UCL Université catholique de Louvain

LOUVAIN School of Management

Louvain School of Management

Le paradoxe personnalisation – protection des données personnelles : Analyse des potentiels mécanismes d'atténuation des risques

Mémoire recherche réalisé par
Floriane Geerinckx

en vue de l'obtention du titre de
Master 120 crédits en sciences de gestion, à finalité spécialisée

Promoteur(s)
Virginie Bruneau
Valérie Swaen

Année académique 2017-2018

Avant toute chose, je tiens à remercier très sincèrement, Madame Virginie Bruneau, d'avoir accepté de m'encadrer dans cette recherche en tant que promotrice. Ses conseils, sa disponibilité et son intérêt pour le travail m'ont été très précieux dans la rédaction de ce mémoire.

Par ailleurs, je souhaite remercier Madame Chantale de Moerloose pour le suivi méthodologique offert et Mademoiselle Pauline Munten pour son aide lors de l'interprétation des résultats de mon étude.

Ma reconnaissance s'adresse également à toutes les personnes ayant pris le temps de répondre à mon questionnaire et l'ayant partagé.

Enfin, un grand merci à ma famille et mes amis pour leur soutien et encouragements tout au long de ce projet.

Table des matières

Liste des figures

INTRODUCTION

La rédaction de ce mémoire intervient dans le cadre de la seconde année du Master en sciences de gestion réalisé à la Louvain School of Management. Au fil de mes études, mon intérêt pour le marketing digital n'a cessé de croître. Après avoir choisi l'option e-business et bénéficié d'un stage chez Emakina, une agence digitale, il m'a semblé naturel de m'orienter vers le domaine du digital pour le sujet de mon mémoire.

Les notions telles que les agents de recommandations virtuels, le *Big Data*, les *AdBlockers* et le *one-to-one* marketing m'ayant toujours fortement intriguée, j'ai décidé de m'intéresser à la personnalisation en ligne dans le cadre de cette recherche. Par ailleurs, le *tracking* et la question de la protection des données personnelles étant au cœur des débats juridiques, techniques, économiques et sociologiques actuels sur les pratiques du web, il m'a semblé pertinent d'intégrer ce concept dans ma recherche. Cette idée est également venue de la constatation personnelle que l'enregistrement sur des sites web nécessite de plus en plus l'acceptation de cookies ou une connexion via des réseaux sociaux permettant aux entreprises d'avoir accès à un nombre considérable de données personnelles, du sentiment de vulnérabilité personnel émanant de ces pratiques ainsi que du fait de l'augmentation des fuites de données en ligne. Dès lors j'ai choisi d'aborder ce qui me semble être un paradoxe entre la personnalisation en ligne et la protection des données personnelles. Grâce à cette recherche, j'espère combler des lacunes existantes dans ce domaine de recherche et trouver des réponses à des questions personnelles telles que : Pourquoi désirons-nous autant recevoir des offres personnalisées ? Les services personnalisés peuvent-ils être performants alors que les individus sont réfractaires aux démarches intrusives? Comment les individus perçoivent-ils le respect de la vie privée dans le contexte de la personnalisation en ligne ? Ces deux concepts sont-ils conciliables ? Entre inquiétudes en matière de vie privée et bénéfices de personnalisation, quel est le plus important pour le consommateur lors de la prise de décision ? Acceptons-nous de divulguer nos informations personnelles sur base d'une décision réfléchie ou impulsive ? Sous quelles conditions acceptons-nous que notre information personnelle soit récoltée ? Que peuvent mettre en œuvre les entreprises afin de faire paraître la personnalisation comme étant moins personnelle et intrusive ?

En outre, ce sujet se révèle d'autant plus important qu'il est au cœur de la préoccupation des consommateurs et un challenge pour toute entreprise, ce qui a été souligné par de nombreux sondages effectués par des entreprises de conseil. En effet, la personnalisation en ligne est un des phénomènes mis en évidence depuis quelques années, nécessitée par et rendue possible grâce aux nouvelles technologies. Cette stratégie marketing part de la constatation que chaque individu a des besoins, intérêts et préférences différents qui évoluent constamment et que face à l'explosion de l'information en ligne, il est nécessaire de lui offrir de manière proactive une meilleure expérience en ligne, avec du contenu plus adapté à ses préférences. Jeff Bezos, le CEO d'Amazon a ainsi déclaré : « Si nous désirons avoir 20 millions de consommateurs, alors nous devons avoir 20 millions de 'magasins'» (Montgomery & Smith, 2009, p.134). Ceci démontre l'importance et la valeur des stratégies de personnalisation dans, entre autres, l'environnement de l'e-commerce. Par conséquent,

l'investissement dans la personnalisation web est une haute priorité stratégique pour toute entreprise et même une nécessité sur le plan de la concurrence (Chau, Ho, Ho & Yao, 2013 ; Chellappa & Sin, 2005).
Les chiffres ci-dessous montrent ainsi que les consommateurs sont demandeurs de ces expériences sur mesure. Accenture Interactive (2018, p.3) dévoile notamment que 91% des consommateurs sont plus susceptibles d'acheter auprès d'une marque qui les reconnaît et fournit des offres et recommandations pertinentes. En effet, selon une étude d'Accenture (Kuin, 2018, para. 2), 74% des consommateurs en ligne sont frustrés lorsqu'ils sont exposés à du contenu non-pertinent sur un site web, à une surcharge d'options ou par manque de personnalisation et d'anticipation de leurs besoins. Parmi ceux-ci, 40 % ont quitté la plateforme digital immédiatement pour faire un achat autre part et 61% ont tout simplement arrêté d'acheter ou de s'intéresser à au moins une entreprise l'année passée suite à une mauvaise expérience client (Kuin, 2018, para. 2). Salesforce Research (2016, p.9) mentionne également d'autres raisons de quitter le site web d'une entreprise telles que le fait d'être traité comme des nombres plutôt que des individus.

Par ailleurs, tout porte à croire que cette demande pour la personnalisation sera davantage présente dans le futur grâce à la généralisation de la technologie des objets. D'autant plus que 50% des individus n'ont plus de préférence entre une interaction avec des humains ou avec des technologies comme l'intelligence artificielle et que 44% utilisent déjà un type d'assistance virtuelle, dont 87% sont satisfaits (Accenture Strategy, 2017, p.15).

Cependant, malgré l'attractivité du marketing personnalisé, 56% des individus se disent être « préoccupés » à « très préoccupés » à propos de la manière dont les entreprises gèrent et utilisent leurs données personnelles, considérant quelque fois le traitement de l'information comme une intrusion dans leur intimité (KPMG, 2017, para. 2). En effet, cela vient du fait qu'une condition nécessaire à la personnalisation en ligne est la compréhension approfondie des comportements des consommateurs par l'entreprise, rendue possible grâce à la capitalisation des données clients divulguées sur internet. Toutes ces traces laissées par les consommateurs sur internet sont désormais la première source de valeur pour les entreprises mais suscitent simultanément de nombreuses inquiétudes auprès des consommateurs et inévitablement une perte de leur vie privée. Sur base de cette constatation, nous pourrions croire qu'en fin de compte les individus ne partageront plus leurs données avec les entreprises et que la personnalisation ne sera, par conséquent, pas aussi fructueuse que nous l'avions décrite. Paradoxalement, les études montrent aussi que, malgré ces préoccupations, les consommateurs finissent tout de même par partager de l'information et donc de bénéficier de la personnalisation.
C'est dans le cadre de ce paradoxe entre la personnalisation et protection des données personnelles que se situe notre travail.

Ainsi, sur base de ces constats, il semble intéressant d'aborder la problématique initiale suivante :

Dans quelle mesure des mécanismes d'atténuation des risques de confidentialité, tels que la confiance, la transparence et le contrôle des données, ont une influence sur l'intention du consommateur de partager des données personnelles, en vue de bénéficier de la personnalisation en ligne ?

Ce mémoire vise (1) à comprendre les raisons pour lesquelles les consommateurs acceptent ou rejettent la participation à la personnalisation en ligne, en fonction des craintes relatives à la protection des données. Pour cela, nous analyserons (2) quels facteurs ont une incidence sur l'intention du consommateur de divulguer des données personnelles en ligne. L'objectif final de notre question de recherche est (3) d'évaluer par quels moyens les entreprises peuvent diminuer les risques perçus des consommateurs en matière de vie privée associés à la divulgation des données dans un contexte de personnalisation, afin d'améliorer les pratiques marketing des entreprises.

Pour ce faire, la méthodologie suivie au cours de cette étude consiste en une analyse théorique sur base d'articles académiques et de rapports de recherches d'entreprises de conseil et marketing, et une partie pratique au moyen d'une étude quantitative. Il a été décidé de prendre la perspective du consommateur et de se concentrer principalement sur les services personnalisés sur les sites web car la personnalisation dans les publicités, e-mails personnalisés, sur les réseaux sociaux ou hors ligne ont déjà fait l'objet de nombreuses études. Par ailleurs, il est important de préciser que nous analyserons la volonté de participation des consommateurs aux services personnalisés en mesurant leur intention de communiquer de l'information personnelle en ligne. La raison de ce choix est que cette intention de divulgation est la condition nécessaire à la personnalisation et insinue par conséquent que les consommateurs souhaitent recevoir du contenu sur mesure en retour de leurs données. Nous souhaitons également mentionner que, dans cette étude, les termes confidentialité, protection ou respect de la vie privée seront utilisés comme des synonymes pour faire référence au concept de « *privacy* » en anglais et que lorsque nous mentionneront le concept de « risques perçus », il s'agira des risques perçus en matière de vie privée dans le cadre de la personnalisation.

En ce qui concerne la structure, ce mémoire est organisé comme suit.
Dans la première partie, nous partirons de la littérature afin de proposer une mise en contexte global. Nous explorerons dans un premier temps l'évolution du consommateur et du marketing à l'ère du numérique afin de comprendre les origines de la personnalisation. Dans un deuxième temps, nous développerons en détail le concept de personnalisation et de la protection de la vie privée. Enfin, dans le dernier chapitre de cette partie, le sujet se précise car nous nous intéresserons plus particulièrement à la combinaison de ces deux concepts principaux et aux facteurs pouvant impacter la relation entre les risques de confidentialité et la participation à la personnalisation. Tout ceci, dans le but d'avoir une base solide sur laquelle appuyer l'argumentation de notre étude et ainsi développer une série d'hypothèses et notre modèle de recherche.
La deuxième partie de ce travail développe l'étude quantitative permettant de tester le modèle de recherche établi et, ainsi, d'éventuellement déceler une combinaison d'éléments diminuant les risques perçus et encourageant la divulgation des données. Cette partie présentera d'abord les procédures de la recherche, suivis des résultats obtenus et l'interprétation de ceux-ci. Nous la conclurons en développant des recommandations managériales, un recul critique, les limites de notre étude et des suggestions pour des recherches futures.

PARTIE I : REVUE DE LA LITTÉRATURE

Cette partie théorique est divisée en 4 chapitres. Nous passerons d'abord en revue la documentation sur l'évolution du consommateur et du marketing, sur la personnalisation ainsi que sur le concept de la protection de la vie privée. Ensuite, nous combinerons ces deux thèmes pour discuter des compromis auxquels les utilisateurs doivent faire face lorsqu'ils souhaitent divulguer de l'information afin de bénéficier de la personnalisation en ligne.

1. Evolution du consommateur et du marketing

Avant toute chose, il est nécessaire d'exposer le contexte définissant l'apparition du phénomène étudié. Ainsi, dans ce chapitre nous présenterons l'évolution du consommateur et du marketing.

L'essor d'internet et de la popularité du *Big Data* ainsi que les progrès dans les technologies numériques et l'adoption généralisée des réseaux sociaux ces dernières années, ont transformé tant les individus que les opérations et stratégies marketing des entreprises (De Filippi, 2016). Comme pour toute nouveauté marketing, il faut prendre le consommateur comme point de départ et partir du constat qu'il a changé, notamment en termes d'habitudes et attentes.

Ainsi, on peut observer que le consommateur est aujourd'hui **ultra-connecté**, actif et omniprésent sur internet. Il interagit davantage avec les marques et possède un grand réseau avec d'autres consommateurs (Hennig-Thurau, Malthouse, Friege, Gensler, Lobschat, Rangaswamy, & Skiera, 2010). Il révèle beaucoup sur lui-même en ligne et toutes ses activités sur internet peuvent être observées par les entreprises. Cette connectivité constante fait partie du quotidien car les individus gèrent tous les aspects de la communication quotidienne principalement via leur smartphone (Saleforce Research, 2016).

Ensuite, il **ne rentre plus dans les cases telles que définies auparavant,** c'est-à-dire sur base de critères sociodémographiques uniquement, car il ne veut pas être catalogué avec une étiquette spécifique (Capgemini Consulting, 2017). En outre, mieux informé et plus exigeant, il souhaite **se sentir mis en avant et s'attend à ce qu'on le reconnaisse**, changeant par conséquent la relation existante avec les entreprises. Il tient désormais les rennes des stratégies commerciales et de l'innovation et demande du sur-mesure tel que des expériences personnalisées, s'attendant à ce que les entreprises soient davantage à l'écoute de ses attentes, préférences et besoins (SAS, 2015). Ce sont les consommateurs qui déterminent en quelque sorte le futur et succès des entreprises, celles qui dirigeront l'ère du digital et celles qui ne fonctionneront pas.

Par ailleurs, les structures web étant larges et compliquées et avec la croissance de l'e-commerce, le consommateur **se retrouve face à une multitude d'informations.** Le choix est, par conséquent, de plus en plus difficile pour lui (Komiak & Benbasat, 2006). Dès lors, la panoplie de messages médiatiques auxquels il est exposé ainsi que la grande quantité d'informations à gérer sur les réseaux sociaux, a rendu le consommateur sans pitié : si le contenu n'est pas pertinent, si c'est du « baratin » commercial, il perd son intérêt pour son émetteur et l'abandonnera (Kuin, 2018). Le consommateur a donc besoin de systèmes intelligents afin de filtrer l'information et d'aide dans sa prise de décision (Chang, Changchien & Huang, 2006). Dans cette optique, il semble que les consommateurs sont convaincus de la nécessité

d'innover et adoptent une attitude davantage positive envers les technologies. Ils exigent également que les sociétés les utilisent, entre autres l'intelligence artificielle comme les *chatbots*, afin de fournir des échanges plus proactifs et intelligents (Salesforce Research, 2016). Bien que plus réceptif à ces technologies, il est important de noter que plus de 40% des individus craignent encore que les services intelligents n'en apprennent trop sur eux et nombreux sont ceux qui sont réticents face à la reconnaissance faciale (Accenture Strategy, 2017, p.17). Trouver un juste équilibre est primordial.

Nous sommes donc entrés dans une ère où les consommateurs dictent ce qu'ils veulent, où et quand. Ils exigent, et non pas uniquement désirent, que les entreprises connaissent leurs préférences et comprennent leurs besoins (Deloitte, 2015). La nécessité d'anticiper leurs besoins est aussi plus évidente que jamais. Dès lors, la **tendance est aujourd'hui à l'omnicanalité, l'immédiateté et à la personnalisation** de l'ensemble du parcours digital du consommateur.

Toutes ces évolutions du comportement du consommateur ont rendu obsolètes les techniques marketing antérieures, notamment de segmentation traditionnelle. Les spécialistes du marketing ont bien compris qu'ils devraient désormais répondre aux attentes et besoins des consommateurs à l'échelle de l'individu, lui redonner sa place en sortant de la logique de masse, avec comme objectif la satisfaction du client et non pas les bénéfices de l'entreprise. L'époque du marketing (segmentation, publicité, communication) de masse, du marketing transactionnel et centré sur le produit est donc révolue, laissant place au **marketing relationnel et au marketing *one-to-one*** (Capgemini Consulting, 2017 ; Baek & Morimoto, 2012). Le marketing relationnel a été définit par Parvatiyard et Sheth (cités dans Noble & Phillips, 2004, p. 289) comme étant « un processus continu d'engagement dans des activités coopératives et collaboratives avec des consommateurs immédiats et finaux afin de créer et enrichir la valeur économique mutuelle, à un coût réduit ». Des exemples de stratégies de marketing relationnel sont les programmes de fidélité, les offres de remise et les offres personnalisées via e-mail (Noble & Phillips, 2004).
Quant au marketing *one-to-one*, c'est le fait d'adapter ses stratégies envers un consommateur individuel. Il s'agit d'individualiser toutes les actions marketing et cibler plus finement. C'est la micro-segmentation poussée à l'extrême, où le segment est de taille 1 (Arora, Dreze, Ghose, Hess, Iyengar, Jing, Joshi, Kumar, Lurie, Neslin, Sajeesh, Su, Syam, Thomas & Zhang, 2008). En effet, il est possible d'affiner les segments traditionnels en intégrant davantage de critères. Alors qu'initialement seuls des données sociodémographiques et géographiques étaient utilisées, aujourd'hui il est nécessaire d'intégrer des critères comportementaux et psycho-graphiques (Capgemini Consulting, 2017). Enfin, bien que ces dernières soient subjectives et donc plus difficiles à mesurer, l'apport des nouvelles technologies et le *Big Data* a rendu cela possible.

Dès lors, il est évident que **les entreprises devront s'adapter** à ces nouveaux comportements et à ce changement de paradigme marketing, mettant les évolutions des préférences et attentes du consommateur au cœur de leur business. En outre, adopter les stratégies marketing citées ci-dessus, ciblant les consommateurs individuels et aidant ceux-ci à accomplir efficacement leurs objectifs, est susceptible d'être un élément clé du succès d'une entreprise (Lee & Cranage, 2011). Dans ce mémoire, nous allons nous concentrer sur la technique de la personnalisation, partie intégrante du marketing relationnel et marketing *one-to-one*.

2. Personnalisation

La personnalisation étant un des éléments clés de notre problématique, nous nous devons d'analyser ce concept en profondeur. Nous commencerons par définir ce concept, soulignerons ensuite l'importance des données clients et de la technologie, présenterons les différents types de personnalisation ainsi que son fonctionnement et ses effets. Nous terminerons avec des exemples concrets.

2.1. Définition du concept

La personnalisation n'est certainement pas un nouveau phénomène car l'idée d'offrir des produits et services personnalisés existait déjà bien avant l'internet. Même avant la révolution industrielle, dans l'absence d'usines et de la production de masse, tout était personnalisé. Cependant, cette technique a repris de l'ampleur dans les années 1980 lorsque l'individu est à nouveau devenu central dans les stratégies de marketing et production et a ensuite pris une **nouvelle tournure avec le développement des technologies de traitement de l'information** (Deloitte, 2015).

En effet, grâce aux possibilités offertes par le web et en réponse aux attentes des consommateurs connectés, la personnalisation est dorénavant possible sur cette plateforme également, on peut même parler d'une expérience client multicanale personnalisée (Li & Unger, 2012). La personnalisation est aujourd'hui plus précise, enrichie par des données récoltées en ligne grâce à des technologies sophistiquées. Selon Salonen et Karjaluoto (2016), les conditions sont excellentes pour que la personnalisation en ligne prospère. Cette prolifération de personnalisation en ligne est aussi due au fait qu'un nombre croissant de consommateurs achètent en ligne. Des recherches ont d'ailleurs montré que les techniques de personnalisation sont plus efficaces en ligne que dans un environnement hors ligne, grâce à la facilité de collecter un grand nombre de données du consommateur et d'offrir des recommandations en temps réel (Adomavicius & Tuzhilin, 2001). Etant un concept qui évolue, on remarque également qu'on est passé d'un marketing personnalisé, où il ne s'agissait que de messages et publicités sur mesures, à des expériences entièrement personnalisées (Accenture Interactive, 2018).

En plus d'être un concept en constante évolution, la personnalisation est un concept multidimensionnel et signifie quelque chose de différent pour chaque entreprise et acteur de la chaîne de valeur. Vesanen (2007) indique notamment que la personnalisation est un concept large comprenant tant l'exécution du marketing personnalisé et les technologies associées, que la publication du marketing personnalisé, la valeur pour le client et la valeur pour l'entreprise. Par conséquent, trouver une définition optimale n'est pas facile. Nous avons sélectionné quelques **définitions** complémentaires ci-dessous.

