressions. A chaque fois que ton questionnaire est montré à cette audience-là et « imprimée » mille fois sur une *timeline*, ou à chaque fois que quelqu'un clique sur le questionnaire, tu vas payer un certain prix. En fait le prix en question est une enchère basée sur une fourchette où on te dit que tu devrais payer à peu près ça mais ça dépend aussi du nombre de gens qui essaient de toucher la même audience en même temps. Pas forcément la même audience, mais ces gens-là peuvent aussi faire partie d'une audience que Mc Donalds essaie de toucher, tous les gens qui habitent en France par exemple, et si McDonalds est prêt à payer 3 dollars du clic pour les toucher et que toi t'es prêt à payer 0.5 dollar du clic, il a plus de chance que la publicité qui va effectivement s'afficher ça va être celle de McDonalds plutôt que la tienne. Tu peux aussi mettre des budgets maximum par jour, mais pour te donner un ordre d'idées en matière de CPC, en général c'est autour de 40-45 centimes du clic.

### **C'est beaucoup !**

Oui et non. Ça dépend. Si je paye 40 centimes du clic et que tu me prends une chambre à 150 euros, le coût d'acquisition est pas énorme en soi. Si je paye 35 centimes du clic et que tu prends un menu Maxi Best of, le coût d'acquisition est moins intéressant. Ça dépend.

### **Du coup Facebook ne vous donne pas directement de données, mais vous autorise à intégrer son système pour y insérer des publicités ?**

Absolument. Il a plusieurs autres manières de le faire, plus fines et n'importe qui peut le faire, même un hôtelier indépendant et de manière très facile. Tu peux aller dans l'administration de ton Opéra ou de ton Briand etc, donc de ton PMS, aller sur Facebook, tu *upload* ta liste de gens et d'adresses emails et Facebook te dit, selon son algorithme d'analyse, lesquels sont présents sur Facebook. En gros tu intègres ta base de données sur le compte Facebook. Ils vont pouvoir te dire que 64 % de tes clients sont présents sur Facebook et tu vas pouvoir afficher ce qui s'appelle le *customer audience* et l'utiliser soit en positif, soit en négatif. C'est-à-dire que soit tu vas leur parler sur Facebook en leur disant « Vous êtes déjà venus chez moi, voici 20 % pour la prochaine fois », ou tu peux l'utiliser de manière négative en disant « Je veux toucher n'importe qui mais pas ceux-ci » puisque tu les as déjà vu et tu vas donner une offre spéciale « première fois dans mon établissement » ou ce genre de trucs. Tu peux même aller plus loin : cette fonction qui permet de coder la base de données et qui s'appelle *custom audience*, t'en as même une autre qui a dû changer depuis, mais qui s'appelle *similar audience* permettait d'analyser les gens que tu venais d'*uploader* et

avec qui ça avait *matché* sur leur propre base de données et de trouver d'autres profils similaires à ceux-là. Ce ne sont pas gens qui sont dans ta base de données, mais similaires dans leur base de données à eux. Ça va te permettre d'évaluer le nombre de gens selon un nombre de critères donnés du genre « j'ai des clients franco-français et je veux baser les audiences sur ce segment en soi, mais avec des gens qui sont basés dans les plus grandes villes françaises », et après tu peux leur envoyer un post sponsorisé ou une bannière par rapport à ton hôtel. Et c'est là que 40 centimes du clic est-ce-que c'est si cher que ça comparé à un affichage dans la gare de Lyon ? Ce que beaucoup d'hôtels font par ailleurs et qui vont te coûter 10 000 euros du mois. est-ce-que tes 10 000 euros ne sont pas mieux dépensés en faisant ça, sachant que c'est là que les gens réservent plutôt qu'un panneau d'affichage quand les gens arrivent à la gare et donc à leur destination. C'est une question de bon sens.

**Imaginons que vous n'ayez aucune contrainte légale, de stockage ou de coût. Quelles sont les données auxquelles vous aimeriez avoir accès ?**