Tam et Ho (2005, p. 271) définissent la personnalisation comme étant « une stratégie de marketing orientée consommateur et visant à offrir le bon contenu, à la bonne personne, dans le bon format et au bon moment ». Imhoff, Loftis et Geiger (cités dans Baek & Morimoto, 2012, p.64) complètent cette définition en mentionnant qu'il s'agit de « la capacité de l'entreprise de reconnaître et traiter ses consommateurs comme des individus à part entière à travers des messages personnels, des bannières publicitaires ciblées, des offres spéciales ou autres transactions personnelles ». Roberts (cité dans Baek & Morimoto, 2012) précise également que cette communication personnalisée pour un individu en particulier se fait sur base de données et préférences exprimées de façon explicite ou implicite. De manière plus large,

Lavie et al. (cités dans Ho & Bodoff, 2014, p. 498) définissent la personnalisation en ligne comme étant « un processus automatisé qui identifie un utilisateur, collecte son schéma de navigation, analyse les préférences connues d'utilisateurs similaires et évalue ses préférences spécifiques afin d'adapter le contenu pour chaque utilisateur ». Enfin, Chellappa et Shivendu (2007) insistent sur le fait que la personnalisation en ligne est généralement offerte gratuitement et Murthi et Sarkar (2003) soulignent que c'est un processus itératif.

En reprenant tous les mots clés qui ressortent de ces définitions, nous pouvons dire que la personnalisation est une stratégie de marketing axée sur le consommateur individuel ainsi qu'un processus et la capacité de l'entreprise à identifier et exploiter les données disponibles des individus afin de découvrir leurs préférences et leur offrir sans frais, des produits, services, contenus et expériences adaptés, qui rempliront aux mieux les attentes et comportements de ceux-ci, notamment dans le bon format et au bon moment.

Dans ce travail, lorsque nous parlerons de personnalisation, il s'agira de la personnalisation en ligne ou *web personnalisation*, un sous-thème de la recherche de la personnalisation (Tuzhilin cité dans Salonen & Karjaluoto, 2016). La taxonomie entre ces deux concepts n'est pas très claire étant donné que la personnalisation est aujourd'hui couramment considérée comme étant liée à l'internet (Salonen & Karjaluoto, 2016). Nous devons donc rajouter à notre définition personnelle que le processus est automatisé et concerne les canaux digitaux. Etant donné que ces termes sont interchangeables, par facilité, nous utiliserons régulièrement la notion de personnalisation dans ce travail, en sous-entendant que cela fait référence au processus dans un environnement web.

2.2. Personnalisation versus customisation

Il est important de ne pas confondre personnalisation et customisation. Bien que les préférences des consommateurs et l'individualisation du contenu soient au cœur de ces deux concepts, la personnalisation est réalisée par l'entreprise et est automatisée, tandis que la customisation est initiée par le consommateur (Montgomery & Smith, 2009). La force de la personnalisation est donc qu'elle nécessite un minimum d'effort de la part du consommateur, contrairement à la customisation où le consommateur doit spécifier de manière proactive les éléments de son marketing mix souhaités et donc investir du temps et des ressources cognitives afin d'adapter le site web (Arora et al., 2008). Prenons l'exemple de la fonctionnalité MyYahoo sur Yahoo.com. Celle-ci permet à l'utilisateur de spécifier les éléments qu'il veut voir apparaître sur sa page d'accueil comme les prévisions météorologiques, les cours des actions, les préférences en termes d'actualité, le thème d'affichage etc. Un service de personnalisation offrirait cela sans l'implication de l'utilisateur, sur base de ses données récoltées et analysées.

Cette distinction entre customisation et personnalisation est importante étant donné qu'elle met la charge sur le spécialiste en marketing dans le cas de la personnalisation et que c'est donc à lui d'anticiper ce que le consommateur veut (Chung, Wedel & Rust, 2016). Ceci est complexe car il est nécessaire de faire des déductions à partir d'actions qui ne sont pas nécessairement liées à la personnalisation attendue. Heureusement, aujourd'hui il y a une multitude de sources d'information disponibles sur internet pour cet apprentissage passif et de systèmes adaptatifs (Montgomery & Smith,

2009). Pour la customisation, c'est évidemment plus simple mais le problème est que l'individu ne sait pas toujours quoi répondre ou n'a pas le temps d'adapter tout lui-même.
Nous insistons sur cette différence car nous verrons par la suite que le fait que ce soit l'entreprise qui ait le contrôle dans le cas de la personnalisation, et non pas le consommateur, peut provoquer des inquiétudes en matière de protection de la vie privée auprès de ce dernier.

2.3. Les données et les technologies au cœur de la personnalisation

A partir des définitions de ce concept, nous pouvons déduire que la personnalisation dépend impérativement de deux facteurs : la **volonté du consommateur de partager des données personnelles** et la **capacité de l'entreprise à obtenir et extraire cette information** (Xu et al., 2011).

Premièrement, afin de pouvoir offrir de la personnalisation et plus de valeur aux clients, les entreprises doivent en effet tirer parti de la moindre source de données personnelles disponible. Ces données représentent « toute information relative à une personne physique identifiée ou identifiable par un numéro d'identification ou un élément qui lui est propre » (Portes, 2018, p.197). Celles-ci sont nombreuses et variées.

Alors que dans un contexte hors-ligne, les données sont limitées à l'information que l'on peut trouver au point de vente, lors d'un contact directe ou suite à une transaction financière, les entreprises en ligne peuvent récolter une multitude d'informations à propos des activités de navigation des individus sur leur site web, même s'il n'y a pas de transaction financière qui ait lieu. Ainsi, Chellappa et Sin (2005) affirment que toute l'information que l'on récolte aujourd'hui grâce à la technologie, est comparable à tout ce qu'on pourrait acquérir si l'on suivait chaque client qui entre dans un magasin avec une caméra, l'observait à travers tous les rayons et analysait chaque produit qu'il touche. Ensuite, même dans l'absence de données à propos d'un consommateur spécifique, un produit ou service peut être proposé et adapté en fonction de la moyenne des autres consommateurs ou d'un segment de consommateurs (Chellappa & Sin, 2005). Capgemini Consulting (2017) affirme d'ailleurs qu'aujourd'hui, il est possible d'en savoir presque autant sur des prospects que sur des clients grâce aux données externes. Dans cette optique de **maximisation des données récoltées**, on constate également que de plus en plus de sites web demandent à un nouvel utilisateur de se connecter via un réseau social car cela donne accès à de nombreuses données et nécessite peu d'effort de la part du consommateur. Par ailleurs, les objets connectés, aujourd'hui fortement intégrés dans la vie quotidienne des individus avec le smartphone considéré comme l'outil préféré pour recevoir et transmettre de l'information, forment une grande source d'information extrêmement précieuse pour les entreprises. Ces données détaillées sont notamment faciles d'accès et plus représentatives de la vraie manière dont les individus utilisent les services digitaux, étant donné qu'ils les emmènent partout avec eux (Sutanto, Palme, Tan & Phang, 2013 ; Park, Matic, Garg & Oliver, 2017 ; Salonen & Karjaluoto, 2016). Ainsi, Heath (2017) va jusqu'à dire que **ces données des consommateurs sont l'or noir du 21ème siècle.**

Enfin, selon une étude d'Emarsys, Forrester, Researchscape et Evergage, 55% des spécialistes marketing utilisent l'information quant aux pages visitées et contenus vus pour la personnalisation, 47% utilisent la localisation, 46% le comportement sur le site, 43% les données démographiques et

39% les types d'outils utilisés (Wander, 2018, para. 2). Cependant, il semble que les consommateurs ne divulgueront ces données que s'ils en voient l'intérêt et /ou font confiance à l'entreprise (nous reviendrons sur ce point plus tard).

Deuxièmement, Lee et Cranage (2011) insistent sur le fait qu'un des enjeux majeurs pour les entreprises est **d'exploiter correctement cette prolifération des données**. Heureusement, en parallèle, les **technologies** que les sociétés peuvent détenir pour élaborer leur marketing se sont enrichies. Le progrès des technologies de l'information facilitent les tâches des entreprises pour collecter, conserver, analyser et utiliser les grandes masses de données fournies, volontairement ou involontairement, par les internautes (SAS, 2015). Ces technologies permettent également de collecter et traiter ces données plus rapidement afin d'offrir des solutions en temps réel. Ainsi, les entreprises peuvent désormais analyser le type d'appareil utilisé par le consommateur, les temps de visites sur chaque page web, le délai d'ouverture d'un e-mail, les mots-clés recherchés et même la géolocalisation exacte, les habitudes alimentaires, de l'information à propos de la santé, ou encore le rythme de sommeil et les routines d'entraînements des individus (SAS, 2015). Certains experts vont jusqu'à analyser la vitesse de passage de la souris sur une page web, ce qui révèle de l'information sur la réaction d'un individu à du contenu, comme son intérêt ou de l'agacement. Le fonctionnement de ces technologies sera expliqué plus loin dans ce travail.

Il est donc clair que les technologies de l'information et, en premier, lieu l'information qu'elles récoltent par rapport au consommateur, sont des facteurs clés de succès pour toute initiative de personnalisation d'une entreprise opérant en ligne (Awad & Krishnan, 2006). Dans la section suivante, nous exposerons les différentes exploitations possibles de ces données.

2.4. Typologie

Selon la littérature, tout comme pour la définition, il n'existe pas une seule manière de catégoriser la personnalisation étant donné qu'elle peut être appliquée sur de nombreux supports et de manière variée. Certains font une distinction en fonction des données utilisées, d'autres en fonction du message communiqué, du canal de diffusion, du design etc. Sur base de la cassification de Chellappa et Sin (2005), Khelladi, Castellano et Limongi (2014), Accenture (Kuin, 2018), Deloitte (2015) et Grenouilleau (2015), nous résumons ci-dessous les différents types de personnalisation en ligne.

Tout d'abord, la personnalisation en fonction des données récoltées comprend deux grandes catégories : **la personnalisation basée sur les caractéristiques individuelles spécifiques et statiques** et la **personnalisation comportementale**. Cette première inclut des données telles que le nom, le sexe, l'adresse, le style de vie, les préférences exprimées du consommateur. Ces données sont généralement communiquées explicitement par le consommateur. Le **marketing personnalisé nominal,** où l'on intègre le prénom ou nom de la personne visée dans le contenu, en fait donc partie et est la forme la plus basique de la personnalisation. C'est une approche entre temps standardisée et utilisée principalement dans les e-mails commerciaux, les newsletters, ainsi que sur Facebook et sur les pages d'accueils des sites internet.

Ensuite, concernant la personnalisation comportementale, il y a en premier lieu le **marketing personnalisé transactionnel** qui se base sur les données liées au comportement d'achat de l'utilisateur, tel que l'historique d'achat. La personnalisation peut également se baser sur le contenu que le consommateur a vu au préalable, son parcours de navigation, le lieu et l'heure de ses visites, appelée **la personnalisation liée au contenu ou contextuelle**. Selon cette catégorie, quelqu'un ayant lu des articles à propos de la Colombie, recevra par la suite des suggestions de packs voyages pour ce pays par exemple. Contrairement au marketing personnalisé nominal, on ne perçoit pas nécessairement que le contenu a été adapté de manière individuelle dans ces deux dernières sortes de personnalisation. Celles-ci utilisent généralement la technique du filtrage collaboratif (expliquée plus loin dans ce mémoire) et se présentent sous forme de recommandations. Troisièmement, **la personnalisation sur base de la localisation géographique.** Ceci consiste à modifier du contenu en fonction de la localisation géographique du consommateur et apparait souvent par le biais de notifications sur un dispositif portable. Les données utilisées dans ce cas sont évolutives et temporaires. Des exemples en sont la redirection automatiquement vers la page web de notre pays, l'adaptation des devises ou les notifications d'offres spéciales d'un magasin à proximité et tous cela en temps réel. Enfin, il y a **la personnalisation causée par l'utilisateur** qui a lieu lorsqu'une plateforme digitale répond aux actions d'un visiteur sur celle-ci. Elle se base sur les sources entrantes. Un nouveau visiteur d'un site web sera par exemple invité à souscrire aux newsletters, alors qu'un fidèle utilisateur ne verra pas ce message.

En termes **d'affichage** des messages ou de contenus personnalisés, Evergage (Sweet, 2016) présente les différentes possibilités de la manière suivante. Des exemples sont présentés à l'annexe n°1.

- *Inline content.* Cela fait référence à l'ajout de sections de contenu sur une page web, en l'intégrant plutôt que d'ajouter un message au-dessus de la page. Généralement, les consommateurs ne vont même pas réaliser qu'il y a de la personnalisation, comme c'est régulièrement le cas pour les personnalisations transactionnelle et contextuelle citées ci-dessus. Un exemple est le fait d'insérer du nouveau contenu sur mesure, directement sur la page du panier d'achat afin d'encourager les consommateurs à acheter plus pour bénéficier de la livraison gratuite. Ceux ayant déjà suffisamment dans leur panier pour bénéficier de la livraison auront un autre contenu ou tout simplement rien, sans que cela impacte le design du site. Il s'agit donc d'une manière harmonieuse de présenter de l'information aux consommateurs dans le contexte et design de la page.
- *Inpage Edits.* Ceci permet de modifier ou supprimer du contenu existant pour un segment ou individu spécifique, plutôt que d'ajouter toute une nouvelle section, contrairement au *inline content*. Cela permet de faire des changements subtiles mais percutant afin d'attirer différents visiteurs sur le site. Le fait de montrer une page avec une image, du texte et des *call-to-actions* liés aux voyages en famille pour le segment « famille » sur un site de vacances est un exemple de ce type de personnalisation, où on propose des vacances liées à la saison en cours. C'est également une technique offrant des expériences harmonieuses, tandis que les deux options suivantes sont de natures plus interruptives mais tout de même efficaces.
- *Les infobars.* Celles-ci correspondent aux messages persistants qui apparaissent généralement en haut d'une page web, en tant que *sticky message*, mais qui peuvent être fermés. Cela sert, par exemple, de rappel ou d'alerte comme en cas de soldes ou d'une offre de livraison gratuite. Ces barres peuvent être adaptées aux caractéristiques ou au comportement du client.

- *Callouts.* Ce sont des messages attirant l'attention sur une certaine zone ou fonctionnalité en particulier. Cela peut être un message afin de souligner les nouvelles fonctionnalités d'un produit ou pour demander l'avis de l'utilisateur concernant le produit. On peut alors définir que, par exemple, uniquement les nouveaux visiteurs voient ce message, ou uniquement lorsqu'ils passent la souris à cet endroit-là. Le but étant que cela apporte de la valeur au client et non pas que cela le gène.
- *Les messages pop-ups.* Ces messages qui apparaissent soudainement au-dessus du contenu sont très efficaces pour attirer l'attention mais également très envahissants. Par conséquent, ils sont à utiliser avec modération et stratégiquement. Ils sont souvent utilisés afin d'attirer les prospects et améliorer les taux de conversions en demandant l'adresse e-mail ou rappeler de passer à l'achat avant de quitter la page. A nouveau, on peut définir le timing ainsi que l'affichage seulement pour un certain type de clients, adapter les effets d'apparition etc.

Tous ces messages personnalisés peuvent utiliser les données de manière **explicite ou implicite**. En d'autres termes, dans le premier cas, le consommateur retrouvera clairement ses données dans le message sur mesure et saura donc quelles données sont détenues et utilisées par l'entreprise, ce que certains consommateurs n'apprécient pas, estimant que c'est bizarre, voire glauque. Tandis que dans le deuxième cas, la recommandation coïncidera également avec ses préférences mais sans que l'information personnelle soit communiquée expressément (Karwatzki, Dytynko, Trenz & Veit, 2017). Dans le même genre, on peut afficher du **contenu personnalisé informatif et non-informatif** (Sahni, Wheeler, & Chintagunta, 2018). Le premier cas est une adaptation basique du message où un contenu non-informatif par rapport au produit, comme un nom, est mentionné. Le contenu est néanmoins pertinent pour la cible et a donc de la valeur. Dijkstra ainsi que Hawkins et al. (cités dans Sahni, Wheeler & Chintagunta, 2018) mentionnent d'ailleurs que les tactiques de personnalisation ne doivent pas nécessairement fournir de contenu persuasif, être liées au comportement attendu ou fournir de nouvelles informations à propos de l'individu pour être performantes. L'autre forme est de modifier le contenu informatif du message, personnaliser les arguments par rapport au produit afin de persuader le destinataire.

En résumé, de nombreuses possibilités existent au niveau du contenu et de la forme, le mieux étant de combiner plusieurs approches. Par ailleurs, il est important d'en vérifier la performance en faisant des tests d'algorithmes afin de savoir précisément quelles recommandations doivent apparaître sur quelle page, quel message, quelle quantité etc. Connaître les attentes d'un client spécifique est primordial pour le bon affichage car certains apprécient, par exemple, de voir leur nom dans le contexte d'un e-mail mais il semble que ce n'est pas le cas pour d'autres canaux comme les bannières publicitaires (Sahni, Wheeler & Chintagunta, 2018).

Enfin, concernant le **canal ou l'outil de diffusion**, on retrouve la personnalisation généralement sous forme de recommandations sur des sites web, des pages d'accueils et fils d'actualités personnalisés comme sur les réseaux sociaux, sous forme de publicités personnalisées, de recherches par mots clés adaptées, d'e-mails personnalisés, d'expériences sur mesure via les *chatbots* par exemple ou via des applications mobiles offrant de la personnalisation en temps réel sous forme de messages *push*. Une étude de Evergage (2018, p.6,14) a notamment montré que les marketers utilisent la personnalisation essentiellement dans les canaux suivant : des campagnes e-mail (77%), sur sites web (52%), sites

mobiles (28%), applications web (31%) et applications mobiles (24%) sous forme de bannières (45%), messages call-out (40%), contenus intégrés (38%), questionnaires (36%), modifications du contenu existant (30%), pop-ups (29%), barres d'information (25%). Les deux types les plus fréquemment utilisés sont donc les publicités et services personnalisés sur des sites web (Awad & Krishnan, 2006).
Dans ce travail, nous nous concentrerons principalement sur les services personnalisés, c'est-à-dire les sites web dont la présentation et le contenu s'adaptent aux caractéristiques et préférences des utilisateurs en fonction des données récoltées. Des exemples sont la langue ou la devise automatiquement adaptée, des publicités appropriées, un message de salutation personnalisé ou des systèmes de recommandations sur un site web.

2.4.1. Systèmes de recommandations

La recommandation est l'application la plus utilisée parmi toutes celles disponibles afin de personnaliser un site web (Ho & Bodoff, 2014). L'agent de recommandation, un logiciel informatique de conseil personnalisé fonctionnant comme un agent de vente virtuel, est capable d'identifier, sélectionner et présenter automatiquement des choix appropriés, à partir d'une grande quantité de contenus ou produits alternatifs, basés sur des données récoltées de tous les utilisateurs. Son objectif premier est d'aider le consommateur à prendre des décisions, dans le sens où ces systèmes atténuent les difficultés associées au choix entre des alternatives (Hauble & Trifts, cités dans Lee & Kwon, 2008 ; Zanker et al., cités dans Ho & Bodoff, 2014).

En pratique, l'agent suggèrera une petite série d'éléments d'information (du contenu ou des produits en fonction du type de site web), sous forme de recommandations qui coïncident avec les préférences de l'utilisateur et ce, sans qu'il ne se soit nécessairement renseigné à ce sujet auparavant. Ceci est généralement effectué grâce à l'outil de filtrage collaboratif. Dès lors, les challenges de ces systèmes de recommandations sont le grand nombre de choix alternatifs, les données manquantes, l'hétérogénéité de l'usage et la flexibilité (Montgomery & Smith, 2009). Les recommandations qui n'utilisent pas le filtrage collaboratif utilisent généralement un questionnaire et une sélection d'options limitées au début du parcours d'achat ou de l'inscription du consommateur afin de connaître explicitement ses préférences.

On peut faire une distinction entre trois types de recommandations (Chen, Hu, Kuo & Liang, 2010). Premièrement, celles **générées par l'utilisateur uniquement**. Cela fait référence aux recommandations basiques en fonction des recherches antérieures du consommateur (*action-to-item*), où l'utilisateur pourrait lire un message tel que « nous avons sélectionné les éléments suivant pour vous ».

Ensuite, les **recommandations basées sur l'objet lui-même**. Il s'agit de suggestions de contenus ou produits complémentaires sur base des attributs des produits déjà achetés ou contenus déjà visualisés (*item-to-item*). Le message suivant pourrait alors éventuellement apparaître: « vous pourriez également aimer ceci ».

Troisièmement, il existe des recommandations plus complexes **basées sur le comportement d'un utilisateur et celui d'autres utilisateurs conjointement, appelé l'environnement social.** Ici, nous

retrouvons les recommandations basées sur ce que d'autres ont visualisé, du genre « Les visiteurs ayant consulté ce contenu/produit ont également consulté celui-ci », ainsi que les recommandations en fonction de ce que d'autres ont acheté, dans le cas d'un site d'e-commerce, par exemple « les consommateurs ont également acheté ceci ». Les produits ne sont alors pas nécessairement complémentaires mais néanmoins adaptés. Smart Insights (2005) insiste sur la différence entre ces deux dernières recommandations car la première serait la méthode la plus efficace en termes de revenus, étant donné que ce n'est pas du « baratin » publicitaire mais encourage l'individu à regarder autre chose, pouvant mener à des décisions d'achats non-planifiées. Cela retient les visiteurs et donne une impression de faire partie d'un large groupe d'utilisateurs. Au contraire, dans le deuxième cas, la suggestion sera nettement moins efficace car c'est trop insistant et non approprié pour ceux n'ayant aucune intention d'achat. Les processus complexes nécessaires à la mise en œuvre de ces suggestions appropriées sont expliqués dans la section suivante.

2.5. Fonctionnement et technologies

Nous avons plusieurs fois énoncé l'importance des technologies pour la personnalisation, dont la technique de filtrage collaboratif, et allons désormais tenter de mieux comprendre leur fonctionnement. Il est en effet important d'éclaircir cela, car bien que ces technologies modernes permettent d'offrir du contenu de valeur aux consommateurs, c'est exactement ces traitements et usage de l'information personnelle qui peut empiéter sur le désir de confidentialité de ceux-ci, comme nous le traiterons plus en profondeur dans le prochain chapitre. Ceci peut à son tour, nuire à l'acception et la performance de la personnalisation.

La personnalisation en ligne est principalement implémentée via des outils CRM en combinaison avec des techniques sophistiquées de *data mining* ou fouille de données (De Filippi, 2016). Ces outils sont utilisés afin d'analyser les interactions des consommateurs, qui circulent sous forme de données, une grande masse provenant de diverses sources, appelée *Big Data* et régulièrement considérée comme l'ADN du marketing. A partir de cette grande base de données, le *data mining* peut tirer des informations qui ne sont pas facilement accessibles à partir de chaque source de données individuelle. Un profil utilisateur sera alors créé puis lié aux produits et services de l'entreprise afin de finalement offrir du contenu et une structure web adaptés aux besoins spécifiques et individuels (Cambria et al. cités dans De Filippi, 2016).

Sommairement, nous pouvons dire que le processus technique de la personnalisation ressemble à ceci : données entrantes (*inputs*) – algorithmes – publication (*outputs*). Nous décrirons plus précisément la mise en place de la personnalisation ci-dessous, basé sur les étapes d'Eirinaki et Vazirgiannis (2003), notamment le profilage des utilisateurs, l'extraction et l'analyse des données, la gestion du contenu, la publication sur le site web et enfin, le renouvellement du procédé.

Le profilage des utilisateurs (inputs). C'est le fait de recueillir et d'enregistrer de l'information spécifique à un visiteur dans le but de créer un profil utilisateur. Cette information peut être une donnée externe ou une donnée client provenant des interactions avec l'entreprise, partagée directement ou indirectement (explicitement ou implicitement) par un individu. Comme déjà évoqué, cette collecte d'informations personnelles est intrinsèque à la personnalisation et la divulgation par les utilisateurs

est donc cruciale (Xu, Luo, Caroll & Rosson, 2011). L'entreprise va ensuite obtenir ces données personnelles en enregistrant toutes les traces des consommateurs sur tout type de plateforme alors qu'ils naviguent sur le web, par l'intermédiaire de technologies appropriées. Ces traces peuvent être le nom du consommateur donné volontairement, l'adresse e-mail communiquée, les préférences exprimées, l'historique d'achat ou de recherches récolté via des cookies, l'emplacement de l'individu, des évaluations de produits et services, des informations postées sur les réseaux sociaux etc. et peuvent quelque fois être obtenues à l'aide de partenariats avec Google ou Yahoo par exemple (Aguirre et al., 2015).

Aguirre et al. (2015) expliquent que les entreprises **récoltent** ces données via la méthode **«*covert* » ou « *overt* ».** Le premier fait référence à une collecte des données sans empiéter sur l'expérience de navigation en ligne car les utilisateurs ne sont pas conscients que l'entreprise enregistre leurs données en temps réel (Montgomery & Smith, 2009). Le fait de récolter cette information discrètement permet à l'entreprise d'obtenir des informations non biaisées et donc une connaissance du consommateur plus riche, permettant de mieux adapter ses services. Cependant, lorsque les consommateurs réalisent que leurs données ont été récoltées sans leur consentement, en voyant, par exemple, une publicité basée sur leurs données, ils seront surpris, auront l'impression de perdre le contrôle de leurs données et se sentiront plus vulnérables. L'étude réalisée par Aguirre et al. (2015) a d'ailleurs démontré qu'une stratégie c*overt* n'a pas d'impact sur les taux de clics de publicités sur Facebook pour cette raison. Au contraire, selon la méthode *overt*, les consommateurs savent que leurs données sont recueillies car l'entreprise a fait l'effort de les informer (Sundar & Marathe, 2010). L'idée derrière cela, est qu'à force d'être informé que les données seront récoltées et que sans s'y opposer, les consommateurs donnent en quelque sorte leur consentement. Cette méthode fait entre autres référence à l'utilisation des cookies, où généralement un message d'information apparait sur le site, informant les consommateurs que l'entreprise garde trace de tout mouvement sur le site web mais ne demande pas nécessairement le consentement explicite. Une demande excessive du consentement des consommateurs pourrait négativement impacter le flux d'achat du consommateur et prolonger le temps nécessaire pour une action en ligne (Aguirre et al., 2015).

Ceci n'est pas à mélanger avec une **manière active de récolter des données**, qui est le fait de poser directement des questions afin de mieux connaître l'utilisateur. Alors que la méthode **passive** oblige le marketer de faire des déductions en ce qui concerne les intérêts des consommateurs, en utilisant des informations de transactions ultérieures ou des données à propos du parcours de navigation (Ho & Bodoff, 2014). Deezer utilise, par exemple, l'approche active et passive car, lors de l'inscription, l'utilisateur sélectionne des chansons ou artistes de musique qu'il aime et par la suite, son comportement est analysé en fonction de ses sélections et des comportements similaires d'autres utilisateurs afin de compléter son profil (voir filtrage collaboratif).

Nous verrons plus loin que, ce qui peut poser problème en termes de protection de la vie privée lors de cette étape de la personnalisation, est que la collecte des données personnelles par l'entreprise peut causer un sentiment de vulnérabilité auprès du consommateur. Nous parlerons même **d'atteinte à la vie** privée auprès du consommateur lorsque cette information est récoltée et utilisée sans que le

consommateur en ait été averti et ai pu donner son consentement. Il s'agit donc du **conflit entre le droit du client de protéger sa vie privée et le droit des entreprises de récolter** et s'approprier ces données personnelles (Portes, 2018).

L'extraction et analyse des données (algorithmes). La deuxième phase est le traitement de l'information enregistrée dans le serveur (*web logs*) en utilisant des techniques de *data mining* comme la classification, le *clustering*, les règles d'association et la découverte de modèles séquentiels. Le *data mining* est un processus informatique consistant à agréger des données sans rapport. C'est-à-dire que l'entreprise sélectionne, lie, classe et synthétise ces données afin d'en extraire de nouvelles informations statiques et découvrir des schémas d'utilisation individuels et spécifiques, qui sont en quelque sorte des déductions de comportements probables des consommateurs (Eirinaki & Vazirgiannis, 2003). Ces schémas d'utilisateurs se compléteront au fur et à mesure des interactions avec l'entreprise et autres utilisateurs. Les techniques de *data mining* permettent donc de transformer les données et souligner des liens entre les informations, mais ne offren pas la possibilité de faire des recommandations. Elles permettent uniquement de générer des hypothèses (Rygielski, Wang & Yen, 2002).

Pour cette phase suivante de recommandation et détermination d'actions à réaliser, 3 approches existent afin de trouver les corrélations entre les schémas et le contenu des pages web : le filtrage basé sur le contenu, le filtrage collaboratif et le filtrage par règle, décrits ci-dessous. Elles diffèrent en fonction des données traitées et algorithmes utilisés (Chang, Changchien & Huang, 2006). Par ailleurs, la performance de ces différentes méthodes est liée à la qualité du système d'association entre l'utilisateur et l'item, que ce soit via une intervention humaine, des analyses statistiques ou l'apprentissage par machine.

Tout d'abord, le **filtrage basé sur le contenu** dépend uniquement des préférences des utilisateurs individuels. Le système suit tout le comportement d'utilisation d'un utilisateur et recommande des éléments similaires à ceux utilisés, consultés ou évalués favorablement dans le passé par cet utilisateur.

Un système de **filtrage collaboratif** va analyser le comportement d'un utilisateur afin de trouver des corrélations avec d'autres profils utilisateurs assez similaires. Les préférences de ces autres consommateurs proches de celles de l'individu en question seront utilisées afin de lui recommander ce qui a déjà été adopté par ceux-ci. Ceci fonctionne selon la supposition que les utilisateurs avec un comportement similaire ont des intérêts analogues (Eirinaki & Vazirgiannis, 2003).

Dans le cas du **filtrage basé sur une règle**, l'élaboration du profil utilisateur est basée sur des règles de sélection, manuelles ou automatiques. L'information est identifiée, analysée puis divisée dans la catégorie correspondante selon un arbre décisionnel, afin de faire correspondre le profil avec les produits ou services adaptés. Pour illustrer ceci, imaginons un vendeur qui identifie que certains consommateurs peu engagés ont tendance à ouvrir leurs e-mails et naviguer sur le site web mais achètent principalement en magasin, rarement en ligne. Ces acheteurs seront alors mis dans un segment appelé « lèche vitrine ». Grâce à des règles spécifiques, les personnes répondant aux

caractéristiques de ce segment seront identifiées, assignées à ce segment et des actions spécialement désignées pour celui-ci seront automatiquement exécutées, comme l'envoi de notifications *push* et d'e-mails. Cette méthode nécessite l'intervention du spécialiste marketing afin d'établir les règles d'approches, alors que les autres sont automatiques.

Notons qu'afin de déduire des conclusions plus précises et combler les lacunes de chacune de ces méthodes, ces approches peuvent être utilisées conjointement. L'approche hybride la plus utilisée est la combinaison du filtrage collaboratif avec celle basée sur le contenu car différentes sources sont alors utilisées, améliorant la précision des recommandations sur mesure.

Ce qui peut s'avérer problématique pour le consommateur dans cette étape, est le fait que l'entreprise puisse **l'identifier et en apprendre trop sur lui** à l'aide de ces techniques très sophistiquées mais également très opaques et intrusives ; celles-ci étant capables de surveiller, stocker, mémoriser et communiquer toute l'information personnelle (Martin & Murphy, 2017). C'est pour cette raison que certains utilisateurs accueillent les nouvelles technologies comme la reconnaissance faciale avec forte suspicion.

Gestion du contenu. C'est le processus de classification du contenu d'un site web en catégories sémantiques adaptées au profil de l'utilisateur analysé, afin de rendre la recherche d'information et la présentation plus facile pour les individus. La gestion du contenu est extrêmement importante pour les sites web dont le contenu augmente quotidiennement, comme pour les sites d'actualité.

Publication sur site web (ou e-mail/app mobile : outputs). C'est la présentation uniforme du contenu adapté à l'utilisateur final. Cet *output* peut concerner n'importe quel élément des 4P : la communication, le produit ou service, le prix et/ou la livraison (Vesanen, 2007). La page web est alors modifiée de manière explicite ou implicite, de façon à ce que l'utilisateur s'en rende compte ou non, comme mentionné dans la section Typologie. Etant donné qu'une session en ligne est particulièrement courte, Ho, Bodoff et Tam (2011) ainsi que Lee et Cranage (2011) conseillent d'afficher les adaptations et recommandations de produits et services avant le début de la recherche de l'individu ou en temps réel afin d'obtenir l'effet désiré.

Cette étape peut également susciter des préoccupations en matière de vie privée auprès des consommateurs dans le cas où **le message personnalisé est illicite, indésirable, trop fréquent et envahissant ou contenant trop d'information personnelle spécifique**. Les consommateurs ont, par conséquent, le sentiment d'être observés et de ne pas avoir le contrôle, ce qui les poussera à adopter des mesures de protection de la vie privée (voir pièges et challenges associés à la personnalisation).

Acquisition de l'information et recherche. La personnalisation étant un processus itératif, cela veut dire que plus l'entreprise retient longtemps un client et interagit avec lui, plus elle acquiert de l'information à son sujet et peut affiner son expérience ainsi que rendre ces étapes plus efficaces via l'expérience accumulée (Murthi & Sarkar, 2003). Il faudra donc constamment évaluer son efficacité à l'aide de feedbacks et acquérir de nouvelles données afin d'enrichir le profil utilisateur. Ces données

n'étant pas toujours stockées dans le serveur du site web en question, il se pourrait que les éditeurs des sites doivent obtenir cela via des partenariats, des connexions aux réseaux sociaux, le dépôt de cookies sur d'autres sites, voire même d'en acheter.

Toutes ces techniques peuvent donc permettre de répondre à un besoin en temps réel ou de prédire certains événements ou tendances (De Filippi, 2016). Il apparait que la réussite de la personnalisation dépend de la capacité à correctement détecter et puis agir sur les préférences actuelles. Evidemment, l'investissement dans ces technologies informatiques et dans les capacités d'analyses nécessaires à la personnalisation n'est pas insignifiant. Il est donc d'autant plus important que les vendeurs en ligne s'intéressent réellement au comportement du consommateur dans un contexte de personnalisation, s'ils souhaitent que l'utilisation des technologies génère les bénéfices espérés, mentionnés ci-dessous.

2.6. Conséquences positives

Nous avons donc pu constater que les entreprises doivent avant tout utiliser la personnalisation afin de répondre aux besoins des consommateurs et de les satisfaire. Grâce à cela, cette stratégie marketing peut également procurer des effets positifs considérables à l'entreprise (Salonen & Karjaluoto, 2016).

Tout d'abord, en ce qui concerne les consommateurs, grâce à la personnalisation offerte par l'entreprise, ils pourraient bénéficier d'**offres spéciales** ainsi que de **produits et services améliorés ou nouveaux.** Ayant plus d'information quant aux préférences des consommateurs, l'entreprise peut en effet mieux adapter son offre et offrir des réductions en échange de ces données (Aguirre et al., 2015). Par ailleurs, ils bénéficieront d'un **meilleur ciblage en ligne** de produits, services et messages publicitaires suggérés afin de les guider dans leurs achats ou autres décisions en ligne (Xu et al., 2011). L'expérience d'achat du consommateur sera alors plus agréable ainsi que plus efficace car ces techniques sont rapides et pratiques, deux éléments fortement appréciés aujourd'hui (Aguirre et al., 2015 ; Chellappa & Sin, 2005).

En effet, le fait d'offrir des recommandations pertinentes **réduit la surcharge d'information** pour le consommateur qui se voit proposer exactement ce qu'il recherche (Li & Unger, 2012 ; Tam & Ho, 2005 ; Liang, Lai & Ku., 2006). Cette amélioration du traitement de l'information signifie que le consommateur devra fournir moins d'efforts cognitifs, ce qui aura par la suite généralement un impact **positif sur sa satisfaction ainsi que sur sa prise de décision,** telle que l'adoption d'un service, le partage de données personnelles ou l'achat (Liang, Lai et Ku, 2006 ; Karwatzki et al., 2017 ; Aguirre et al., 2015). Cela empêche donc le paradoxe du choix ainsi que la probabilité que les consommateurs quittent le site rapidement sans faire aucune transaction, car se trouver face à trop d'options les effraye généralement (Lee & Kwon, 2008).

Puis, concernant les entreprises, elles bénéficieront sans aucun doute du fait que la personnalisation en ligne influence de manière favorable l'attitude et le comportement du consommateur envers le site web et le service (Aguirre et al., 2015 ; Lee & Cranage, 2011). Ainsi, suite au pistage de ces comportements et éventuellement le partage volontaire de données, les entreprises bénéficient d'une **meilleure compréhension de leurs clients**, permettant d'identifier chaque consommateur individuellement (Di Filippi, 2016 ; Chau et al., 2013).

Ensuite, en tant que dimension du marketing relationnel, la personnalisation pourrait avoir des conséquences positives sur la relation que les entreprises entretiennent avec leurs consommateurs. Effectivement, selon plusieurs chercheurs, offrir des messages personnalisés suscite des sentiments positifs concernant l'interaction avec l'entreprise, ce qui augmente la probabilité de création de **relations durables et de confiance** (Ansari & Mela, 2003 ; Li & Unger, 2012). L'étude d'Evergage (2018, p.1) confirme cela, car il est apparu que 98% des marketers sont d'accord avec le fait que la personnalisation améliore les relations client, dont 74% affirment voir un impact qualifié de « fort » ou « extrême ». Cependant, bien que certaines recherches ont affirmé qu'offrir un ciblage personnalisé permet de développer la fidélité et est même une condition nécessaire à cela, d'autres n'ont pas réussi à démontrer cette relation (Salonen & Karjaluoto, 2016). Par ailleurs, Ansari et Mela (2003) et Lee et Cranage (2011) confirment l'impact positif sur la **rétention** des clients actuels. Ceci est, entre autres, dû au fait que les coûts de transfert vers une autre entreprise en ligne sont élevés pour un consommateur car cela implique de devoir divulguer à nouveau son information et apprendre le fonctionnement de la personnalisation du nouveau fournisseur. On appelle ceci l'effet d'immobilisation. Plus le niveau de personnalisation est élevé, plus le consommateur est engagé et les coûts de transferts pour celui-ci sont élevés (Son & Kim, 2008).

Concernant les revenus, Ho et Bodoff (2014) révèlent que la personnalisation **augmente les revenus de la publicité** via des meilleurs taux de clics et taux de ventes. Ainsi, Ho et Tam (2005) et Tucker (2014) ont affirmé que dans le contexte de la publicité personnalisée en ligne, on peut observer plus d'achats ; ces publicités étant deux fois plus efficaces que les traditionnelles.
Par ailleurs, dans le cas des entreprises d'e-commerce, Li et Unger (2012) ainsi que Arora et al. (2008) mentionnent que la personnalisation en ligne offre la possibilité aux entreprises de plus facilement **renforcer la différenciation des produits** et donc d'obtenir des **marges de profits plus élevées.** L'étude d'Evergage (2018, p.22) avait également établi que 87% des marketers disent avoir observé une amélioration considérable de plus de 10% de leurs résultats commerciaux provenant des efforts personnalisés, soit trois fois plus que s'ils n'avaient pas intégré cela (Abraham, Mitchelmore, Collins, Maness, Kistulinec, Khodabandeh, Hoenig & Visser, 2017). De meilleurs résultats de ventes peuvent également être liés **à des couts de marketing et production plus faibles**, grâce au fait que la personnalisation permet de simplifier la gamme de produits et services tout en créant une demande plus prévisible (Deloitte, 2015). Une autre raison pourrait être **le taux plus élevé d'acquisition de nouveaux visiteurs sur le site web et la transformation de ceux-ci en acheteurs**. Ensuite, selon Vesanen (2007) la personnalisation justifie aussi une facturation de prix supérieurs et l'étude de Li et Unger (2012) a révélé que les clients susceptibles d'utiliser la personnalisation seront prêts à payer pour cela. Puis, cela permettrait même éventuellement de pratiquer une tarification discriminatoire (Blattberge & Deighton cités dans Chellappa & Sin, 2005). Enfin, Chellappa et Shivendu (2007) indiquent également que la valeur la plus évidente et directe de l'information du consommateur pour l'entreprise réside dans **les revenus du partage ou de la vente de ces données** à des partenaires de la chaine d'approvisionnement ou à des annonceurs.

Pour finir, la masse de données récoltées représente un réel **avantage concurrentiel** pour l'entreprise car elle peut être mise à profit afin de se différencier de ses concurrents en lui permettant de mieux

comprendre et prédire la demande, gérer l'inventaire, créer des chaines d'approvisionnement efficaces et générer de nouvelles opportunités (Chellappa & Shivendu, 2007). En outre, Cao et Li (cités dans Salonen & Karjaluoto, 2016) jugent que la personnalisation est l'outil le plus efficace afin **d'assurer le succès d'une entreprise en ligne**. Cependant, l'efficacité de la personnalisation est contestée par certains chercheurs comme Shen et Dwayne ainsi que Zhang (cités dans Salonen & Karjaluoto, 2016) qui n'ont trouvé aucun appui pour cela. Cela peut être dû au fait que la personnalisation est un concept attirant mais difficile à implémenter en tant qu'outil. Pour terminer, Lee et Cranage (2011) précisent que la personnalisation n'est pas une condition nécessaire et suffisante à la réussite, la personnalisation étant actuellement disponible et abordable pour tous et ne représentant donc plus un avantage concurrentiel. Ce qui fait la différence et rend les services de personnalisation réellement utiles et profitables, c'est d'implémenter des pratiques de protection des données, afin de protéger l'information du consommateur (ceci sera abordé plus tard).