(Silence) C'est difficile à dire. Parce que tu sais pas en avance si elles vont t'intéresser et c'est ça le *catch* quelque part. Je pourrais dire que j'aimerais savoir combien quelqu'un dépense en note de supermarché, mais au final, est-ce que c'est vraiment intéressant ? J'en sais rien. Savoir qu'ils dépensent 3 euros par mois en produits surgelés et qu'il pourrait bien être clients chez HolidayExpress, peut-être, mais je peux pas le savoir avant d'avoir les données en question. Du coup c'est pas forcément si facile que ça de donner une réponse à cette question puisqu'au final ça dépend de la pertinence du point de donnée en soi. T'as par exemple des vendeurs qui te proposent d'utiliser des données assez intéressantes comme par exemple Adarra qui sont avec des gens comme Calcana et qui vont en fait pouvoir te dire « Vous allez pouvoir vous servir de cette publicité avec des personnes qui viennent de *booker* un vol vers une destination particulière ». Et ça c'est intéressant parce que tu te dis que la corrélation doit être importante : ils viennent de *booker* un vol et ils vont sûrement chercher un endroit où séjourner très bientôt. S'insérer dans ce moment est intéressant. Si Voyages SNCF était une boîte innovante, ils auraient déjà un système en place qui permettrait de proposer ce type de choses à des publicitaires dans le domaine du voyage comme le fait très bien Facebook par exemple ou Skyscanner. Si j'étais eux je l'aurais fait depuis au moins deux ans et ils seraient beaucoup moins en déficit que maintenant. C'est sûrement un peu tard maintenant. T'as peut-être plus intérêt de jouer sur des inter échanges, même si c'est un peu difficile évidemment, mais il a sûrement un coup à jouer en la matière et d'aller par exemple voir Air France pour voir si ça ne les intéresserait pas de faire un partage de données là-dessus puisque ça pourrait être intéressant pour nous d'acheter ce type de données en tant qu'hôtelier. Nous on fait ça pour des *record card* avec les agences de location de voitures, vu qu'ils sont tout en bas dans la chaîne, et où tu vas en général après avoir *booké* ton avion, du coup ça va les intéresser de toucher nos clients, comme nous ça va nous intéresser de toucher les leurs et celles des compagnies aériennes.

**Du coup, est-ce-que vous revendez vos données ?**

On ne les revend pas exactement, c'est plutôt des partenariats stratégiques. C'est-à-dire qu'on ne fait pas payer pour, c'est plus une question d'échanges en nature on va dire. Si nous ça nous intéresse de savoir quels sont leurs clients parce qu'ils peuvent aussi nous intéresser, tout comme nos clients les intéressent. Du coup on va faire un partenariat basé sur nos systèmes de fidélité. On donne la possibilité à leurs clients de gagner des points sur notre programme de fidélité et vice-versa. Si par exemple le client loue une voiture chez Hertz, ils peuvent gagner des points sur notre carte de fidélité. Pour nous c'est intéressant parce que ça nous d'ajouter de nouveaux clients à notre carte de fidélité, mais également de les faire utiliser nos points encore plus et on va aussi pouvoir leur donner des placements dans nos communications, comme à nos clients. C'est aussi intéressant pour nos clients de toutes les façons. Tout est question de pertinence. Ça peut être utile pour nos clients de dire qu'ils peuvent avoir 30 % de réduction sur une voiture. Ils n'y auront pas pensé, mais vu que le prix est intéressant, *why not* ? C'est intéressant pour nos clients et c'est intéressant pour nous aussi.

**Quel est votre ressenti par rapport à l'aspect éthique de tout ça ? Parce que de manière générale, les gens ne se rendent pas compte que les entreprises les connaissent très bien grâce aux traces qu'ils laissent un peu partout, du coup ça se fait un peu à leur insu.**

Oui et non. Parce que tu connais des groupes, des agrégats de personnes, mais tu connais jamais une personne en soi. Je vois ce que tu veux dire, c'est que quand t'es sur un site et que tu vois une publicité pour Amazon qui te dit que tu as visité un site pour une crème pour, je sais pas, les hémorroïdes par exemple, et ils vont te proposer de la

racheter, tu vas te dire « c'est gentil mais j'utilise le PC de quelqu'un d'autre », mais en soi est-ce que c'est pertinent pour moi, oui quelque part. Je préfère voir ça plutôt qu'une publicité pour les couches sachant que j'ai pas d'enfant. Mais à choisir entre une publicité qui m'intéresse et une autre pour quelque chose qui ne m'intéresse pas du tout, ben je préfère ne pas avoir de publicité du tout. Mais sachant que c'est quand même ça qui fait tourner Internet de nos jours, c'est un moindre mal je dirais. Tant que ça reste en agglomérats, tant que ça reste quelque chose qui est anonymisé, c'est faisable. Après ça être effectivement utilisé pour ce genre de choses et c'est très puissant. Je sais pas si tu as vu la polémique avec Banksy qui a été démasqué. En analysant plein de données qui étaient dispo publiquement sur le web, ils ont réussi à faire des corrélations entre la localisation d'un mec et Banksy, grâce à ses graffitis et ils ont affirmé à 98 % qu'il s'agissait de ce mec-là. Du coup tu te dis que c'est assez pervers quand même comme manière de faire les choses. Mais c'est une réalité et malheureusement c'est une chose qu'il est possible de faire maintenant de manière très facile quand t'as le nom de quelqu'un ou même juste leur photo ou quelque chose comme ça, tu facilement savoir s'ils sont en vacances, le voir sur Instagram et te dire « je vais aller me faire leur maison pendant qu'ils sont pas là ». Pour te dire, il y a même eu une campagne l'année dernière du gouvernement ici à Londres pour prévenir les gens en leur disant « ne postez pas de choses sur Instagram, assurez-vous que vous n'êtes pas pré-traçables, mais c'est impossible, c'est quand même une incitation digitale pour venir cambrioler votre maison ».