Il est important de préciser que la majorité de ces effets semblent dépendre de facteurs contextuels comme le type de consommateur (Che et al., cités dans Salonen & Karjaluoto, 2016), la phase dans le processus d'achat (Lambrecht & Tucker, cités dans Salonen & Karjaluoto, 2016), la qualité perçue du service (Chellappa & Shivendu, 2007) ainsi que le timing de la personnalisation (Ho & Bodoff, 2014). De plus, étant donné la nature fluctuante des préférences du consommateur, l'effet de la personnalisation aura tendance à changer avec le temps.
Bien qu'il semble clair que la personnalisation en vaut la peine pour l'entreprise et que cela se confirmera dans le futur avec l'essor de l'intelligence artificielle, cette stratégie suscite également de nombreuses discussions.

2.7. Les pièges de la personnalisation et challenges associés

Il y a, en effet, un revers à la médaille car la personnalisation peut impliquer des coûts ou effets indésirables pour les consommateurs qui se répercutent sur l'entreprise. Certains de ces effets négatifs peuvent empêcher l'utilisation généralisée de la personnalisation.

Tout d'abord, la personnalisation augmenterait les **perceptions des risques quant au respect de la vie privée** en raison de la divulgation d'information nécessaire et impacterait de ce fait négativement le comportement du consommateur, tel que sa volonté d'utiliser ces services. Un chapitre entier sera dédié à ce coût principal.

Deuxièmement, comme mentionné dans les avantages en ce qui concerne la rétention, **l'effet d'immobilisation** suite à la personnalisation peut décourager les consommateurs à l'idée d'utiliser des méthodes de personnalisation alternatives. Par conséquent, il est très difficile d'attirer des clients de ses concurrents, d'autant plus qu'acquérir un nouveau consommateur peut coûter jusqu'à 10 fois plus que retenir un client (Chellappa & Sin, 2005).

En outre, les consommateurs craignent que la personnalisation soit synonyme de **frais additionnels ou tarifs discriminatoires** (Vesanen, 2007). Cela vient de l'idée que, comme les entreprises en savent beaucoup sur les préférences des consommateurs et habitudes, ils adaptent l'information au cas par cas

mais pourraient éventuellement aussi adapter leurs prix. Une pratique qui est souvent reprochée aux compagnies aériennes.

Un autre problème est la « **dépersonnalisation** » en cas de campagnes personnalisées non-désirables, illicites, trop fréquentes et personnelles ou des messages dont la valeur perçue est faible. En effet, selon Accenture Interactive (2018, para.4), 27% des répondants disent être anxieux suite à une expérience trop personnelle ou invasive avec la marque et 64% du fait que l'entreprise détenait de l'information qu'ils n'avaient pas partagé délibérément. Ceci est généralement dû à une mauvaise compréhension des préférences par l'entreprise (Baek & Morimoto, 2012). Ce type de personnalisation non-appropriée crée le sentiment d'être observé et oppressé, tout particulièrement lorsque l'individu reçoit des promotions sur son smartphone, comme les messages instantanés au moyen de techniques de géolocalisation lorsqu'on passe devant le magasin (Baek & Morimoto, 2012 ; Accenture Interactive, 2018). Afin de récupérer le sentiment de contrôle de leurs données et éviter tout acharnement publicitaire, les individus adoptent des mesures de protection. Son et Kim (2008) citent 6 formes de réactions à ces messages de marketing hautement ciblés ou illicites : le refus de donner de l'information personnelle, l'utilisation de fausses données, la demande de retrait de ses informations des bases de données des entreprises en ligne, le bouche-à-oreille négatif et le dépôt de plaintes directement auprès des entreprises en ligne ou indirectement auprès d'organisations tierces. Wirtz, Lwin et Williams (2007) rajoutent qu'ils peuvent avoir recours à des technologies afin de rétablir l'équilibre de pouvoir et se protéger, tels que les *AdBlockers*, le filtrage d'e-mails, la navigation privée sur internet, les logiciels antispywares etc. Ces actions nuisent fortement à l'entreprise car cela peut donner lieu à des analyses erronées des préférences et tendances de consommation et, dans une plus large mesure, à une analyse incorrecte du marché cible. Par ailleurs, cela constitue une perte de fidélité et impacte la réputation de l'entreprise à long terme, ce qui conduit les spécialistes du marketing à encourir des coûts plus élevés.

Concernant les challenges, force est de constater qu'aujourd'hui, certaines entreprises sont encore loin du compte en matière de personnalisation de l'offre au client dans le respect de sa vie privée et selon un ciblage individuel approprié et juste. Premièrement, les consommateurs ne se sentent pas toujours privilégiés par la personnalisation ou autre programme de marketing relationnel, estimant que ces pratiques **ne vont en général que dans un sens** et ainsi renforcent l'inégalité entre l'entreprise et les individus. Plus précisément, certains consommateurs considèrent que les entreprises utilisent certaines stratégies marketing encore trop dans une idée de performance et agissent dans leur propre intérêt à l'aide de tactiques trompeuses et manipulatrices, alors qu'elles devraient porter plus d'attention envers les consommateurs et ainsi, les servir de manière éthique et unique sans les submerger (Aguirre et al., 2015). Certains individus refuseront de partager de l'information et offrir aux entreprises le contrôle de leur interaction pour ces raisons.

Ensuite, les individus doutent des éventuelles initiatives prises par les entreprises afin de protéger leur vie privée. Bien que prospérant dans un monde dominé par la technologie, les consommateurs veulent être **traités comme des humains** et non des séries de données. Le défi est donc d'humaniser l'expérience digitale en passant de communications marketing à des conversations digitales où l'entreprise comprend le consommateur, anticipe ses besoins et le protège (Salesforce Research, 2016).

Nous approfondirons notre compréhension de ce besoin de protection et pratiques loyales dans le prochain chapitre.

Un autre challenge concerne le **manque d'expertise des consommateurs**. Le fait qu'ils ne comprennent pas toutes les techniques de traitement de l'information qui se cachent derrière les processus de personnalisation contribue à créer un sentiment de méfiance et freine les individus à partager leurs données personnelles (Li & Unger, 2012).

Enfin, la personnalisation n'est pas encore exploitée au maximum car les entreprises estiment qu'elles n'ont pas encore assez de données. Or, les données clients ne manquent pas. Ce qui pose problème est que les entreprises n'utilisent pas encore suffisamment les technologies adaptées ou du moins, pas assez efficacement (Vesanen, 2007). Seulement une minorité, 26% selon Evergage (2018, p.16) utilise la personnalisation basée sur le *machine learning* par exemple. Pourtant, Evergage (2018) indique que les entreprises ne seront que satisfaites de leurs stratégies de personnalisation à partir du moment où elles adopteront le *machine learning*, car elle permet d'offrir des expériences *one-to-one* à grande échelle ne nécessitant aucun effort humain.

2.8. Bons exemples de personnalisation

Nous finirons ce chapitre concernant la personnalisation en citant quelques exemples d'entreprises ayant réussi à implémenter des techniques de personnalisation performantes. Ces techniques sont utilisées dans de nombreux environnements en ligne, que ce soit des moteurs de recherches, des détaillants en lignes, des annonceurs, des réseaux sociaux ou autres. Les plus avancés sont probablement les géants du web comme Google, Amazon et Netflix.

Prenons tout d'abord Google et Yahoo qui peuvent affiner les recherches des utilisateurs en incorporant les recherches antérieures (Aguirre et al., 2015). Avant, des utilisateurs qui différaient en termes d'objectif, d'expertise d'âge, de carrière etc., recevaient des résultats identiques lorsqu'ils faisaient une même requête sur un moteur de recherche. Le but premier de la personnalisation, dans ce contexte, est de ne plus offrir de résultats statiques mais d'adapter les résultats de recherche sur base des recherches précédentes et peut-être même sur base de la connaissance observée de l'utilisateur afin de faciliter la recherche. Par exemple, pour une recherche avec le mot « voyage », selon les techniques de personnalisation, il aura été identifié sur base de l'historique de recherche qu'un certain utilisateur voyage régulièrement pour le travail à San Francisco, contrairement à un autre individu. Des résultats de recherche leur seront proposés tels que des packs voyages très différents (Montgomery & Smith, 2009). De plus, l'enregistrement des données permet l'auto-remplissage sur de nombreux sites partenaires de Google. Lorsque quelqu'un fournit son adresse de domicile à Google, cela permettra à Google de remplir cette information automatiquement lorsqu'il choisira l'option de livraison ou de carte de crédit avec des vendeurs Froogle par exemple, faisant gagner du temps à l'individu (Chellappa & Shivendu, 2007).

Amazon, cette entreprise d'e-commerce avec une offre dont on a parfois l'impression qu'il n'y a pas de limite, rend également la recherche extrêmement facile en offrant brillamment de la personnalisation. Considérée comme pionnier en matière de recommandations personnalisées, elle

offre en effet de nombreuses suggestions sur mesure via les fonctionnalités « Your Amazon », « l'affaire du jour », « la liste de souhaits et cadeaux », les recommandations par catégorie, les e-mails personnalisés avec le nom du client etc. Amazon utilise également les *chatbots*, des logiciels conversationnels qui permettent de répondre aux requêtes des consommateurs en temps réel. Tout ceci grâce à son utilisation minutieuse du *Big Data* et le filtrage collaboratif, basé sur les transactions effectuées antérieurement par le client, ce qu'il a visualisé, ses évaluations des produits et/ou l'agrégation des tendances d'achat et préférences d'utilisateurs similaires (Lee & Cranage, 2011). La personnalisation est d'ailleurs un axe clé de différentiation pour Amazon, désirant devenir l'entreprise la plus orientée client. Ainsi, Amazon aurait un taux de rétention deux fois plus important qu'une librairie normale, entre autres grâce à la personnalisation et il semblerait même que certaines personnes aillent sur le site uniquement car ils sont curieux de voir ce qu'il va leur être suggéré afin de potentiellement découvrir de nouvelles choses (Montgomery & Smith, 2008).

Autre grand acteur, Netflix, une plateforme de streaming de vidéos sur abonnement. Cette entreprise a réussi à révolutionner le secteur de la location de films grâce à sa stratégie commerciale alternant création de films, séries et documentaires originaux mais également grâce à son algorithme de recommandation de contenu et l'exploitation d'une large base de données sur le comportement de ses utilisateurs. Les recommandations personnalisées comme « Top Choix de Floriane » ou « Sortis récemment » sont basées sur les visionnages et évaluations préalables de l'individu ainsi qu'en comparant le profil d'un individu avec celui d'autres utilisateurs similaires.

Nous pouvons aussi prendre Nivea comme exemple, qui affiche des vitrines électroniques différentes en fonction de si c'est la première visite de l'utilisateur, un nouvel acheteur ou un consommateur fidèle. Les deux premiers pourront visualiser des produits peu chers afin d'encourager l'engagement, les utilisateurs fidèles, quant à eux, recevront des produits de plus haute valeur en forme de pack afin d'augmenter la quantité du panier. Un test A/B a dévoilé qu'après cette implémentation, le taux de conversion de navigation s'est amélioré de 70% et les transactions ont augmenté de 150% (Abraham et al., 2017, para. 4).

Enfin, les sites destinés à l'actualité sont également connus pour leur personnalisation sous forme de pages d'accueil personnalisées avec des articles sélectionnés en fonction des intérêts de l'utilisateur, ainsi que les agences de voyages proposant constamment, via site web, mail ou publicités, des packs voyages basés sur les destinations favorites et les recherches antérieures, ou encore les systèmes de recommandations de musique tels que Spotify et Deezer. Nombreuses sont les entreprises ayant tenté de tirer parti de la personnalisation comme les géants du web le font mais peu y arrivent étant donné les investissements importants, le temps et les capacités nécessaires ainsi que dû aux autres challenges cités ci-dessus.

2.9. Conclusion

En résumé, les entreprises ont dû mettre la personnalisation au premier plan de leur agenda stratégique afin de répondre au mieux aux nouvelles attentes des consommateurs. Non seulement, les interactions personnalisées apportent de la valeur au client, mais elles permettent également aux entreprises de collecter encore plus d'information sur lui. Nous retenons surtout que cette technique marketing peut

se présenter sous des formes variées et diverses, mais pas toujours adaptées. Les utilisateurs peuvent, en effet, percevoir la personnalisation comme une atteinte à la vie privée, s'ils n'ont pas consenti à la collecte de données ou si la personnalisation apparait comme trop personnelle. Lorsque le consommateur considère l'entreprise comme opportuniste, il pourrait prendre des mesures de protection afin de rétablir l'équilibre de pouvoir et ainsi fortement nuire à l'entreprise, qui dépend d'informations précises sur les consommateurs. L'équilibre entre confidentialité et personnalisation, semble donc rarement atteint. Dans cette perspective, Fan et Poole (cités dans Salonen & Karjaluoto, 2016) spécifient qu'avant la question était de savoir comment personnaliser alors qu'aujourd'hui il s'agit de réfléchir à comment le faire de manière adéquate, c'est-à-dire dans le respect de la vie privée des individus. Nous étudierons plus en profondeur ce deuxième thème central dans les prochains chapitres. Ce qui nous intéresse tout particulièrement dans ce chapitra est la valeur que la personnalisation apporte aux consommateurs. Ainsi, nous avons pu constater que l'anticipation des bénéfices de la personnalisation, comme des offres spéciales, un gain de temps, une réduction de la quantité d'information et du contenu ciblé et intéressant, favorisent l'intention de divulgation du consommateur. Ceci nous amène à inclure ces attentes en tant qu'incitation dans notre modèle de recherche. Par conséquent, nous posons l'hypothèse suivante :

> *Hypothèse* : Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention du consommateur de partager des données personnelles.

Maintenant que nous maitrisons ce qu'est la personnalisation et ce qu'elle peut apporter au consommateur, nous tenterons, dans les prochains chapitres, de comprendre d'autres critères sur lesquels se base majoritairement la décision de dévoiler les données personnelles.

3. Protection de la vie privée (*privacy*)

Ce troisième chapitre est nécessaire car il est apparu ci-dessus que l'effet indésirable majeur de la personnalisation est l'intrusion par l'entreprise dans la vie privée des consommateurs suite à la connaissance de leurs données personnelles. Ceci pourrait constituer un frein à l'intention de divulgation de ses informations privées et mérite donc d'être pris en compte dans notre étude.

Comme mentionné dans le chapitre précédent, **pour les entreprises**, récolter et gérer les données des consommateurs peut être un procédé précieux afin de mieux connaitre et cibler son consommateur ainsi que pour saisir de nouvelles opportunités. Toutefois, cela constitue également un risque considérable si elles les perdent ou en abusent (KPMG, 2017). En faisant référence à la théorie des deux facteurs, Salonen et Karjaluoto (2016) expliquent notamment que la confidentialité est ce qui s'appelle un facteur d'hygiène, dans le sens où une protection élevée a un effet limité sur l'amélioration des performances, alors que les failles et abus des données ont un impact significatif sur la performance et la fiabilité de l'entreprise. En effet, cela pourrait causer une impression de vulnérabilité et violation de contrat auprès d'une communauté entière, et par conséquent une perte de réputation (Pavlou & Gefen, cités dans Xu et al., 2011). Quand on voit comment tout s'est déroulé dans le cadre du scandale Facebook – Cambridge Analytica, on imagine de fait rapidement l'amplitude des conséquences lorsque les données sont perdues ou partagées. Puis, en savoir beaucoup, voire trop, sur son consommateur et le

faire ressentir peut être perçu comme un abus de confiance par les consommateurs et peut les faire fuir (SAS, 2015). Les entreprises doivent donc trouver un juste milieu et ont tout intérêt à minimiser les inquiétudes des consommateurs, car elles sont nombreuses.

C'est exactement cette nature asymétrique de la relation entre entreprises en tant que collecteurs de données et consommateurs, fournisseurs de données, qui aggrave le problème de respect de la vie privée dans le contexte de la personnalisation en ligne et qui nous intéresse. Dès lors, nous allons à présent aborder le **positionnement du consommateur** par rapport aux données personnelles qu'il serait amené à partager.

3.1 Définition des concepts

Bien que le thème de la protection de la vie privée ne soit pas nouveau, son sens et son importance ont évolué dans un contexte où l'exploitation de l'information du consommateur est désormais devenue la première source de valeur pour les entreprises mais également source de craintes pour les consommateurs.

Effectivement, le **concept de la confidentialité ou protection de la vie privée** a été défini initialement, dans le domaine de la psychologie, par Warren et Brandeis (cités dans Baruh, Secinti & Cemalcilar, 2017, p.26) comme « *the right to be let alone* », c'est-à-dire le droit d'être laissé seul, d'avoir son intimité. La protection de la vie privée faisait alors essentiellement référence à la protection physique d'une personne, tandis qu'aujourd'hui, il s'agit également de la protection de l'information personnelle *(information privacy)*. Ce travail s'inscrit donc dans ce deuxième type de protection de la vie privée. Ainsi, dans le domaine des systèmes d'informations, la protection de la vie privée ou confidentialité (*privacy*) a été généralement définie comme étant « la capacité d'un individu de contrôler les conditions selon lesquelles son information personnelle est recueillie et utilisée » donc contrôler quand, comment et jusqu'où son information personnelle est communiquée aux autres (Westin, cité dans Chellappa & Sin, 2005, p. 186). La définition plus récente de Turn (cité dans Alhouti, Johnson & D'Souza, 2016, p.23) rejoint cette idée; la confidentialité faisant référence aux « droits des individus en termes de collecte, stockage, traitement, distribution et utilisation de l'information personnelle ». En d'autres termes, la protection de l'information donne aux consommateurs les droits suivants : la connaissance de l'information récoltée à leur propos, la capacité d'accéder à cette information et de contrôler ces informations, notamment pouvoir décider de ne pas divulguer les données personnelles aux autres (Rognehaugh, cité dans Sutanto et al., 2013).

Ce qui constitue la vie privée du consommateur est déterminé contextuellement, individuellement, et culturellement selon Lanier et Saini (2008). Ainsi, au cœur de ce concept de confidentialité se trouvent **les inquiétudes** et les **risques perçus** en matière de vie privée. Bien qu'étant tous les deux fondés sur l'évaluation du risque, ces concepts sont différents. Les inquiétudes, d'une part, concernent l'anticipation des risques **en général** et sont donc un trait de personnalité, une tendance individuelle à être préoccupé au sujet de la vie privée, tandis que les risques perçus en matière de vie privée, d'autre part, concernent également l'anticipation individuelle des risques associés à la divulgation de l'information mais ce, dans une situation **spécifique** (Malhotra, Kim & Agarwal, 2004).

Sachant que l'utilisation d'informations personnelles peut être perçu comme un risque d'atteinte à la privée, il est logique de penser que ces inquiétudes et risques perçus auront une influence non négligeable sur l'intention comportementale du consommateur que nous étudions et donc sur l'utilisation des services personnalisés. Nous développerons ces facteurs ayant un potentiel impact sur la divulgation d'informations personnelles ci-dessous.

3.2 Inquiétudes en matière de protection de la vie privée

Au fur et à mesure que les entreprises intensifient leurs efforts pour recueillir et utiliser les données des consommateurs, ceux-ci s'inquiètent de plus en plus de la confidentialité de leurs données personnelles et du risque d'atteinte à la vie privée. Les préoccupations relatives à la protection de la vie privée sont aujourd'hui, **à l'ère du numérique, un des problèmes majeures** et la raison la plus souvent invoquée par les consommateurs pour refuser d'utiliser Internet ou un service spécifique (Chellappa & Sin, 2005 ; Dinev & Hart, 2006 ; Xu et al. 2010 ; Zhao, Lu & Gupta, 2012). Ainsi, ces craintes sont apparues comme le meilleur moyen de comprendre les sentiments des consommateurs à l'égard de la protection de leurs données personnelles et méritent donc d'être prises en compte dans notre modèle de recherche.

Les inquiétudes au sujet de la confidentialité ne sont pas nouvelles mais c'est principalement les avancées technologiques de traitement de l'information en combinaison avec la multiplication des failles de sécurité, devenues monnaie courantes ces dernières années, qui ont renforcé ce sentiment de vulnérabilité (Junglas et al., cités dans Alhouti, Johnson & D'Souza, 2016). Ces différents scandales et l'attention publique accrue envers ces problèmes de confidentialité ont en effet permis de conscientiser les individus aux enjeux qui pèsent sur les données, tels que les pratiques de publicité, les méthodes de collecte, l'utilisation et le transfert de données vers des parties tierces. Cependant, il semblerait qu'ils n'ont pas pour autant fait changer les comportements de ceux-ci, comme nous le constaterons plus tard (Aguirre et al., 2015).

Ce qui pose concrètement problème est le potentiel **comportement opportuniste des entreprises**, davantage présent avec l'accès généralisé aux données, notamment en matière de collection des informations personnelles, l'usage secondaire non autorisé, l'accès inapproprié à ces données du consommateur et les erreurs. Effectivement, Smith, Milberg et Burke (1996) ont élaboré l'échelle « *Concern For Information Privacy* » comprenant ces quatre dimensions, qui depuis lors est considérée comme un des instruments les plus fiables afin de mesurer les préoccupations des individus à l'égard des pratiques organisationnelles en matière de protection de la vie privée.

3.2.1 Inquiétudes relatives à la récolte et l'enregistrement des données personnelles

Premièrement, des craintes surgissent en raison des **pratiques de récolte de données des entreprises qui sont quelque fois discutables et perçues comme une atteinte à la vie privée**. Les consommateurs suspectent notamment que les entreprises collectent et traquent leur information privée entièrement à leur insu ou que les technologies récoltent les données sur base du principe de non-adhésion, c'est-à-dire que par défaut et sans l'approbation consciente du consommateur, les données seront enregistrées tant que le client ne s'y oppose pas explicitement (Lee & Cranage, 2011 ; Li & Unger, 2012 ; SAS, 2015).

En effet, comme nous l'avons évoqué dans le chapitre de la personnalisation, pour le profilage, les entreprises récoltent un maximum d'information au sujet des consommateurs, que ce soit de manière explicite ou implicite. Le problème se situe principalement dans cette pratique implicite, car cela peut porter atteinte à certains droits fondamentaux des consommateurs, tels que le **droit d'accès à l'information** et le **droit à la vie privée**, donc le droit de retenir de l'information à son propos et empêcher le partage de ses données (De Filippi, 2016). Or, ce contrôle est important. Si on se réfère au sondage de SAS, 33% des personnes interrogées estiment n'avoir aucun contrôle sur les informations personnelles qu'elles partagent avec les entreprises et 51% un peu de contrôle, seulement 7% juge avoir un contrôle complet (SAS, 2015, p.5). Par ailleurs, seulement 20% estiment que les entreprises sont ouvertes et transparentes en matière de règles et de traitement de l'information (SAS, 2015, p.6). Mis à part le fait que les entreprises offrent rarement cette transparence et ce contrôle souhaité, les individus n'ont souvent pas l'expertise ou les connaissances nécessaires pour contrôler leurs données ou les conséquences potentielles d'une violation de la vie privée. Cette **incapacité de se protéger et vérifier** la réalité des activités des entreprises en matière de collecte et traitement de données privées suscite, par conséquent, un sentiment de vulnérabilité générale auprès des consommateurs. Par ailleurs, les individus sont inquiets par **la quantité, parfois injuste, et la pertinence des données personnelles demandées, collectées et enregistrées** car ils estiment que l'échange n'est pas toujours équitable (KPMG, 2017).

3.2.2 Inquiétudes relatives à l'usage secondaire non-autorisé des données personnelles

Étant donné que les informations à propos des clients sont devenues un actif vital sur le marché en ligne, il y a toujours le danger qu'elles soient utilisées à d'autres fins, sans le consentement des propriétaires, que celles pour lesquelles elles ont été recueillies. Ceci constitue une atteinte à leur vie privée. Ainsi, il apparait que les consommateurs sont les plus préoccupés par la possibilité de **vente ou partage de leurs données à des parties tierces sans leur accord,** une pratique marketing stratégiquement importante dans de nombreuses industries, notamment pour obtenir des listes de diffusion (Treiblmaier & Pollach, 2007). L'usage secondaire des données fait également référence à l'utilisation des données pour des **campagnes marketing non désirées risquant le harcèlement, ou sollicitées mais très ciblées.** En effet, bien que de nombreuses recherches insistent sur les inquiétudes suite à des messages non sollicités, White, Zahay, Thorbjørnsen et Shavitt (2008) ont démontré que des messages sollicités peuvent également porter problème lorsqu'ils suggèrent un niveau inapproprié de familiarité avec les préférences et comportements du consommateur. Les consommateurs craignent que les communications sur base de leurs données soient trop personnelles, allant au-delà de la reconnaissance amicale.

Par ailleurs, les individus craignent **le spamming** ou que leurs données soient **utilisées à d'autres fins que celles du marketing,** moins légitimes comme l'utilisation commerciale abusive ou la surveillance gouvernementale (De Filippi, 2016 ; Karwatzki et al., 2017). KPMG (2017) mentionne qu'il y a un sentiment croissant que les citoyens sont **sous surveillance**, que les téléphones et navigateurs peuvent enregistrer l'information à propos d'eux sans qu'ils ne s'en rendent comptent. Ainsi, le fait d'envoyer des messages publicitaires contenant des données selon lesquelles il apparait clairement que le consommateur est observé, est très risqué. Il semblerait que les consommateurs répondent positivement à l'usage implicite de l'information personnelle par l'entreprise comme pour des

recommandations de produits mais n'aiment pas l'utilisation explicite de leurs données comme lorsqu'on les salue avec leur prénom, car cela augmente le sentiment d'intrusion (Karwatzki et al., 2017). Van Doorn et Hoekstra (cité dans Salonen & Karjaluoto, 2016) ont constaté le même effet pour des publicités en ligne. Cela dépend probablement des individus ou du support utilisé car l'étude de Shani, Wheeler et Chintagunta (2018) a montré qu'inclure le nom dans un e-mail est justement bénéfique. Bien que certains acceptent cette intrusion, estimant que c'est le prix de la commodité des offres adaptées, d'autres se posent la question de savoir si la fin justifie les moyens.

3.2.3 Inquiétudes relatives à l'accès non-autorisé et aux erreurs

Ensuite, ils sont également préoccupés par les **accès non-autorisés à leurs données** car cela pourrait causer des failles de sécurité, ainsi que par le manque **de mécanismes de protection** appropriés en cas d'erreurs dans l'enregistrement des données personnelles (Li & Unger, 2012).
Nous avons ainsi repris toutes les raisons pour lesquelles les individus pourraient ressentir une certaine vulnérabilité par rapport à la confidentialité générale des sites web, et examinerons maintenant les effets de ceux-ci sur les intentions comportementales.

3.2.4 Effets des préoccupations relatives à la vie privée

Il existe un consensus général dans la littérature selon lequel les traits de personnalités impactent, dans une certaine mesure, la sensibilité au risque. Effectivement, la tendance d'une personne à s'inquiéter de la protection de ses données personnelles influencerait la façon dont cette personne perçoit une situation particulière dans laquelle une entreprise sollicite ces données. Plus précisément, les internautes qui se préoccupent beaucoup de la protection de la vie privée sont susceptibles de percevoir beaucoup de risque au moment de l'évaluation de ceux-ci. Cette proposition est conforme à la théorie de l'action raisonnée, qui suggère que les caractéristiques individuelles influencent les croyances fondamentales (Malhotra, Kim & Agarwal, 2004 ; Dinev, Xu, Smith & Hart, 2013 ; Kehr, Kowatsch, Wentzel & Fleisch, 2015). En conséquence, nous supposons la relation de causalité suivante :

> *Hypothèse* : Les inquiétudes en matière de vie privée ont une influence positive sur les risques perçus de la personnalisation.

Nous verrons par la suite que la confiance en tant que croyance peut être considérée comme une contrepartie de ces inquiétudes et aurait, par conséquent, également un quelconque effet sur ce processus de prise de décision, mais positif (Bansal, Zahedi & Gefen, 2010).

3.3 Risques perçus

Le rôle des préoccupations relatives à la vie privée étant établi, nous nous devons désormais d'observer celui des risques perçus sur l'intention de divulguer des données en vue de bénéficier de la personnalisation. Bien que mesuré de manière relativement similaire, reprenant également les éléments de l'échelle de « *Concern For Information Privacy* » de Smith, Milberg et Burk (1996), ce concept et ses effets sont différents de ceux du trait de personnalité que nous venons d'aborder.

Précisons que dans ce contexte de protection de la vie privée, le risque perçu concerne la confidentialité des données. Ainsi, selon Kim, Ferrin et Rao (2008), ce risque lié à l'information est le

type de risque prédominant en ligne, décourageant les utilisateurs de divulguer leurs informations ou les encourageant à soumettre de fausses informations en ligne.
Ce concept est généralement défini comme étant « le degré selon lequel un individu estime qu'une perte éventuelle importante de confidentialité est associée à la divulgation de ses données personnelles » et ce, dans une situation précise (Xu et al., 2011, p.44 ; Karwatzki et al., 2017, p.372). Cela reflète les perceptions du consommateur quant à la façon dont la collecte, le stockage et l'utilisation des données personnelles, ou (le manque de) transparence ou de contrôle, l'affecte négativement dans un contexte en particulier (Smith , Milberg & Burke, 1996 ; Malhotra et al., 2004).

La perception de ce risque dépendra de l'évaluation de la vulnérabilité de chacun dans la situation spécifique de divulgation des données et du niveau de craintes de celui-ci, comme développé ci-après.

3.3.1 Inquiétudes et risques hétérogènes en matière de vie privée

Alhouti, Johnson et D'Souza (2016) affirment que la perception du consommateur de la protection de ses données est aussi hétérogène que les consommateurs eux-mêmes, car chacun évalue le rapport entre partage de données et risques liés à la vie privée d'une manière différente. Ceci est en accord avec la théorie *Communication Privacy Management* (CPM) également appelée *Information Boundary Theory* (IBT), qui explique les processus psychologiques selon lesquels un individu contrôle la protection de l'information privée. Celle-ci suggère effectivement que chaque consommateur définit ses propres frontières qui déterminent quelle information il est prêt à divulguer ou non à d'autres parties (Sutanto et al., 2013).

Les consommateurs peuvent notamment être catégorisés selon leur niveau d'inquiétude et par conséquent, leur **niveau de besoin de préserver leur information** afin d'établir leur espace personnel. Ainsi, selon la classification de Westin (cité dans Awad & Krishnan, 2006) on retrouve des « fondamentalistes » (25% de la population) qui sont extrêmement préoccupés par la protection de leur vie privée, percevront plus de risques et seront donc moins enclins, voire pas du tout, à partager de l'information personnelle et de participer aux services (de personnalisation par exemple), malgré les fonctionnalités de protection de la vie privée implémentées. Puis, il y a « les non-préoccupés » qui semblent dévoiler leurs données facilement (25% de la population) et enfin, « les pragmatistes » qui sont modérément inquiets et représentent la proportion de consommateurs la plus large (Taylor, cité dans Alhouti, Johnson & D'Souza, 2016). Outre cette tendance à être préoccupé par la protection de sa vie privée, certains chercheurs mentionnent que des personnes ayant un caractère de type « innovateur », en opposition au type « suiveur », percevront moins de risques pouvant les freiner à divulguer leur information (Xu et al., 2010).

Diverses études ont également établi que **l'éducation** et **le niveau d'utilisation d'internet** pourraient être liés au degré de préoccupation en matière de protection de la vie privée sur internet. Ainsi, les individus ayant un niveau d'éducation élevé seraient plus méfiants car ils consacreraient plus d'attention à la recherche de produits adaptés et d'information quant au traitement de leurs données et veillerons à mieux utiliser leurs données en ligne (Punj, cité dans Alhouti, Johnson & D'Souza, 2016 ; Hoy & Milne, 2010). En ce qui concerne l'utilisation d'internet, un usage fréquent signifierait que l'individu a acquis une certaine expertise et est plus sensible à la confidentialité et aux fonctionnalités pouvant

protéger ses données (Kobsa, 2007). Néanmoins, un usage élevé pourrait aller de pair avec un niveau de divulgation des données plus élevé mais plus responsable. En outre, bien qu'ils ne soient pas unanimes à ce sujet, certains chercheurs mentionnent l'influence de l'âge sur le niveau d'inquiétudes relatives à la vie privée. Les jeunes seraient moins préoccupés mais ne divulgueraient pas pour autant plus d'information personnelle en ligne. Au contraire, il semblerait qu'ils prennent plus de mesures de protection. La internautes âgés de 45 ou plus, seraient quant à eux, soit fortement préoccupé soit pas du tout (Sheehan, 2002 ; SAS, 2015).

Autre élément impactant les préoccupations en matière de la vie privée est **la culture** de l'individu**.** En effet, chaque culture accorde une importance différente à la vie privée selon Hofstede et son modèle de différenciation culturelle (Baruh, Secinti & Cemalcilar, 2017 ; Veltri, Krasnova, & Elgarah, 2011). Il s'agit alors d'identifier le niveau de tolérance d'exposition aux risques et de l'acceptation des situations ambigües, qui varie fortement au sien des populations, appelée l'évitement de l'incertitude ainsi que la dimension masculine ou féminine et le caractère collectiviste ou individualiste de la culture.

Outre les traits de personnalités et autres caractéristiques individuels, lors de l'évaluation des risques perçus, les individus peuvent également être influencés par le design du site web, par la réputation de l'entreprise, par des expériences antérieures avec la marque, par le type de données sollicitées, par le secteur auquel l'entreprise appartient, la confiance dans la marque, le dispositif utilisé afin de divulguer de l'information ou encore des attributs du site web comme des fonctionnalités de contrôle et transparence qui seront discutés plus tard (Dinev & Hart, 2006 ; KPMG, 2017 ; Hérault & Belvaux, 2014).

3.3.2 Effets de risques perçus sur le comportement du consommateur

De nombreuses recherches préalables comme celles de Chellappa et Sin (2005) ou Awad et Krishnan (2006) ont affirmé que la perception des risques joue un rôle majeur sur les intentions et éventuellement aussi sur les comportements de ceux-ci.

Premièrement, un niveau élevé de risques perçus **diminuerait l'intention de partager de l'information personnelle en ligne** dans le but de minimiser les potentiels effets négatifs (Capgemini Consulting, 2017).

En outre, plus les risques perçus sont grands, plus le consommateur peut avoir recours à diverses **mesures de protection de la vie privée** tels que fournir des fausses informations ou incomplètes aux entreprises, installer des logiciels anti-spam et cookies ou tout simplement refuser de partager de l'information (Sheehan & Hoy, 2000). Comme déjà mentionné, cela peut miner totalement les efforts des spécialistes du marketing en ligne (Wirtz, Lwin & Williams, 2007).

Enfin, une moindre volonté de communiquer de l'information personnelle suite à la perception de ces risques, diminuerait l'intention d'achat ainsi que la somme d'argent dépensée en ligne et impacterait négativement la confiance envers l'entreprise (Milne & Boza,1999 ; Li & Unger, 2012).

Etant donné le caractère négatif de ces effets, nous proposons la relation de causalité suivante :

Hypothèse : Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation).

Cependant, nous verrons dans les chapitres suivants que les individus prennent certains éléments en compte pouvant modérer cette relation entre les risques et l'intention de divulgation.

3.4 Paradoxe de la vie privée et *privacy calculus*

Bien que nous venons de supposer qu'il existerait un lien négatif entre les risques et l'intention de divulgation des consommateurs, les recherches ont, d'autre part, démontré à maintes reprises que ces perceptions et craintes ne sont pas toujours reflétées dans leur choix de gestion de la confidentialité. Au contraire, il apparait qu'ils finissent par divulguer une grande partie de leur information personnelle, que ce soit via les réseaux sociaux, sites internet ou applications mobiles (Hargittai & Marwick, 2016). Cette dichotomie entre l'attitude à l'égard de la protection de la vie privée et le comportement a été appelée *privacy paradox* ou paradoxe de la vie privée (Baruh, Secinti & Cemalcilar, 2017). En effet, il semblerait que les consommateurs, pourtant soucieux, finissent par divulguer une grande partie de leur information personnelle via les réseaux sociaux, sites internet ou applications mobiles par exemple (Hargittai & Marwick, 2016).

Nous allons à présent essayer de comprendre pourquoi, malgré les craintes, les individus finissent par vouloir partager des données personnelles et tenterons donc de déceler d'autres déterminants de l'intention de divulgation. Selon Barnes (cité dans Hargittai & Marwick, 2016), cela pourrait être dû à un manque de compétence ou de connaissances des différentes attitudes responsables possibles en termes de protection de la vie privée. Ensuite, il se pourrait que les consommateurs manquent de compréhension au niveau du fonctionnement d'un système de traitement des données ou des risques de confidentialité. Par exemple, de nombreux utilisateurs se connectent sur des sites web via les réseaux sociaux, sans se rendre compte jusqu'à quel point tout est traçable et que ces applications enregistrent et utilisent l'information de leur profil en entier (Li & Unger, 2012). A son tour, Acquisti (cité dans Kokolakis, 2017) indique que certains individus estiment qu'il n'est pas réaliste de se protéger contre toute intrusion possible de la vie privée, car il est impossible de ne pas laisser de trace sur internet. Ils doutent alors qu'adopter une stratégie stricte de protection de la vie privée finisse par porter ses fruits.

Bien que l'étude réalisée par Hargittai et Marwick (2016) confirme tous ces éléments, certains participants ayant fait preuve de connaissance et d'adoption de moyens pour se protéger de toute invasion à la vie privée, n'ont pas pour autant arrêté de divulguer de l'information. Par conséquent, la présence simultanée de ces éléments suggère que le paradoxe de la vie privée ne peut uniquement être dû à un manque de compréhension ou d'intérêt pour la protection de la vie privée.

Dans le cas de l'étude de Hargittai et Marwick (2016), qui étudie ce paradoxe dans le contexte des réseaux sociaux, il apparait qu'une raison pour ce manque de protection de la part des consommateurs est que ces médias, procurant de nombreux bienfaits tels que la satisfaction de socialiser et l'auto-expression, et ayant pris une telle importance dans la vie quotidienne des individus, diminuent la perception du risque d'être surveillé. En d'autres termes, tant que les consommateurs considèreront que les plateformes comme Facebook et Google sont nécessaires, les avantages l'emporteront sur les risques et la perte relative de la vie privée sera considérée comme le prix à payer en échange de l'accès gratuit.

Dès lors, on constate que la décision d'un individu de partager de l'information est basée sur l'évaluation de la valeur ressentie suite à cette divulgation dans un contexte spécifique. Il a effectivement été démontré que les individus effectuent ce qu'on appelle le ***privacy calculus,*** le calcul de confidentialité lorsqu'ils doivent prendre une décision en termes de divulgation d'information et que la protection de la vie privée entre en jeu (Culnan & Armstrong, 1999). Le résultat de cette **évaluation est fonction du ratio entre les bénéfices et risques perçus suite à la divulgation des données**, ce qui coïncide avec le concept de valeur perçue (Xu et al., 2011). Ce modèle est basé sur la théorie du comportement planifié, de l'échange social, du contrat social et de la maximisation de l'utilité. Ces relations contraires sur l'intention de partager des données personnelles, confirment l'utilité des hypothèses que nous avons posées précédemment concernant les risques et bénéfices perçus de la personnalisation.

Par conséquent, une des raisons principales pour cette incohérence entre inquiétudes et intention comportementale peut être que les risques perçus d'une possible intrusion de la vie privée sont considérés comme moins importants que les bénéfices reçus suite au partage de l'information (Baruh, Secinti & Cemalcilar, 2017). Autrement dit, que les retours de cette divulgation compensent le risque que leur vie privée soit compromise (Dinev & Hart, 2006 ; Culnan & Armstrong, 1999). Au contraire, les consommateurs auront l'intention de prendre des mesures de protection lorsqu'ils soupçonnent que leur information personnelle n'est pas protégée (Baek & Morimoto, 2012).

3.5 Conclusion

En conclusion, le respect de la vie est crucial pour les consommateurs dans le contexte actuel, où les données personnelles sont devenues une source de valeur clé pour les entreprises. Ce consommateur de plus en plus anxieux, n'est toutefois pas entièrement fermé à partager ses données en ligne. Selon le *privacy calculus*, les individus seront uniquement réticents à l'idée de divulguer de l'information personnelle s'ils s'attendent à ce que les résultats négatifs soient supérieurs aux gains espérés. **Il convient désormais de vérifier si les bénéfices de la personnalisation peuvent compenser les effets négatifs relatifs à la confidentialité,** tout comme c'est le cas pour les réseaux sociaux, qui eux ont l'avantage d'être plus hédoniques qu'utilitaires.

4. Paradoxe de la personnalisation et de la vie privée

Ce quatrième chapitre aborde le dilemme auquel un consommateur doit faire face lors la prise de décision de divulguer de l'information dans un contexte de personnalisation, ce qui constitue l'essence même de ce mémoire. Ensuite, nous tenterons d'identifier les facteurs pouvant influencer ce choix.