### **Du coup, quelles sont les perspectives d'avenir ? Que ce soit pour les technologies, votre entreprise ?**

C'est une bonne question. Je crois qu'il y a plein de manières de potentiellement se développer et ça va être très difficile d'aller contre de toute façon. Je pense pas que les gens soient prêts à payer pour des choses comme Facebook, Gmail et autres. Si tu me dis maintenant que Tinder va être payant à 10 euros par mois et que Facebook c'est 25 euros, etc. je pense pas qu'il y ait tellement de gens qui vont s'intéresser à l'idée de payer pour ça. Peut-être que certains oui, mais je pense qu'il faut savoir accepter le fait que si tu utilises des services gratuits, le prix à payer c'est tes données. Je pense que c'est surtout un problème d'éducation des personnes, savoir bien éduquer les gens pour qu'ils connaissent mieux les risques. Les paramètres de Facebook ne servent pas à grand-chose d'une certaine manière. c'est plus être conscient de ce qu'on poste, quand est-ce qu'on le poste, en tant que qui. Mais bon voilà, après tout dépend de ce que l'on poste. Je veux dire, toute ma vie est sur Internet, si je veux savoir ce que je fais jour après jour, c'est pas très compliqué d'aller le chercher. Mais je le sais, je connais les risques et c'est mon choix. Je pense que pour la plupart des gens c'est pas, en l'état actuel, un choix rationnel et c'est là que c'est potentiellement dangereux. (Silence).

### **En même temps, vous avez peut-être intérêt à ce que ça ne se sache pas trop.**

Non, parce qu'avec nous ça va toujours être très transparent. On a un programme qui permet à nos clients fidèles de mettre ce qu'on appelle des *perfecties* donc ils peuvent nous dire « je préfère une chambre proche de l'ascenseur lorsque je pars de l'hôtel, je préfère être à l'étage supérieur ou à l'étage inférieur », ce genre de choses. Nous on leur demande ce qu'ils veulent nous donner. S'ils veulent nous donner leur date d'anniversaire ils nous la donnent, si il veulent pas ils le font pas. C'est leur choix. Je pense qu'il est super important de donner ce choix aux gens, de donner cette possibilité de ne pas nous les donner, justement pour éviter les *bad clash* par la suite. Donc c'est plus intéressant que ce soit un choix dès le départ parce que du coup tu vas avoir une bien meilleure compréhension de ce pour quoi la donnée est demandée et de ce pour quoi elle va être utilisée et t'as un bien meilleur retour par la suite. Par exemple tu indiques sur le formulaire d'inscription ce pour quoi la donnée va être utilisée et que tu mets des exemples de ce qu'on va t'envoyer, t'auras un bien meilleur retour, ton taux de d'inscription va être moindre, mais ton taux de désinscription aussi. Du coût l'audience que tu vas avoir va être beaucoup plus fournie et beaucoup plus qualitative. Il a des gens qui vont cliquer plus, qui resteront beaucoup plus longtemps sur ton site parce que les gens vont être beaucoup plus intéressés par ce que tu vas leur dire. C'est super important de rester comme ça parce qu'au final c'est plus intéressant pour la compagnie, c'est plus intéressant pour les clients également. Plus, que ça, même d'un point de vue pécuniaire d'une certaine manière, la confiance est quelque chose de défini parce que notre donnée principale va jamais être le revenu pour une transaction en particulier. C'est ce qu'on appelle le *last time revenue*. Ça va être plus intéressant pour nous de se dire que ce mec de 16 ans qui vient d'arriver chez moi et qui a commencé par un séjour chez Holiday Express à 45 euros la nuit, je vais essayer de le traiter mieux que la personne qui va payer une chambre 500 balles chez Intercontinental parce que qu'on pense que dans sa *timeline*, sa durée de vie va être beaucoup plus importante. Et c'est là que le sentiment de confiance est primordial. Tu peux pas te permettre de commencer cette relation avec le client en lui disant « Haha, on t'a bien eu, on a réussi à choper