Si nous reprenons toute l'information quant aux deux sujets principaux que nous avons analysés auparavant, il s'avère que dans le cas de la personnalisation en ligne, les consommateurs nourrissent également un véritable paradoxe. D'une part, ils veulent profiter des bénéfices offerts par les nouvelles technologies et exigent des services ou produits personnalisés, adaptés à leurs besoins individuels, que les parcours soient fluides etc. D'autre part, ils désirent un relatif anonymat et un contrôle absolu de leurs données personnelles. Awad et Krishnan (2006, p.1) ont appelé ceci le « ***Personalization Privacy Paradox*** ». Le problème de cette attitude paradoxale des consommateurs est que la littérature a

clairement mis en évidence que la personnalisation, outre la capacité de l'entreprise à collecter les données, repose sur la volonté des consommateurs de partager leurs données personnelles (Chellappa & Sin, 2005). La question est alors de savoir si ces attentes sont conciliables.

4.1 L'évaluation du dilemme

Dans ce contexte, les consommateurs vont évaluer si le résultat de la divulgation de l'information personnelle dans le but d'être profilé en ligne (*to be profiled online*) pour un service de personnalisation vaut la peine de sacrifier ses données et risquer de recevoir des recommandations inappropriées. Autrement dit, il s'agit d'éventuellement trouver un compromis entre personnalisation et vie privée, que Chellappa et Shivendu (2007, p.6) ont nommé « *the personalization for privacy (P4P) ratio* ». Ce ratio peut être exprimé comme la valeur marginale des services personnalisés et des préoccupations en matière de protection de la vie privée, mesurée en termes de coûts de protection de l'information personnelle. Le *privacy calculus* représente une approche afin de mesurer ce compromis.

Comme expliqué dans la section concernant les risques perçus, cette évaluation est différente pour chaque consommateur et fonction du contexte. Dans le cas de la personnalisation, cela relève du **poids relatif** que le consommateur donne à la personnalisation et à la confidentialité, ce qui a son tour dépend de facteurs individuels spécifiques tels que le **seuil de protection de la vie privée**. Certaines personnes accorderont plus de valeur aux bénéfices de la personnalisation alors que pour d'autres, c'est la protection de la vie privée qui prime (Karwatzki et al., 2017). Cela dépend également de l'**attitude** générale d'un individu envers la divulgation des données et envers les nouvelles technologies.

Ensuite, la perception des bénéfices ou des risques peut découler du **type de données sollicitées**, ce qui est à nouveau très subjectif. Bien que selon l'étude de Chellappa et Sin (2005) tant les données identifiables directement ainsi que celles identifiables indirectement créent une certaine appréhension, une enquête récente indique que les consommateurs acceptent généralement le partage et l'utilisation de données en matière d'historique d'achats et parcours de navigation mais ont plus de mal avec des données liées à l'identité, préférences politiques, données liées à leur entourage etc. (Ryan, 2017). Selon Home et Home (cités dans Norberg, Horne & Horne, 2007) certaines personnes sont plus sensibles en ce qui concerne l'utilisation de données médicales, financières et familiales que pour des données concernant la consommation de produits et marques ou leur usage de médias. Xu et al. (2011) ajoutent que les risques perçus sont supérieurs lorsque les entreprises ont accès à des données de localisation, étant donné que la position en temps réel de l'individu est alors connue. Par ailleurs, il semblerait que l'intention de divulgation est liée au **type de personnalisation offert**, les avantages étant plus apparents pour des services personnalisés que pour la publicité en ligne par exemple (Awad & Krishnan, 2006).

Ces bénéfices ne sont pas uniquement tangibles comme des réductions monétaires mais peuvent aussi être intangibles (Chellappa & Sin, 2005). La valeur de la personnalisation en ligne dépend notamment principalement de si le produit ou service offert concorde avec les attentes de l'individu et aide à la prise de décision. Ainsi, Culnan et Bies (2003) ont introduit le concept de « deuxième échange » pour décrire cet échange non monétaire impliqué dans le *privacy calculus*, où les consommateurs divulguent des données personnelles en échange de services de meilleure qualité et offres

personnalisées. Ceci, en opposition au « premier échange » faisant référence à un échange utilitaire où des biens et services sont échangés contre de l'argent ou d'autres biens.

Les consommateurs sont donc plus ouverts à partager leurs données personnelles avec une entreprise dans un contexte de personnalisation si on leur offre une incitation. Au contraire, lorsqu'il n'y a pas de gain perçu pour le consommateur, la tolérance diminue fortement : 84% refusent que des applications mobiles aient accès à leurs contacts, photos et historique de navigation et 68% n'acceptent pas que les publicités en lignes exploitent le contenu d'e-mails privés (KPMG, 2017, para. 3).

Nous supposons que si les consommateurs continuent à utiliser les services personnalisés et les entreprises à les offrir, c'est qu'il y a une raison et que cela fonctionne. Bien que cela dépende des personnes et du contexte, l'étude de Chellapa et Sin (2005) a ainsi démontré que la personnalisation et la confidentialité sont deux concepts indépendants, et que **les bénéfices perçus de la personnalisation, l'emportent sur les craintes en matière de confidentialité**. Les consommateurs acceptent donc la possibilité de vulnérabilité dans l'espoir de recevoir une valeur ajoutée en échange ; des services plus personnalisées, du contenu plus pertinent, des offres spéciales, un gain de temps etc. (Xu et al., 2011 ; Awad & Krishnan, 2006). Certains parlent alors d'un contrat social implicite. C'est-à-dire que, dans le cadre de l'échange social d'information personnelle contre de la personnalisation, les consommateurs divulguent de l'information car ils estiment qu'ils peuvent faire confiance à l'entreprise qu'elle prendra ses responsabilités en terme de gestion adéquate de ces données et qu'elle ne partagera donc pas aveuglément ces informations personnelles (Chellappa & Sin, 2005 ; Xu et al., 2010 ; Xu et al., 2011).

4.2 Limites et extension du *privacy calculus*

Le problème de cette notion de calcul est qu'elle se base sur l'idée que les consommateurs agissent en tant qu'agents économiquement rationnels, effectuant un raisonnement cognitif en considérant les risques et bénéfices avant de divulguer de l'information et formant ainsi une perception de la valeur du service personnalisé (Chellappa et Shivendu, 2007 ; Awad and Krishnan, 2006). Or, le consommateur prend souvent des décisions en matière de protection de la vie privée dans un temps limité, sur base de réactions instinctives et intuitives car l'influence d'éléments environnementaux ou heuristiques de jugement peut être forte. Il calcule également selon un processus affectif (Shiv & Fedorikhin, cités dans Aguirre et al., 2015). Par exemple, il se peut que l'individu base sa prise de décision uniquement sur la confiance envers la marque, sur le design d'un site web ou son humeur du moment. Dans cette idée, Norberg, Horne et Horne (2007) estiment que lorsqu'un individu s'exprime par rapport à son intention de divulgation d'information, le risque jouera un rôle significatif dans sa réponse, alors que dans une situation réelle d'auto-divulgation, l'individu n'entamera pas de processus cognitif mais aura tendance à compter sur la confiance ou à divulguer de l'information automatiquement pour des demandes courantes, telles que la requête d'une adresse e-mail.

Similairement, il apparait que les individus ont tendance à mal évaluer les risques, surtout lorsqu'il s'agit de la protection de la vie privée. Effectivement, la vie privée étant un sentiment abstrait, les individus semblent avoir du mal à lui donner une valeur absolue (Xu et al., 2011). Ils n'ont pas accès à toute l'information nécessaire pour porter un jugement éclairé sur les compromis qui interviennent

dans les décisions relatives à la protection de la vie privée et donneront par conséquent, dans la majorité des cas, de l'information personnelle si des bénéfices sont fournis en échange. Il est en effet plus facile de percevoir les bénéfices puisqu'ils peuvent être immédiats, tandis que le risque de divulgation peut être invisible ou apparaître uniquement dans le futur (Acquisti,cité dans Xu et al., 2010).

Malgré l'importance croissante du *privacy calculu*s, nous constatons donc que la littérature souligne certaines faiblesses de ce modèle. Ainsi, il apparait que les décisions en matière de divulgation de l'information personnelle impliquent plus que la simple analyse coûts-bénéfices dont il est question dans la section ci-dessus (Morosan & DeFranco, 2015). **Des facteurs pourraient en effet renforcer ou atténuer ces relations, offrant une explication à ce paradoxe entre personnalisation et protection des données personnelles.** Conformément à cette logique, notre étude étendra ce modèle de calcul de confidentialité de base. Plus précisément, nous nous concentrerons davantage sur le lien de causalité négatif entre les risques perçus et l'intention de divulgation car nous estimons que les entreprises se focalisent moins sur la façon de diminuer les réticences à la divulgation, contrairement à leur forte attention portée à la mise en valeur des bénéfices. Dans cette perspective, la théorie de l'échange social, de la justice et du pouvoir mettent en avant l'effet de la confiance et de mécanismes d'information et de contrôle des données en tant que réducteur d'incertitudes et assurance d'un échange équitable. Nous nous attarderons sur ces modérateurs potentiellement présents lors de la prise de décision dans les sections suivantes.

4.3 La confiance

A maintes reprises, ce concept a été cité dans ce travail et nécessite donc des éclaircissements. Dans pratiquement tous les échanges, susciter la confiance apparait comme crucial. Ceci est d'autant plus vrai pour des interactions en ligne en raison de l'asymétrie de l'information, des risques de confidentialité et sécurité et, par conséquent, de l'incertitude du consommateur intrinsèque à Internet (Pavlou, 2003). Par ailleurs, la confiance est au cœur de toute relation, particulièrement lorsque celles-ci comportent un certain degré de risque ou d'incertitude et que le résultat de la décision est important pour l'individu (Doney & Cannon, 1997 ; Milne & Boza, 1999, Gefen, Karahanna & Straub, 2003 ; Krasnova, Spiekermann, Koroleva & Hildebrand, 2010).

Dès lors, la confiance est extrêmement pertinente et centrale dans un contexte de marketing relationnel et davantage dans celui de la personnalisation en ligne, où attirer de nouveaux consommateurs est difficile en raison des coûts de changement d'entreprise ainsi que les coûts élevés liés à une potentielle perte de confidentialité (Chau et al., 2013). Selon Aguirre et al. (2015), confiance et confidentialité sont effectivement étroitement liées. Le manque de confiance serait l'une des principales raisons pour lesquelles les consommateurs hésitent à divulguer de l'information en ligne (McKnight, Choudhury & Kacmar, 2002).

L''importance de la confiance dans ces relations peut s'expliquer par la théorie de l'échange social. Contrairement à l'échange purement économique, l'échange social, qui inclut le transfert de données intangibles comme dans le cas de la personnalisation, traite des situations où il n'y a pas de contrat explicite ou détaillé liant les parties ou lorsque le contrat est insuffisant pour fournir une protection juridique complète à toutes les parties concernées. Ces relations impliquent donc des obligations

futures non spécifiées et bénéfices non garantis, les parties sont vulnérables dans une certaine mesure et peuvent uniquement s'appuyer sur les liens sociaux. Par conséquent, la confiance, augmentant la certitude perçue concernant le comportement attendu de l'autre partie et réduisant la peur d'être exploité conséquent, est essentielle et détermine les attentes des individus à l'égard de la relation (Gefen, Karahanna & Straub, 2003 ; Chellappa & Sin, 2005).

En d'autres termes, **la confiance est une construction centrale dans des échanges impliquant l'interdépendance, l'incertitude, le risque et pouvant inclure un comportement opportuniste indésirable**. Ceci est le cas de la personnalisation en ligne, où l'on traitera de la confiance d'un individu envers l'entreprise.

4.3.1. Définition

Bien qu'il y ait un accord sur l'importance de la confiance, il y a un désaccord en ce qui concerne sa définition exacte car c'est un concept multidimensionnel. Dans la littérature marketing, elle est généralement considérée comme une attitude ou une croyance. Ainsi, Moorman, Deshpande et Zaltman (cités dans Norberg, Horne & Horne, 2007, p. 107) définissent la confiance comme « la volonté de s'appuyer sur un partenaire d'échange ». Dans un contexte social, la confiance a également été caractérisée par « la volonté de se placer en position de risque » et donc de devenir vulnérable à l'entreprise en ligne (Milne & Boza, 1999, p.7). Puis, Doney et Cannon (1997, p.36) définissent la confiance comme « la crédibilité perçue et la bienveillance de la cible de confiance ». Selon Pavlou (2003, p.106) la confiance fait référence à « la croyance que l'autre partie se comportera d'une manière socialement responsable et, par conséquent, répondra aux attentes de l'individu sans tirer parti de ses vulnérabilités ».
On constate que les deux premières définitions mettent l'accent sur les croyances et l'attitude de la partie qui fait confiance, tandis que les deux suivantes soulignent les attributs de l'autre partie. Prises ensemble, les définitions supposent implicitement que la partie qui fait confiance est confrontée à l'incertitude et à l'inquiétude quant aux résultats possibles d'un échange avec le partenaire et qu'il y a des résultats positifs à atteindre grâce à la confiance. Dès lors, dans notre contexte de personnalisation, nous définissons la confiance comme l'espoir du consommateur de pouvoir compter sur le fait que l'entreprise traite son information personnelle dans le respect de la vie privée et par conséquent, comme la volonté d'assumer les risques de la divulgation.

Par ailleurs, considérant la confiance comme un concept tridimensionnel, Mayer, Davis et Schoorman (1995) ont proposé un modèle afin de mieux comprendre cette notion. Ce modèle promeut l'utilisation de trois facteurs de fiabilité perçue, notamment la compétence, la bienveillance et l'intégrité. Ceci démontre que les individus en ligne évaluent les entreprises non pas en termes généraux mais en termes d'attributs spécifiques. La compétence correspond à la capacité de l'entreprise et l'expertise nécessaire afin de répondre aux besoins du consommateur (crédibilité, réputation et savoir-faire), l'intégrité est la capacité de l'entreprise à tenir ses promesses, à être cohérent et honnête, et la bienveillance signifie que l'entreprise est disposée à agir dans l'intérêt du consommateur et d'éviter de faire quelque chose de préjudiciable, en opposition à l'opportunisme.

4.3.2. Effets modérateurs

La littérature a porté beaucoup d'attention à la confiance en ligne, ayant incorporé cette dernière dans de nombreuses études sur le comportement des consommateurs. Ces études ont notamment montré l'influence négative de la confiance sur les risques perçus, l'impact positif sur l'achat en ligne, sur l'intention d'acheter ou sur la divulgation de l'information en vue de la transaction, ainsi que l'impact sur le nombre de clics et sur l'acceptation de la publicité (Martin & Murphy, 2017).

En ce qui concerne le contexte de la vie privée, le **lien entre la confiance, le risque et la vie privée** a été régulièrement souligné. Bansal, Zahedi et Gefen (2010) ont également constaté que la confiance jouerait un rôle plus important sur la divulgation d'information que les **préoccupations** relatives à la protection de la vie privée. Par ailleurs, plusieurs auteurs affirment que la confiance est une des raisons pour lesquelles on observe une telle dichotomie entre les préoccupations exprimées en matière de vie privée et la divulgation d'information personnelle (Acquisti & Gross cités dans Krasnova et al., 2010 ; Norberg, Horne & Horne, 2007). Toutefois, **le rôle exact de la confiance dans ce contexte n'est pas encore clair** car elle n'a pas été modélisée de façon uniforme dans la littérature ; certains considèrent la confiance comme un antécédent ou le résultat des risques perçus, d'autre comme une variable indépendante des risques pouvant influencer la divulgation de l'information. Malgré ce manque d'uniformité, Gefen et al. (cités dans Krasnova et al., 2010) mentionnent que, dans toutes les situations où le risque est inhérent à une activité, la confiance servira de stratégie de réduction du risque et le risque, à son tour, aura un impact direct sur le comportement. Ceci part du principe que, suivant la théorie de l'action raisonnée, la confiance favorise des attitudes positives envers l'entreprises en ligne, ce qui impactera la perception des risques et par la suite le comportement. Les études de Malhotra, Kim & Agarwal (2004), Kim et al. (2008) et Krasnova et al. (2010) ont souligné ce même lient entre confiance, risque et intention de divulgation. Suivant cette logique, nous établirons que les risques perçus, étant influencé par les préoccupations, pourraient être atténués, voire surmontés, si l'individu a une confiance préétablie dans le site web. Dans le cadre du *privacy calculus*, nous attribuons donc un effet modérateur à la présence de la confiance et établissons l'hypothèse suivante :

> *Hypothèse* : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données sera moins forte si le consommateur a confiance dans l'entreprise en ligne.

Etant donné que les entreprises peuvent faire que très peu afin de mitiger les incertitudes du consommateur en matière de confidentialité, le fait d'avoir la possibilité de l'influencer de manière indirecte via le développement de confiance peut être intéressante pour celles-ci.

Mis à part cet effet pertinent pour notre recherche, la confiance **crée aussi les coûts de changement de fournisseur**, ce qui augmente la probabilité que le client continue à entretenir une relation avec l'entreprise traitant équitablement l'information du client (Culnan & Armstrong, 1999).

Enfin, bien que la confiance et le respect de la vie privée soient deux concepts proches, ils diffèrent dans leurs effets sur la personnalisation. Le respect de la vie privée est considéré comme étant l'antécédent de la confiance et ainsi, de l'intention d'adoption de la personnalisation (Li & Unger, 2012). Alors que **la confiance serait un antécédent et un résultat de la personnalisation,** ces concepts étant constamment positivement liés (Komiak & Benbasat, 2006).

4.4 Autres mécanismes d'atténuation

Pour déterminer d'autres facteurs pouvant atténuer l'impact des inquiétudes et risques en matière de protection de la vie privée sur la divulgation d'information, nous allons nous baser sur la théorie de la justice et du pouvoir. Selon ces théories, il ne suffit pas que les organisations offrent à leurs clients des avantages attrayants pour encourager la divulgation des données personnelles. Elles devraient également **fournir un contexte équitable et digne de confiance** pour la collecte et l'utilisation de ces données afin que les consommateurs perçoivent la divulgation comme une proposition à faible risque (Zhao, Lu & Gupta, 2012).

Sur base de la littérature existante, nous avons en effet pu constater qu'un des problèmes de la divulgation de l'information en vue de la personnalisation est qu'il existe une certaine inégalité entre les contributions et les bénéfices reçus des utilisateurs, par rapport à ceux de l'entreprise. Il est évident que lorsqu'un service est offert gratuitement, tel que la personnalisation, cela signifie souvent que c'est basé sur l'exploitation des données des consommateurs. Ainsi, Tim Cook, CEO d'Apple, affirmait « *when an online service is free, you're not the customer. You're the* product » (Hackett, 2015, para. 3). Toutefois, la réelle difficulté pour le consommateur est que cette inégalité semble parfois inévitable car l'entreprise n'aide pas dans la prise de décision, elle ne propose aucune alternative ou explication, par exemple.

Afin de réduire ces inégalités, il existe des pratiques de rééquilibrages. Un consommateur peut, par exemple, quitter une relation afin de rééquilibrer le pouvoir, donner de fausses informations personnelles, faire du bouche-à-oreille négatif etc. avec de lourdes conséquences pour les sociétés (Mosteller & Poddar, 2017). Cependant, les entreprises pourraient elles-mêmes intervenir de manière proactive afin d'assurer cet équilibre, en adoptant **des stratégies éthiques et équitables du traitement de l'information**. Il semblerait en effet que la **perception des consommateurs quant à l'éthique et l'équité peut influencer leur volonté de divulguer des données personnelles,** tout en réduisant simultanément les comportements négatifs précités (Culnan & Bies, 2003).

Dès lors, ces théories conseillent aux entreprises d'adopter des approches d'intervention en matière de protection de la vie privée et donc d'atténuation des risques perçus, adhérant aux trois types de perception de la justice (Culnan & Bies, 2003) :

- *la justice distributive :* La perception de cette justice se réfère à l'évaluation des avantages que les consommateurs tirent du partage de leurs données par rapport à leurs contributions, c'est-à-dire la mise en danger de leur vie privée.
- *la justice procédurale :* Cela concerne **l**'équité perçue des politiques et des procédures de protection de la vie privée.
- *la justice interactionnelle :* C'est l'évaluation de la qualité reçue du traitement interpersonnel, comprenant l'honnêteté et le respect des promesses.

Nous nous concentrerons sur la justice procédurale car la justice distributive est fortement liée au concept de *privacy calculus* et suppose ainsi que les bénéfices perçus de la personnalisation sont des facteurs pouvant encourager la divulgation des données. Quant à la justice interactionnelle, cela

reprend des éléments du concept de la confiance dont nous avons également déjà discuté (Martin & Murphy, 2017).

Suivant le principe de la justice procédurale, les pratiques de rééquilibrage seront représentées dans notre étude par le **contrôle des données** et la **transparence à propos du traitement des données**. Ces mécanismes pourraient, selon certains chercheurs, surmonter les craintes et ainsi atténuer les risques perçus car ils augmentent la perception d'équité des pratiques de l'entreprise. L'importance du contrôle et de la transparence est également établie dans la théorie de la réactance psychologique, qui propose que les consommateurs résistent à la restriction de leurs choix (Beke, Eggers & Verhoef, 2018). La transparence et le contrôle en tant que mécanismes d'atténuation des risques ont été déduits de la définition de la confidentialité (*privacy*).

En effet, rappelons que la confidentialité est « la capacité d'un individu à contrôler les termes selon lesquels l'information personnelle est recueillie et utilisée » (Westin, cité dans Chellappa & Sin, 2005, p. 186). A partir de cette définition, nous pouvons interpréter qu'une manière d'atténuer les risques perçus en matière de respect de la vie privée est d'augmenter le niveau de contrôle d'un consommateur sur l'information personnelle (Awad & Krishnan, 2006). Ce concept de contrôle est composé de deux éléments clés : le contrôle actif et passif. Autrement dit, ce contrôle effectif du consommateur dépend tant de la conscientisation de la collecte et de l'usage des données que de la capacité à influencer ceux-ci (Beke, Eggers & Verhoef, 2018). Dès lors, si les entreprises veulent respecter la vie privée des consommateurs, elles devraient mettre en œuvre des pratiques expliquant quelles informations elles collectent, comment elles stockent ces informations et à quelles fins elles les utiliseront (transparence ou mécanisme de notification). De plus, les entreprises devraient permettre aux consommateurs de pouvoir refuser la collecte et l'utilisation des informations ainsi que de pouvoir exiger la suppression de leurs données enregistrées (contrôle ou mécanisme choix).

Bien que cette notion de contrôle soit incorporée dans la plupart des arguments et définitions de la confidentialité, la transparence et les fonctionnalités de contrôle peuvent également être conceptualisés comme des variables liées mais distinctes des risques perçus en matière de confidentialité. Ceci est en accord avec la conceptualisation de Laufer et Wolfe (cités dans Xu et al., 2008) qui ont considéré ce contrôle, actif et passif, comme étant une variable médiatrice dans le système de la protection de la vie privée. Dès lors, nous supposons que ce concept de contrôle effectif, comprenant la transparence et les solutions technologiques de contrôle, sont deux approches principales pouvant être utilisées par des organisations afin d'assurer une certaine équité et ainsi résoudre la pression conflictuelle entre le désir de recevoir de la personnalisation et celui de garder son information privée. Nous tenterons de mieux comprendre la nature et les effets de ces deux concepts distincts dans les prochaines sections.

4.4.1. La transparence

Nous avons pu constater que les chercheurs mentionnent régulièrement que les obstacles et les incohérences dans la prise de décisions en matière de protection de la vie privée peuvent être dus, du moins en partie, à l'**asymétrie de l'information**. Effectivement, tandis que les utilisateurs sont de plus en plus transparents en ligne car ils ont tendance à divulguer beaucoup d'information, les entreprises ne le sont pas. Celles-ci tiennent, au contraire, les clients dans l'ignorance en utilisant des pratiques et

technologies invisibles et opaques (Portes, 2018). Ainsi, les consommateurs qui sont confrontés à des décisions sensibles en matière de protection de la vie privée ne savent pas, ou que très peu de choses sur la façon dont leurs données sont recueillies et utilisées, et avec quelles conséquences. Ce manque de connaissance provoque le sentiment d'être manipulé et une réticence à divulguer de l'information (Malhotra, Kim & Agarwal, 2004). Ce défi a été principalement attribué au manque de transparence, et plus particulièrement aux politiques de confidentialité qui ne communiquent pas efficacement les pratiques de traitement de l'information et les risques d'atteinte à la vie privée pour les consommateurs (Adjerid, Acquisti, Loewenstien & Brandimarte, 2013).

4.4.1.1. Définition du concept

Contrairement au contrôle, la transparence est une **dimension passive de la protection de l'information**. Elle reflète la mesure selon laquelle une personne est informée des pratiques de traitement de l'information (Malhotra, Kim & Agarwal, 2004). Par ailleurs, pour Martinez (cité dans Portes, 2018, p. 33) « la transparence ne décrit pas une action mais un état de parfaite communication entre le spectateur et le phénomène qui lui apparaît malgré l'obstacle ». En effet, c'est un droit de savoir comment et pourquoi les données privées sont recueillies, quels tiers seraient susceptibles d'avoir accès à ces données demandées, les utilisations prévues, les mesures qui seront prises pour protéger la confidentialité, les conséquences de fournir ou de retenir de l'information, et tout moyen de recours disponible pour l'individu (Malhotra, Kim & Agarwal, 2004). Etre mis au courant que l'information personnelle sera partagée ou non avec d'autres entités serait le facteur ayant le plus d'influence sur la perception des risques et la divulgation de l'information en ligne (Cranor, Reagle & Ackerman, 1999).

4.4.1.2. Dépasser le stade de conformité

À l'heure actuelle, la méthode principale par laquelle les sites web communiquent cela, est la publication de politiques de confidentialité, entre autres car c'est une obligation légale. Le problème est que seulement 0,5 à 1% des utilisateurs font l'effort de lire ces déclarations, la longueur et la complexité de ces déclarations étant un obstacle à la prise de connaissance de ceux-ci (Kobsa, 2007). Par ailleurs, la difficulté à trouver ces politiques lorsque les consommateurs en ont besoin et le langage trompeur utilisé, amènent à se demander ce que l'entreprise cache.

Ensuite, les technologies ayant évolué plus rapidement que les législations, cela laisse beaucoup de liberté aux entreprises en ce qui concerne la collecte et le traitement des données. Le dernier texte concernant la protection des données personnelles avant le RGPD datait notamment de 1995 alors qu'internet était encore à ses débuts (Hargittai & Marwick, 2016 ; Garcia-Rivadulla, 2016). En conséquence, la simple présence des politiques de confidentialité, en tant qu'approche de transparence, n'est pas suffisante pour conscientiser les internautes sur ces pratiques et l'aider dans sa prise de décision en matière de protection de la vie privée. Il est nécessaire de dépasser le stade de conformité en étant réellement engagé dans la protection des données de ses consommateurs.

Nombreux sont donc les appels pour plus de transparence afin de résoudre l'équation personnalisation versus confidentialité (Kobsa, 2007 ; Sutanto et al., 2013 ; Awad & Krishnan, 2006). Ainsi, Malhotra, Kim et Agarwal (2004) estiment qu'augmenter la conscientisation du consommateur en notifiant clairement et mettant en valeur les pratiques de traitement de l'information est un élément clé de la confidentialité

en ligne. Cela veut dire qu'il est nécessaire de **fournir l'information pertinente de manière intelligente**, en améliorant la lisibilité de ces pratiques et en mettant cette information à disposition des consommateurs **au moment et endroit approprié** (p.ex. moment de la divulgation). La transparence signifie également qu'il faut alerter de manière proactive que les consommateurs sont traqués (ex. cookies) et présenter des arguments pour encourager la divulgation, c'est-à-dire les avantages obtenus en échange de l'information fournie (Martin & Murphy, 2007). Mise à part la supposition que cela puisse être bénéfique pour le client, cette mise en valeur des règles en matière de respect de la vie privée pourrait également être un atout concurrentiel pour l'entreprise lui permettant d'améliorer sa réputation.

Dans cette idée de présentation des pratiques de l'entreprise de manière explicite et contextuelle, un simple logo qui indiquerait que les données sont utilisées pour des fins marketing, par exemple, avertirait le consommateur. Puis, une infobulle au moment de la requête d'information ainsi qu'une phrase comme « pourquoi reçois-je cette publicité ? » menant vers une page expliquant le ciblage est plus parlant que de l'inclure dans des déclarations de confidentialité standards.

4.4.1.3. Effets modérateurs

Cet enjeu de la transparence est au cœur des débats actuels car il est communément admis qu'un défaut de transparence peut entraîner des comportements de méfiance et même de résistance de la part des clients. Cela suit le principe que les gens ne prennent des décisions éclairées que si une information adéquate leur est fournie. Etant donné que les mécanismes d'amélioration de la transparence cités-ci dessus sont destinées à solutionner ces asymétries de l'information, cela accroit l'équité procédurale perçue et par conséquent, atténue les risques perçus en matière de protection de la vie privée, tout en favorisant la réciprocité (Treiblmaier & Pollach, 2007). Nous développons par conséquent l'hypothèse suivante :

Hypothèse : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données.

Nous avons posé l'hypothèse que la transparence atténuerait la relation entre les risques perçus et la divulgation selon l'idée que conscientiser les consommateurs peut les rassurer. Cependant, il est important de noter que dans certaines études, cet effet n'était pas significatif (Karwatzki et al., 2017). Cela pourrait être dû au fait que les consommateurs ne considèrent pas la transparence comme utile car ils ne peuvent pas être certains que les entreprises respectent leur engagements en matière de protection de la vie privée. Par ailleurs, un effet contraire de la transparence pourrait être justifié par la dualité des caractéristiques de la transparence : l'information donne un signal d'équité mais rend simultanément les problèmes de vie privée explicites. Ceci suppose qu'étant plus averti en ce qui concerne la quantité d'information collectée et la manière dont elles sont utilisées, les individus pourraient finalement être plus méfiants et réticents à la divulgation d'information (Portes, 2018).

4.4.2. Le contrôle

Comme déjà mentionné, certains individus considèrent la divulgation de l'information en vue de bénéficier de la personnalisation comme un échange inégal. En effet, le problème avec la personnalisation en ligne est que c'est un processus qui place la responsabilité de la protection des données personnelles entre les mains de ceux qui recueillent, utilisent et vendent des données personnelles, éventuellement sans que le consommateur s'en aperçoive, causant des inquiétudes de confidentialité auprès du consommateur (Xu, Teo, Tan & Agarwal, 2009). Les fonctionnalités de contrôle forment alors une **solution afin de rétablir la perception de l'équité dans l'échange et la relation avec l'entreprise.** C'est un acte de rééquilibrage car, en cédant les commandes de gestion de sa confidentialité à l'utilisateur, l'entreprise renonce de façon proactive à un certain pouvoir. Cette solution équitable part du principe que « le pouvoir n'est pas une propriété de l'organisation, mais de la relation » et apparait comme le moyen le plus efficace afin de protéger le consommateur car il fait lui-même partie de la première ligne de défense (Palmatier, Stern, and El-Ansary cités dans Mosteller & Poddar, 2017, p.29).

4.4.2.1 Définition du concept

Une définition que nous estimons appropriée pour cette variable de contrôle est celle de Treiblmaier et Pollach (2007, p.3) : « Le contrôle se réfère à la mesure dans laquelle les utilisateurs peuvent influencer les données recueillies à leur sujet ou peuvent choisir de rejeter la récolte de leurs données », en précisant qu'une modification ou un rejet peut être fait à n'importe quel moment. Par ailleurs, Alge (cité dans Malhotra, Kim & Agarwal, 2004) souligne l'importance de la liberté d'accepter ou rejeter les pratiques de l'entreprise en matière d'information.

En pratique, ceci donne la possibilité à un consommateur d'activement contrôler les paramètres de confidentialité d'un site, ayant le choix d'approuver, rejeter ou limiter le type et la quantité d'information demandée ainsi que de préciser ses préférences en ce qui concerne l'utilisation de ses données futures et l'accessibilité de celles-ci à autrui (*opt-in*). Les fonctionnalités de contrôle fournissent aussi, par la suite, l'option d'avoir accès à l'information détenue par l'entité à son sujet, de vérifier et éventuellement modifier cela (*opt-out*).

Tout comme pour la transparence afin d'assurer des effets positifs, ces fonctionnalités se présentent idéalement sous forme de **paramètres de confidentialité visibles, granulaires et faciles à comprendre, proposés au moment de la requête d'information** personnelle et accessible à n'importe quel moment par la suite (Sheehan & Hoy, 2000). Elles doivent simultanément bien s'intégrer dans le design d'une page web afin de ne pas être trop envahissantes (Alekh, 2017).

Cette notion de contrôle est inhérente au *permission marketing* et suit la logique de *l'empowerment*. La demande de permission d'utilisation est effectivement primordial car nous avons constaté qu'un consentement implicite ne suffit pas, les consommateurs se sentant fortement vulnérables lorsqu'ils réalisent que de l'information à leur propos a été collectée sans leur accord. Effectivement, les selon une étude de Microsoft, 84% des consommateurs souhaitent avoir l'option d'explicitement accepter ou non les conditions, tout en étant informé, et par la suite également avoir la possibilité *d'opt-out* à tout moment (Sterling, 2015, para. 8). Dans cette perspective, il serait opportun d'obliger les entreprises à

proposer le niveau maximum de confidentialité par défaut et non pas comme sur Facebook, où l'on autorise le partage par défaut et doit désactiver manuellement les autorisations pour chaque type de données (Garcia-Rivadulla, 2016).

4.4.2.2 Effets modérateurs

En ce qui concerne les effets de cette fonctionnalité, les recherches antérieures ont affirmé qu'en général, les personnes perçoivent moins de risques d'atteinte à la vie privée lorsqu'elles estiment être en mesure de contrôler la divulgation et l'utilisation subséquente de leurs données (Xu, Parks, Chu & Zhang, 2010; Tucker, 2014). Par ailleurs, le fait de donner aux utilisateurs, les moyens d'agir témoignent de l'engagement d'une entreprise à protéger la vie privée des consommateurs. Ceci réduit l'incertitude et la perception d'opportunisme (Krasnova et al., 2010 ; Zhang, Wang & Jin 2014 ; Mosteller & Poddar, 2017). Compte tenu de cette relation négative entre le contrôle perçu et les risques en matière de protection de la vie privée suggérée par des recherches antérieures, nous supposons que cela rendra le consommateur plus disposé à partager ses données et considérons donc le contrôle comme une variable modératrice. Ainsi, nous proposons l'hypothèse suivante :

> *Hypothèse* : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données sera moins forte si le consommateur perçoit du contrôle sur ses données.

Toutefois, Adjerid et al. (2013) estiment que dans la pratique, ce sentiment intense de contrôle sur la publication des données personnelles peut paradoxalement entrainer des divulgations accrues et plus risquées.

Enfin, Martin, Borah & Palmatier (2017) insistent sur la **présence simultanée de ces deux éléments** car lorsqu'une entreprise offre beaucoup de transparence mais peu de contrôle, les consommateurs perçoivent plus de violation que de sûreté. Il semble donc qu'expliquer exactement au consommateur comment l'entreprise collecte et utilise ses données sans qu'il puisse activement agir sur ces pratiques est très dangereux. D'un autre côté, la combinaison de peu de transparence des données avec beaucoup de contrôle crée une situation d'autonomie non informée. Les clients ont alors la possibilité de changer leurs préférences, donc ils répondent favorablement, mais leurs choix d'*opt-in* et d'*opt-out* sont quelque peu aveugles car ils n'ont pas eu de connaissance complète de ce que l'entreprise fera de leurs données.

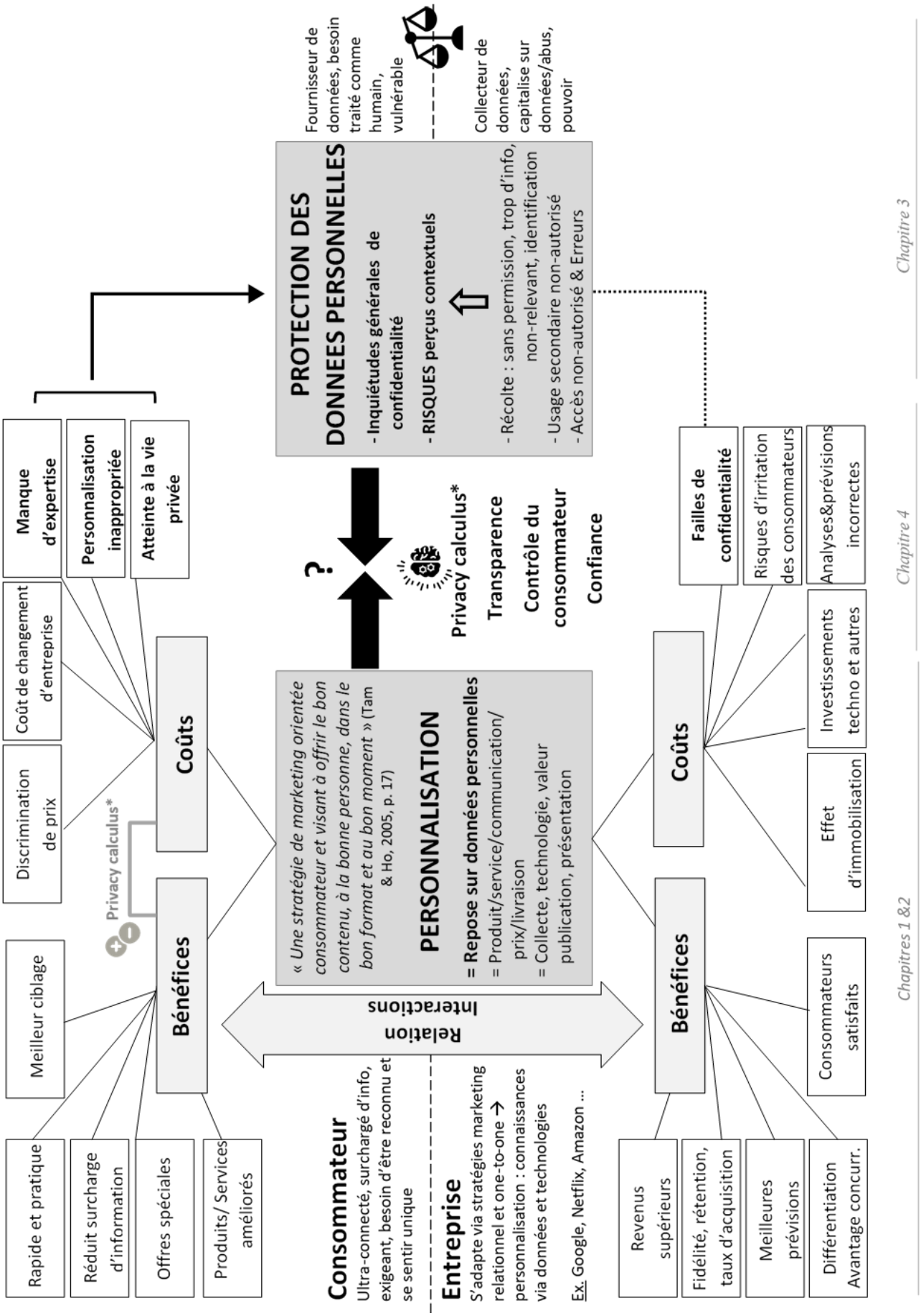
4.5. Conclusion

En conclusion, bien que peu d'études ont appliqué le *privacy calculus* à un contexte de personnalisation, l'étude de Chellappa et Sin, qui se rapproche le plus de notre recherche, a montré que la valeur de la personnalisation l'emportent sur les craintes en matières de vie privée. Cependant, ces résultats ne sont pas nécessairement généralisables car la littérature a également souligné que l'évaluation des risques et bénéfices constituant ce calcul de confidentialité dépend fortement du contexte et d'éléments environnementaux. Il est donc nécessaire d'approfondir ces recherches, tout en incorporant d'autres variables car ce modèle de base possède des limites.

Ainsi, plusieurs études identifient des facteurs venant modérer l'effet des risques perçus. La confiance en tant que variable modératrice trouve sa source dans les théories de l'échange social. De ce fait, il apparait que la confiance intervient lorsqu'un consommateur perçoit de l'incertitude et des risques de résultats négatifs au cours d'un échange. S'il place une quantité considérable de confiance dans une entreprise, cela aidera à atténuer son évaluation négative des risques dans une situation donnée, réduisant par conséquent sa résistance à partager de l'information. Quant à la théorie de la justice et du pouvoir, il apparait que les mécanismes de contrôle et de transparence peuvent générer un fort sentiment d'autonomisation et d'équité, et ce davantage lorsqu'ils apparaissent en combinaison. Ceci peut aider à la prise de décision rationnelle et modérer l'effet des risques perçus provoqués, entre autres, par leur sentiment général de vulnérabilité.

5. Conclusion de la littérature

Figure n°1 : Résumé et vue schématique de la revue de la littérature



6. Elaboration du modèle de recherche

La revue de littérature, comprenant des études effectuées antérieurement dans le domaine du marketing et des applications technologiques, nous a permis de mieux comprendre le processus de divulgation des données personnelles en ligne en vue de bénéficier de la personnalisation et d'étayer l'objet de notre recherche. Ainsi, nous avons pu identifier les facteurs clés encourageant ou entravant cette prise de décision, à savoir les bénéfices et risques perçus de la personnalisation, formant le *privacy calculus*. Les chercheurs exposent également d'autres éléments secondaires jouant un rôle dans ce processus, notamment les inquiétudes en matière de vie privée, la confiance, la transparence et le contrôle des données par le consommateur. Dans ce chapitre, sur base de ces éléments, nous présenterons et expliquerons le cadre conceptuel de notre étude.