ton adresse mail et on va pouvoir t'envoyer des emails tout le temps ! ». Non. C'est pour ça qu'on a des caps sur le nombre d'email qu'on envoie à nos clients, de mémoire on doit être à 3,5 tous les mois mais on a aussi des caps à 0,5, ce qui est assez étrange. En avril on était à 3,5 emails par mois par client et on n'en avait pas plus. Parce qu'on se dit que plus que ça, non seulement ils les liront pas, mais ils vont aussi se désinscrire. Même si ils se désinscrivent pas, ils ne vont plus les ouvrir donc c'est pas très intéressant. De la même façon que tu peux pas matraquer quelqu'un en le suivant partout sur un site, même si tu sais que tu peux leur faire du reciblage en leur rappelant qu'ils ont consulté la page du Crown Plaza de Manchester et qu'ils peuvent avoir une réduction sur cet hôtel. Tu vas leur montrer une fois, tu vas leur montrer deux fois, mais si tu vois qu'à la troisième fois ça a toujours pas cliqué tu t'arrêtes là. On pourrait suivre quelqu'un 17 fois, mais ça sert à rien. Aussi, juste un point là-dessus, c'est que c'est pas dans notre intérêt non plus parce que ça coûte de l'argent. Envoyer un email, faire apparaître une publicité devant toi, etc. ce sont des choses qui coûtent de l'argent donc on va le faire que si on sait qu'on va avoir un retour derrière, que ce sera intéressant et que ce sera une bonne manière de pouvoir te contacter. Si tu n'es pas quelqu'un qui est intéressé et que tu t'es inscrit pour un jeu concours pour gagner des lunettes, c'est d'ailleurs pourquoi on fait très rarement ce genre de choses, on sait que ça attire les vautours d'une certaine manière et qui viendront chez nous donc c'est pas très intéressant. La pertinence, toujours la pertinence. (rires). C'est plus intéressant pour tout le monde et au final. C'est un peu le débat à Bruxelles en ce moment. Est-ce qu'il vaut mieux réguler tout ce qui gestion des données etc. ou les laisser au marché ? Les partisans du libéralisme vont dire que c'est plus intéressant de les laisser au marché parce que de toute façon ça va s'auto réguler et ceux qui vont en abuser vont couler de toute façon mais ça prend du temps et en attendant ça peut être un peu dangereux. Les gouvernementalistes vont te dire que les laisser au marché c'est très bien mais le danger c'est que pendant 6 ans ce soit le bordel.

#### **Le but au fond est d'arriver à avoir une relation équilibrée avec le client.**

En tant que *shopper* par exemple. Plein de gens aiment ce genre de choses et vont aller chez H&M ou chez Machin parce qu'ils aiment bien le produit, ils aiment bien comme il fonctionne. Au final c'est le challenge pour la compagnie de répondre à tes besoins et ça peut être intéressant. Le nombre de fois où on nous a dit, parce qu'on était trop prudents sur nos communications, « non, on va pas l'envoyer à cette personne là parce qu'elle a déjà reçu trois emails et que du coup on va pas leur envoyer notre dernière offre. Par contre il est déjà arrivé qu'on le mette sur la page Facebook ou ce genre de choses et que les gens nous disent « on fait partie des gens à qui on envoie des mails d'habitude, et j'aurais pu l'utiliser lors de mon dernier séjour ». Donc t'as aussi cette demande des clients. Il faut juste savoir le faire bien et le donner aux gens que ça va intéresser. Si tu t'abonnes à la *newsletter* de Zara et qu'il t'envoient des *voucher* avec - 30 %, j'imagine que ça va t'intéresser. (Silence) Si tu perds la confiance d'un client ça va aussi se transposer sur les autres marques, il va te voir comme intrusif parce que tu auras trahi sa confiance. C'est pour ça qu'il faut faire vachement attention avec l'achat de jeux de données et de trucs comme ça parce que tu as des problèmes de pertinence qui arrive toujours derrière. (Silence).

#### **Pensez-vous que l'exploitation de donnée soit aujourd'hui, et dans l'avenir de toute façon, sera quelque chose d'incontournable ? Vu que l'hôtellerie est un secteur ultra concurrentiel est-ce que ce sera une manière de sortir son épingle du jeu ?**

Oui c'est inévitable. Mais pour moi c'est quelque chose qu'on fait depuis des siècles et qui n'est pas nouveau. Simplement, les technologies qui permettent de le faire sont nouvelles et plus pratiques qu'à l'époque. Tout le monde utilise Facebook sans savoir comment ça marche, sans savoir ce qui tourne derrière et personne ne s'en retrouve plus mal. On en arrive à faire des voitures qui vont bientôt pouvoir se *driver* toutes seules et les gens n'ont pas besoin de savoir l'algorithme qu'il y a derrière et comment ça marche. Ils veulent juste que ça marche tranquille et être en sécurité quand tu te fais conduire par la voiture. De la même manière, je pense qu'il est important de savoir s'insérer dans les habitudes clients et de nos collègues en hôtel pour que ça leur permette de faire leur travail comme ils le font normalement depuis des années. Quand on a introduit les nouvelles listes d'arrivées enrichies, au final le format est quasiment de même que les listes d'arrivées précédente, la manière de les gérer est quasiment la même etc. Et tout ça ce sont des choses que l'on a planifié à l'avance en se disant que si tu changes trop les habitudes de quelqu'un il ne va pas bien le faire. Alors que si tu lui proposes quelque chose qui ressemble à celui qu'il utilise déjà mais qui apporte une facilité supplémentaire, alors ils vont l'utiliser. On dit toujours qu'il est très difficile de changer les habitudes mais c'est pas vrai. C'est très difficile quand tu changes des habitudes dont il n'as pas l'habitude entre guillemets. Si c'est quelque chose qui répond à un besoin et qui fait sens, les gens vont l'adopter directement. Quand tu vois le succès de trucs comme Leboncoin et les plateformes comme ça qui sont sorties de

nulle part marchent très fort. Pourquoi ? Parce qu'en France tu as cette économie parallèle entre l'idée de faire du troc, soit de la vente de particulier à particulier. Ce que Leboncoin a réussi c'est de s'immiscer dans une des plus vieilles habitudes et de la rendre plus facile. Ça marche très fort parce que ça répond à un besoin originel qui était déjà là. Tant que tu arrives à trouver des moyens de répondre aux besoins client ça devrait très très bien marcher. Airbnb a fait la même chose en proposant des chambres pas chères du tout parce qu'avant t'avais pas trop le choix : Si t'étais 6 et que tu voulais pas prendre 3 chambres d'hôtel ou aller en auberge de jeunesse où t'étais pas tout seul, t'avais pas d'autre solution. C'est pour ça que ça marche aussi fort. Et l'interface est très très puissante, bien faite et super ludique. c'est la même chose pour Uber. Le changement marche très très bien du moment où il répond à un besoin. La simplicité est une grosse composante mais le prix reste déterminant. Par exemple, Airbnb est souvent plus compliqué que quand tu veux réserver une chambre dans hôtel ou une auberge de jeunesse, tu dois récupérer les clés, faire le ménage, mais comme c'est moins cher et plus *fun* comme expérience. (Silence).

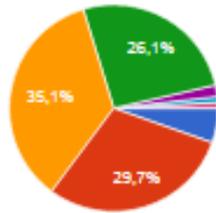
#### Conclusion de l'entretien

**Souhaitez-vous ajouter ou revenir sur quelque chose ?**

Je ne pense pas, j'ai déjà dit beaucoup de choses ! (Rires)

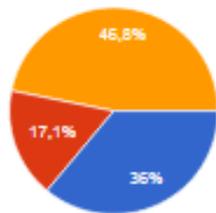
## Annexe F - Résultats bruts du questionnaire en ligne

À quelle catégorie votre établissement appartient-il?



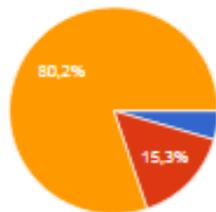
1*	6	5.4 %
2*	33	29.7 %
3*	39	35.1 %
4*	29	26.1 %
5*	2	1.8 %
Palace	1	0.9 %
Non classé	1	0.9 %

Appartenez-vous à une chaîne?



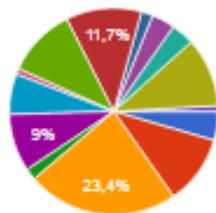
Volontaire	40	36 %
Intégrée	19	17.1 %
Non	52	46.8 %

Depuis combien de temps votre établissement existe-t-il?



Moins de 5 ans	5	4.5 %
Entre 5 et 15 ans	17	15.3 %
Plus de 15 ans	89	80.2 %

Dans quelle région vous situez-vous?



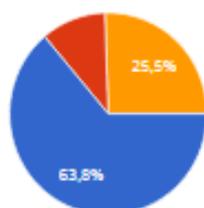
Alsace Champagne-Ardenne Lorraine	5	4.5 %
Aquitaine Limousin Poitou-Charentes	12	10.8 %
Auvergne-Rhône-Alpes	26	23.4 %
Bourgogne Franche-Comté	2	1.8 %
Bretagne	10	9 %
Centre Val-de-Loire	7	6.3 %
Corse	1	0.9 %
Île-de-France	12	10.8 %
Languedoc-Roussillon Midi-Pyrénées	13	11.7 %
Nord-Pas-de-Calais	2	1.8 %
Normandie	4	3.6 %
Pays-de-la-Loire	4	3.6 %
Provence-Alpes-Côte d'Azur	12	10.8 %
		0.9 %

Avez-vous déjà entendu parler du Big Data?



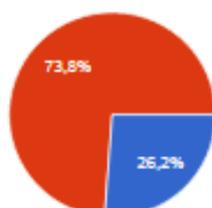
Oui	54	48.6 %
Non	57	51.4 %

Pourquoi ne les utilisez-vous pas?



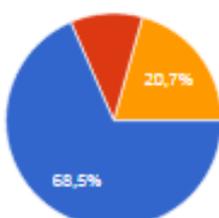
Manque de connaissances/compétences	30	63.8 %
Dépense trop lourde	5	10.6 %
Je n'en vois pas l'utilité	12	25.5 %

Utilisez-vous un logiciel CRM?



Oui	11	26.2 %
Non	31	73.8 %

Stockez-vous vos données sur des supports personnels ou dans un espace géré par un tiers (cloud)?



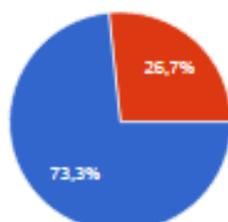
Support personnel	76	68.5 %
Cloud	12	10.8 %
Les deux	23	20.7 %

Votre établissement possède-t-il un site internet?



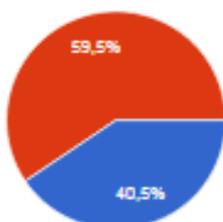
Oui	111	100 %
Non	0	0 %

Analysez-vous son trafic et son contenu?



Oui	55	73.3 %
Non	20	26.7 %

Votre établissement possède-t-il une application mobile?



Oui	45	40.5 %
Non	66	59.5 %

# Glossaire

**Business intelligence** : solutions IT apportant une aide à la décision aux professionnels avec, en bout de chaîne, des rapports et tableaux de bord de suivi des activités de l'entreprise à la fois analytiques et prospectifs.

**Cookies**: Mouchard électronique. Un cookie est un fichier envoyé par un serveur sur le disque dur d'un utilisateur pour stocker des informations sur sa navigation (par exemple, un panier d'achats électroniques). Les utilisateurs peuvent rejeter les cookies au risque de rendre certains sites ou pages inutilisables.

**Custom audience** : offre publicitaire créé par Facebook permettant par exemple de cibler des fichiers clients en important leurs adresses emails sur une base de données sécurisée pour réaliser un matching. Les publicités pourront alors toucher spécifiquement les utilisateurs de Facebook ayant déjà réalisé un achat.

**Data-as-a-Service (DaaS)**: technique consistant à faire payer un abonnement pour l'accès à un dépôt de données via une interface fournie par le fournisseur.

**Global Distribution System (GDS)**: Réseau informatique mondial de réservation utilisé dans les industries du tourisme (transport aérien, agences de voyage, hôtels, etc).

**Local Area Network (LAN)**: Système de communication permettant de relier quelques centaines d'ordinateurs et de périphériques dans un rayon de quelques kilomètres

**Log** (diminutif de logging) : pour désigner un fichier, ou tout autre dispositif, permettant de stocker un historique des événements attachés à un processus. Ces événements sont horodatés et ordonnés en fonction du temps. En clair, le log est un peu le « journal de bord » d'un système

**Monitoring** : Le monitoring d'un site web est la pratique par laquelle on surveille, la disponibilité, les temps de réponse et les temps de chargement d'un site web.

**Online Travel Agency (OTA)**: Ensemble d'intermédiaires de commerce en ligne qui consacrent des services ou des offres de voyage à caractère privé ou professionnel entre un hôtel-restaurant et un consommateur. Ces sites offrent la réservation du voyage, des avis de consommateurs, ou la combinaison des deux. Afin que les hôtels-restaurants puissent indiquer leurs prix et leurs disponibilités, les OTA mettent à disposition des moteurs de réservation. Les OTA facturent les établissements sous forme de pourcentage sur les ventes pour ce « service ».

**PRISM**: Programme de surveillance électronique de la NSA.

**Property Management System (PMS)**: Système global qui gère l'opérationnel hôtelier comme la réservation, la facturation, la gestion des débiteurs, la gestion de la relation client, la gestion des automates, la gestion des relations avec les autres services.

**Radio Frequency Identification (RFID)**: Technologie d'identification automatique qui utilise le rayonnement radiofréquence pour identifier les objets porteurs d'étiquettes lorsqu'ils passent à proximité d'un interrogateur.

**Reporting**: Famille d'outils de Business intelligence destinés à assurer la réalisation, la publication et la diffusion de rapports d'activité selon un format prédéterminé.

**Retargetting**: pratique publicitaire qui consiste le plus souvent à cibler un individu qui a visité un site Internet, mais pour lequel il n'y a pas eu achat ou transformation lors de cette visite.

**Scoring** : technique qui permet d'affecter un score à un client ou prospect. Le score obtenu traduit généralement la probabilité qu'un individu réponde à une sollicitation marketing ou appartienne à la cible recherchée. Il mesure donc l'appétence pour l'offre potentielle.

**Stratégie extensive** : Politique de marketing privilégiant la conquête de nouveaux clients.

**Structured Query Language (SQL)**: langage de requête structurée) est un langage informatique normalisé servant à exploiter des bases de données relationnelles.

**Systemes d'information**: ensemble organisé de ressources (matériels, logiciels, personnel, données et procédures) qui permet de regrouper, de classifier, de traiter et de diffuser de l'information sur un environnement donné.

**Technologies de l'Information de la Communication (TIC)** : Ensemble des techniques et des équipements informatiques permettant de communiquer à distance par voie électronique. (Dictionnaire Larousse en ligne)

**Up-sell** : pratique qui consiste à proposer un produit ou service légèrement supérieur et plus cher que celui auquel s'intéresse le prospect.

**Wide Area Network (WAN)**: Réseau généralement constitué de plusieurs sous-réseaux hétérogènes et s'étendant sur une région ou un pays entier.

## Table des figures

Figure A - Évolution des recherches associées au terme <i>Big Data</i> .....	13
Figure B - Système de de production et de captation des <i>Big Data</i> .....	19
Figure C - Déroulement du plan marketing.....	42
Figure D - Évolution du marketing sous l'impulsion d'Internet.....	43
Figure E - Tableau des variables clients et leur définition .....	44
Figure F - Répartition des hôtels et des chambres selon le type d'hôtel et leur classement en 2014.....	66
Figure G - Usage des données récoltées par les hôteliers.....	74
Figure H - Sites d'un hôtelier 1* indépendant et 5* de chaîne intégrée .....	74
Figure I - Éléments requis lors de l'installation d'une application .....	76

## Table des tableaux

Tableau 1 - Différences entre décision classique et décision <i>data</i> centrée.....	38
Tableau 2 - Parc hôtelier français depuis 2010.....	65
Tableau 3 - Classement des 10 chaînes intégrées présentes en France (en nombre d'hôtels) ..	68
Tableau 4 - Classement des chaînes volontaires en France .....	69
Tableau 5 - Répartition des répondants en fonction de leur catégorie et de leur statut.....	71

# Table des matières

Remerciements .....	5
Introduction générale .....	8
Partie I. Qu'est-ce que le <i>Big Data</i> ? .....	10
Introduction .....	11
Chapitre 1. Les concepts fondamentaux.....	12
1. Définition.....	12
2. Historique .....	12
2.1 L'origine du <i>Big Data</i> .....	12
2.2 Évolution.....	13
a) Des usages.....	13
b) Des supports.....	14
3. Dimensions du <i>Big Data</i> .....	15
3.1 V pour Volume .....	15
3.2 V pour Vitesse (ou vélocité) .....	16
3.3 V pour Variété .....	16
3.4 Les autres V .....	17
4. <i>Big Data</i> et concepts associés.....	17
4.1 L' <i>open data</i> .....	17
4.2 Le <i>crowdsourcing</i> .....	18
4.3 Le <i>cloud computing</i> .....	18
Chapitre 2. La collecte de données: comment et pourquoi? .....	19
1. Les technologies du <i>Big Data</i> .....	19
1.1 Le processus d'exploitation des données .....	19
1.2 Les <i>data center</i> et les <i>datavarhouse</i> .....	20
1.3 Les bases de données.....	20
1.4 Les logiciels et algorithmes .....	21
a) Le <i>data mining</i> .....	21
b) L'analyse prédictive.....	22
c) Le <i>machine learning</i> .....	23
d) La sémantique.....	23
1.5 Les principales solutions <i>Big data</i> .....	24
2. La nature des données.....	25

2.1	Les données structurées et non-structurées .....	25
2.2	Données personnelles et données anonymes .....	25
2.3	Les données référentielles, transactionnelles et opérationnelles .....	26
3.	Les sources de production de données.....	27
3.1	Les sites web .....	27
3.2	Les applications pour <i>smartphones</i> .....	27
3.3	La géolocalisation.....	27
3.4	Les réseaux sociaux.....	28
3.5	L'Internet des objets.....	28
4.	Les grands acteurs du <i>Big Data</i> .....	29
4.1	Les développeurs.....	29
4.2	Le GAFA .....	29
4.3	Les collectivités et acteurs territoriaux.....	30
Chapitre 3. Approches disciplinaires .....		31
1.	Approche économique.....	31
2.	Approche éthique et juridique .....	32
Conclusion .....		34
Partie II. <i>Big data</i> et marketing : quelles opportunités pour les entreprises ? .....		35
Introduction.....		36
Chapitre 1. Le <i>Big data</i> dans les entreprises .....		37
1.	Vers un changement des processus décisionnels .....	37
2.	Les secteurs impactés .....	38
3.	La mise en place d'une démarche <i>Big data</i> .....	39
3.1	Constat.....	39
3.2	L'émergence de nouveaux métiers .....	40
Chapitre 2. Les effets du <i>Big Data</i> dans la stratégie marketing .....		41
1.	La stratégie marketing .....	41
1.1	Définition.....	41
1.2	L'arrivée d'Internet et son impact sur le marketing.....	42
2.	Les apports dans les options stratégiques fondamentales .....	44
2.1	La segmentation .....	44
2.2	Le ciblage.....	45
2.3	Les sources de volume .....	46
2.4	Le positionnement.....	46

3.	Apports dans le mix marketing.....	47
3.1	La politique de produit.....	47
3.2	La politique de prix.....	48
3.3	La politique de distribution .....	48
3.4	La communication .....	49
Chapitre 3. Enjeux et limites.....		51
1.	Les enjeux .....	51
1.1	L'analyse des données comme facteur d'innovation .....	51
1.2	Se départir de l'influence des OTA .....	52
1.3	Le client avant tout .....	53
1.4	Créer des partenariats .....	54
2.	Les limites .....	55
2.1	Assurer la protection des données et leur bonne utilisation.....	55
2.1	L'ensemble de ces outils ne constituent pas une fin en soi .....	56
Conclusion .....		58
Partie III. Méthodologie et terrain d'application .....		59
Introduction .....		60
Chapitre 1. Méthodologie.....		61
1.	Déroulement de la recherche .....	61
2.	Rappel des objectifs.....	61
3.	Les outils utilisés .....	62
3.1	L'étude qualitative .....	62
a)	Objectif.....	62
b)	Profil de la personne interrogée.....	62
c)	Conduite de l'entretien.....	62
3.2	Analyse quantitative.....	63
a)	Objectif.....	63
b)	Échantillon et mode administration.....	63
4.	Les limites .....	64
Chapitre 2. Le parc hôtelier français.....		65
1.	Panorama du parc hôtelier français.....	65
2.	L'hôtellerie indépendante .....	67
3.	L'hôtellerie de chaîne .....	67
3.1	Les chaînes intégrées .....	67

3.2	Les chaînes volontaires .....	69
Chapitre 3.	Analyse des actions des acteurs .....	71
1.	Profil des répondants .....	71
2.	Constat face à l'utilisation des données .....	71
2.1	Niveau de connaissance du sujet .....	72
2.2	Utilisation des données au sein de l'entreprise .....	72
2.3	Outils utilisés et données recueillies .....	73
a)	Les sites Internet .....	74
b)	Les applications mobiles .....	76
c)	Les réseaux sociaux .....	77
2.4	Accès à d'autres données complémentaires .....	77
3.	Le stockage des données.....	77
Chapitre 4.	Perspectives futures.....	79
1.	Possibilités de stage en Master 2.....	79
2.	Travail de recherche envisagé .....	79
	Conclusion .....	80
	Conclusion générale.....	81
	Bibliographie.....	83
	Annexes .....	85
	Table des annexes .....	86
	Glossaire.....	106
	Table des figures .....	108
	Table des tableaux.....	108
	Table des matières.....	109

## RÉSUMÉ

L'avènement d'Internet a, depuis près d'une vingtaine d'années, transformé notre rapport à la technologie et à l'utilisation qui peut en être faite. Les progrès techniques continus en la matière sont à l'origine du flux massif de données auquel les entreprises doivent aujourd'hui faire face, communément désigné par le terme de *Big data*. Ce mémoire a pour objectif de mettre en avant les enjeux que représente cette nouvelle révolution numérique, et de son influence dans la transformation des stratégies marketing, en particulier pour les acteurs de l'hôtellerie française.

**Mots-clés :** *Big data*, hôtellerie, technologie, stratégie marketing

## SUMMARY

The advent of the Internet almost twenty years ago has transformed our relation with technology and the way we can use it. Technological advances in this field are the origin of the massive *data* feed more commonly known as "*Big data*". The aim of this memoire is to bring forward the challenges brought by this new technologic revolution and its influence on the marketing strategy particularly related to French hospitality.

**Key words:** *Big data*, hospitality, technology, marketing strategy