5.1. Questions de recherche et hypothèses

Pour rappel, la question que nous avons choisie pour notre mémoire est la suivante:

Dans quelle mesure des mécanismes d'atténuation des risques de confidentialité, tels que la confiance, la transparence et le contrôle des données, ont un influence sur l'intention du consommateur de partager des données personnelles en vue de bénéficier de la personnalisation en ligne ?

Les études effectuées auparavant nous ont amenés à développer des sous-questions de recherche exposées ci-dessous, nécessaires à l'élaboration de notre modèle de recherche. Les hypothèses associées sont des réponses possibles à ces questions, qui seront testées statistiquement dans la deuxième partie de ce travail.

1. ***Quels sont les déterminants (motivations et freins) de l'intention de divulgation des données personnelles en ligne en vue de bénéficier de la personnalisation? Certaines données sont-elles plus facilement divulguées que d'autres ?***

- *Dans quelle mesure les bénéfices perçus de la divulgation influencent-ils l'intention de la divulgation de données personnelles en ligne en vue de bénéficier de la personnalisation?*

  Hypothèse 1 : Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation).

  H1.1 Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **explicites** (nom, âge, adresse...).

  H1.2 Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **implicites** (Géolocalisation, parcours de navigation, adresse IP…).

- *Dans quelle mesure les risques perçus de la personnalisation influencent-ils l'intention de divulgation en vue de bénéficier de la personnalisation?*

  Hypothèse 2 : Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation).

  H2.1 Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles **explicites** (nom, âge, adresse...).

  H2.2 Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles **implicites** (Géolocalisation, parcours de navigation, adresse IP…).

2. Quels éléments personnels et fournis par l'entreprise peuvent influencer la perception des risques de la personnalisation ?

- *Dans quelle mesure les inquiétudes d'un consommateur liées à la confidentialité influencent-elles les risques perçus en matière de privée de la personnalisation?*

  Hypothèse 3 : Les inquiétudes en matière de vie privée ont une influence positive sur les risques perçus de la personnalisation.

- *Dans quelle mesure le niveau de confiance du consommateur en l'entreprise fait-il varier l'impact des risques perçus de la personnalisation sur l'intention de divulgation ?*

  Hypothèse 4 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur a confiance dans l'entreprise en ligne (modération).

  H 4.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur a confiance dans l'entreprise en ligne

  H 4.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur a confiance dans l'entreprise en ligne

- *Dans quelle mesure le niveau de contrôle du consommateur fait-il varier l'impact des risques perçus de la personnalisation sur l'intention de divulgation ?*

  Hypothèse 5 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur perçoit du contrôle sur ses données (modération).

  H 5.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur perçoit du contrôle sur ses données.

  H 5.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur perçoit du contrôle sur ses données.

- *Dans quelle mesure le niveau de transparence observé par le consommateur fait-il varier l'impact des risques perçus de la personnalisation sur l'intention de divulgation ?*

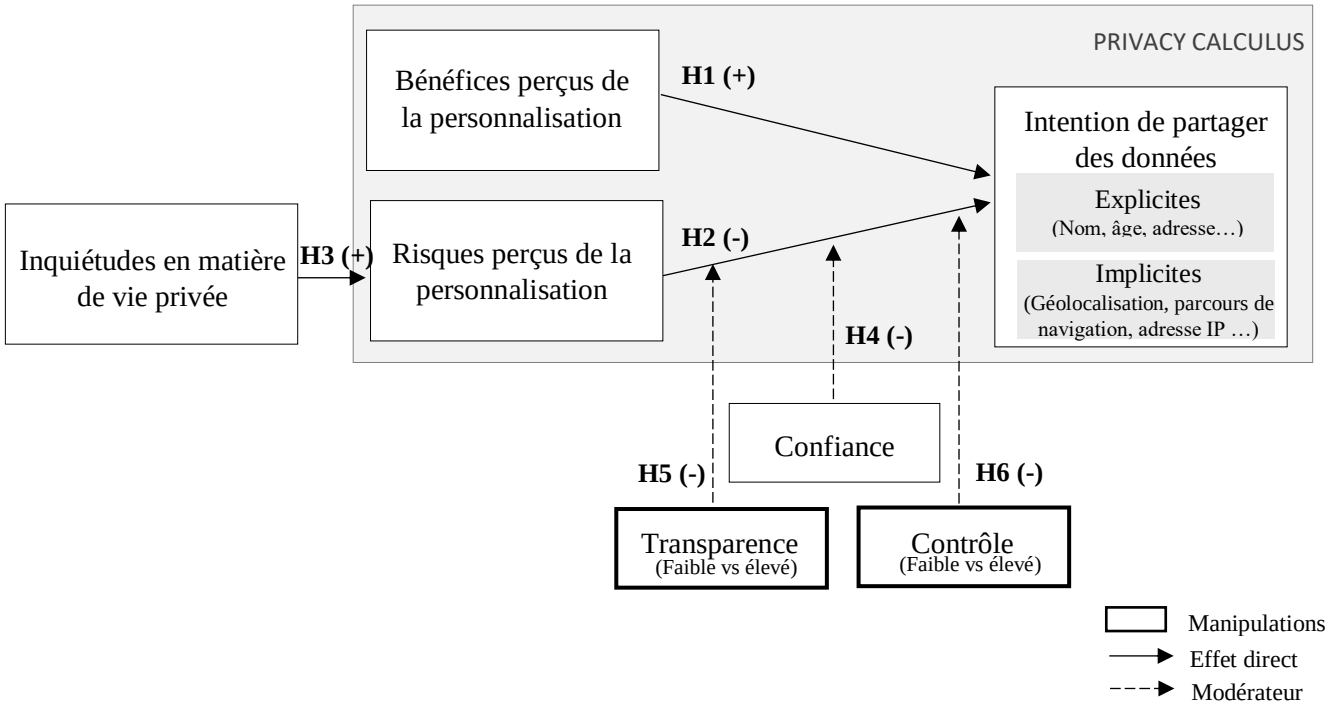
  Hypothèse 6 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données (modération).

  H 6.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données.

  H 6.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données.

Afin d'aborder la question des risques perçus en matière de vie privée dans un contexte de personnalisation en ligne, **notre étude étend le modèle du *privacy calculus*** pour également explorer le rôle des fonctionnalités de contrôle et de transparence ainsi que les inquiétudes en matière de vie privée et la confiance sur la décision de divulgation. Ce modèle de recherche est présenté ci-dessous.
Nous visons principalement à améliorer la compréhension de l'effet des mécanismes de protection et information fournis par les entreprises ainsi que le rôle de la confiance (potentiels mécanismes d'atténuation des risques) sur la relation entre les risques perçus de la personnalisation et la divulgation des données. Ainsi, ces potentiels indicateurs prévisionnels de l'intention de divulgation de l'information pourraient égayer l'entreprise sur la façon d'optimiser la divulgation des consommateurs et les stratégies de personnalisation. Les raisons du choix du modèle sur lequel nous nous appuyons et la différenciation des variables choisies par rapport à d'autres études sont expliquées dans la section suivante. Le raisonnement qui sous-tend les relations hypothétiques est, quant à lui, le reflet de la littérature susmentionnée.

Figure n°2. Modèle de recherche de notre étude



5.2. Choix du modèle de recherche et différenciation

Premièrement, nous avons choisi d'utiliser le ***privacy calculus* comme base** de cette étude car ce modèle est considéré comme étant le plus utile afin de comprendre l'acceptation de la collecte d'informations (Culnan et Bies, 2003, p. 326). L'intention de divulgation des données personnelles en tant que variable dépendante de ce modèle est appropriée dans notre cas, car la personnalisation est uniquement possible si le consommateur partage ses données. Nous avons décidé de faire une distinction au niveau du type de de données personnelles partagées car selon certains chercheurs, cela pourrait impacter la décision. Les recherches explorant le *privacy calculus* dans le but d'expliquer la divulgation d'information sont abondantes dans les contextes tels que les réseaux sociaux, le commerce en ligne, les services de géolocalisation etc. Au contraire, la personnalisation en ligne a reçu moins d'attention. Un soutien empirique au *privacy calculus* a donc été obtenu lors d'études antérieures mais les antécédents et conséquences peuvent varier en fonction du contexte.

Ensuite, concernant les **bénéfices de la personnalisation**, cette étude sera la première à tester empiriquement les items de mesure de Treiblmaier et Pollach (2007). A notre connaissance, aucune des recherches traitant de la personnalisation n'a repris ces items étant donné qu'ils ont soit utilisé une autre variable telle que « la qualité de la personnalisation » (Li & Unger, 2012) ou la « valeur accordée à la personnalisation » (Chellappa & Sin, 2005), soit utilisé un nombre limité d'items comme Xu et al. (2011) qui se sont uniquement concentrés sur la facilité d'utilisation de la technologie et le gain de temps. D'autres ont eu recourt à des attributs propres à une application de la personnalisation comme pour l'application mobile de santé de Guo, Zhang et Sun (2016) et le site de voyage de Lee et Cranage (2011).

Troisièmement, suite à l'appel de plusieurs chercheurs (Xu, et al., 2011 ; Li & Unger, 2012, Martha & Murphy, 2017), cette recherche explore plus en profondeur la nature et le rôle du contrôle, constitué à partir des **concepts de notification (transparence) et de choix (contrôle),** sur la protection de la vie privée dans le cadre de la personnalisation en ligne. En effet, malgré l'attention croissante accordée à la protection de la vie privée et aux perceptions globales des risques, trop peu se sont intéressés aux perceptions sous-jacentes des individus, à savoir celles liées à la transparence et au contrôle dans une situation spécifique de divulgation. Bien que certains chercheurs ont suggéré leurs impacts, peu les ont testés, ou du moins, aucun ne l'a fait conjointement ou dans le contexte du *privacy calculus* à notre connaissance. Or, la littérature laisse entendre que la collecte des données personnelles par l'entreprise est perçue comme étant équitable et moins envahissante, uniquement lorsque le consommateur a été mis au courant et a le pouvoir de gérer ses données (Malhotra, Kim & Agarwal, 2004 ; Culnan & Armstrong, 1999). Nous tenterons donc de remédier à cela et soulignerons la différence entre ces deux variables. Enfin, nous nous différencions de certaines recherches également de par le fait que nous analysons l'effet modérateur de ces variables ainsi que celle de la confiance sur les inquiétudes en matière de confidentialité. D'autres études prises en considération, également basées sur le *privacy calculus*, n'incorporaient pas le concept de confiance ou analysaient l'effet direct de ces pratiques de protection et d'information sur la divulgation de l'information, ne donnant pas de résultat significatif.

La littérature ayant été parcourue en profondeur et le modèle de recherche posé, nous pourrons dans la seconde partie de ce mémoire analyser en pratique le processus de divulgation de données personnelles en vue de bénéficier de la personnalisation.

PARTIE II : ETUDE QUANTITATIVE

Dans la partie précédente de ce mémoire, nous avons établi le cadre théorique de notre recherche. Dès lors, nous allons à présent développer notre propre contribution aux recherches dans ce domaine, en testant notre modèle de recherche à l'aide d'une étude quantitative. Pour ce faire, nous commencerons par décrire la manière dont nous avons procédé afin de pourvoir vérifier ce modèle et justifierons toutes les décisions prises relatives à cela. Dans un second temps, nous effectuerons les opérations préliminaires aux tests, découvrirons les liens entre les variables et la force de celle-ci, pour finalement confirmer ou infirmer notre modèle et hypothèses posées.

1. Méthodologie de la recherche

La méthode choisie afin de vérifier nos hypothèses est l'approche quantitative. Nous estimons en effet qu'après avoir obtenu une bonne connaissance théorique du sujet, une étude quantitative est une méthode appropriée afin de mieux comprendre ce phénomène complexe et répandu qu'est le paradoxe entre la personnalisation et la vie privée. Cela nous permet désormais de réaliser des inférences objectives concernant ce cadre théorique déjà bien reconnu, en mesurant les opinions et intentions d'un grand nombre de consommateurs. Ceci est primordial étant donné qu'il est apparu que l'importance que les individus portent à la protection de leur vie privée diffère fortement d'un individu à l'autre. Dès lors, nous avons eu recours à un questionnaire en ligne dont l'objectif est de mesurer l'impact des inquiétudes générales concernant la vie privée ainsi que les bénéfices et risques perçus de la personnalisation sur l'intention de divulguer de l'information personnelle. Sur base de la littérature, nous pensons que cet impact entre les risques perçus et la divulgation peut être modérée par :

- La transparence fournie par l'entreprise par rapport au traitement des données
- Le contrôle du consommateur sur ses données
- La confiance dans l'entreprise en ligne

1.1. Expérimentation

Notre recherche quantitative s'est faite sous forme expérimentale (2x2), pour laquelle nous avons manipulé les variables de transparence et contrôle, afin d'observer et mesurer l'évolution d'autres variables (Effet des risques perçus sur l'intention de divulgation). Pour cela, quatre scénarios ont été créés faisant varier les niveaux de transparence (faible vs élevé) et de contrôle (faible vs élevé). L'idée est d'analyser si les répondants perçoivent ces facteurs et si ces fonctionnalités de transparence et de contrôle impactent leur intention comportementale. Ces 4 versions se trouvent à l'annexe n°2.

Pour ce faire, le choix de l'entreprise en ligne qui servirait de base à notre étude et aux scénarios était très important. En effet, il était nécessaire de sélectionner une entreprise réelle car nous devions mesurer la confiance envers cette entreprise. La difficulté était alors de trouver une entreprise qui nécessite de récolter de l'information afin de personnaliser son contenu, qui soit connue de tous afin que les répondants puissent au mieux s'imaginer les scénarios mais devant en même temps susciter des niveaux de confiance différents. Pour cela, nous avons choisi l'Avenir, estimant que cette entreprise réponde aux critères que nous venons de citer. Des sites de presse écrite en ligne sont effectivement un bon exemple de sites offrant de la personnalisation, notamment au niveau de la

sélection de leurs articles et du contenu publicitaire grâce au profilage et données clients. Les items de mesure pouvaient ainsi tous être adaptés au contexte de cette étude. Par ailleurs, nous avons voulu éviter de trop grandes références telles que La Libre Belgique ou Le Soir, pour lesquels il nous semblait que la confiance était généralement acquise, du fait de leur renommée.

Une fois l'entreprise choisie, nous avons été amenés à créer un design d'une page web de l'Avenir contenant, le cas échéant, des éléments de transparence et de contrôle ou rien tels que référencés dans la littérature. Ainsi, nous avons tenté de faire correspondre au mieux nos items de mesures avec ces éléments se trouvant dans le design de ces différents scénarios. Nous avons également parcouru plusieurs sites web afin de nous en inspirer pour rendre les scénarios les plus réalistes possibles. Il fut très intéressant de comparer les différences existantes en termes de contrôle et transparence offertes par certains sites. Ensuite, nous avons veillé à ce que ces designs soient les plus similaires possibles afin que les répondants soient dans des conditions identiques. Pour ce faire, un format de base (transparence et contrôle faible) a été réalisé auquel nous avons rajouté quelques éléments de transparence et de contrôle pour satisfaire les trois autres conditions.

Plus précisément, ces scénarios reprennent la page web de l'Avenir avec un message pop-up proposant aux individus de se connecter ou créer un compte afin de bénéficier de nombreux avantages. Dans les conditions de transparence faible, nous avons fourni aux utilisateurs de l'information à propos du type de données récoltées, la durée de conservation des celles-ci et leur usage future. Cette zone de texte est remplacée par de l'information neutre, à savoir une description de ce qu'est l'Avenir, dans le cas de la transparence faible. Ensuite, dans les conditions de contrôle élevé, nous avons intégré des paramètres de gestion de la confidentialité afin d'offrir la possibilité au consommateur de choisir qui aura accès à quel type de données personnelles et à quelles fins elles seront utilisées. Nous avons également mentionné qu'ils auraient la possibilité de changer cela plus tard. Les scénarios contraires contenaient à cet endroit de l'information concernant les différents types de journaux existants de l'Avenir.

Enfin, nous n'avons montré qu'un des scénarios à chaque personne, de manière aléatoire, afin de ne pas biaiser nos résultats.

1.2. Conception du questionnaire

Dans le but de respecter la standardisation du questionnaire (Lambin & de Moerloose, 2008), nous avons construit des questionnaires identiques pour chaque version de scénarios.
Ce questionnaire se compose de 7 sections afin de fournir un maximum de clarté aux répondants. La première partie correspond à la question filtre afin de s'assurer que les répondants connaissent l'Avenir. Ceci est nécessaire car nous mesurons dans la deuxième section le niveau de confiance de chaque individu envers cette entreprise. La troisième partie est la présentation d'un des quatre scénarios choisi de manière entièrement aléatoire par le logiciel et avec un temps d'observation non-limité. Après le visionnage du scénario, les répondants étaient amenés à passer à la page suivante du questionnaire. Celle-ci est composée des variables indépendantes, notamment les questions relatives à l'intention de divulgation, compte tenu du scénario visualisé. En cinquième lieu, les questions relatives au *privacy calculus*, notamment les bénéfices et risques perçus, sont exposées. La sixième partie est

constituée des manipulations checks, c'est-à-dire les questions liées à la transparence, le contrôle et le réalisme du scénario visualisé. Dans la dernière partie, les caractéristiques sociodémographiques de l'échantillon sont récoltées (sexe, âge, diplôme obtenu et statut professionnel) ainsi que deux autres questions concernant leur profil, la fréquence d'utilisation d'Internet et leur tendance générale à être inquiets en matière de protection de la vie privée. Ce questionnaire était précédé d'un texte d'introduction pour en expliquer son utilité, motiver l'individu à participer et l'incitant à répondre le plus sincèrement possible, sans se baser sur ce qu'ils croient être socialement acceptable. Un message de clôture les remerciait de leur participation. Cela durait environ 8 minutes pour y répondre.

Nous avons fait attention à ce que ces questions suivent un ordre logique. D'une part, l'ordre respecte le principe de l'entonnoir car le corps du questionnaire, composé de questions plus générales, est présenté au début et celles plus personnelles et précises, dont le profil sociodémographique, à la fin du questionnaire. Cela permet d'éviter de décourager d'entrée les individus en paraissant trop indiscret. D'autre part, afin de ne pas influencer la réponse du consommateur et d'éviter « l'effet de contamination », certaines questions ont dû être posées avant d'autres. Le programme nous a également permis d'empêcher les répondants de retourner en arrière vers les questions déjà répondues. Ceci est important afin de refléter au mieux la réalité des scénarios, sachant qu'un consommateur ne prendra que quelques secondes afin de déterminer si oui ou non il divulguera de l'information, sans nécessairement regarder en profondeur si des éléments de contrôle et de transparence sont proposés.
Ensuite, nous avons également fait en sorte que les questions soient facilement comprises par tous et non-orientées ainsi que tenté de garder le questionnaire le plus court possible afin d'éviter que ceux-ci n'abandonnent avant sa finalisation. Ce questionnaire contient uniquement des questions fermées et donc des réponses limitées aux solutions, excepté pour la question relative à la profession du répondant contenant une option « autre ». Ce type de questions permet d'obtenir des données cohérentes et une faible variabilité des résultats (Malhotra & Birks, 2017).

Pour mesurer nos différentes variables et s'assurer de la qualité des réponses, nous nous sommes inspirées d'échelles de mesure évoquées et validées dans la littérature existante. Nous les avons traduites en français et adaptées aux scénarios. Ces échelles sont toutes mesurées sur des échelles de Likert à 7 points allant de « pas du tout d'accord » à « tout à fait d'accord », excepté celles pour l'intention de divulgation et le profil sociodémographique du répondant. Celles-ci sont respectivement mesurées sur une échelle sémantique à 7 points et des échelles nominales et ordinales. Il semble que l'échelle de Likert est régulièrement utilisée pour ce type de questionnaire et adéquate car elle est simple à concevoir et à manipuler pour les chercheurs. Par ailleurs, bien qu'elle puisse nécessiter plus de temps de remplissage, elle est simple à comprendre pour les répondants (Malhotra & Birks, 2017). Effectivement, du fait qu'elle possède un nombre impair de catégories, cette échelle ne force pas les réponses des répondants. Elle permet au contraire d'exprimer leur indifférence à l'égard de certaines questions grâce à l'option neutre. Par ailleurs, nous avons veillé à intégrer une description aux catégories des échelles de Likert, afin de réduire davantage l'ambiguïté causée par leurs variétés. Le tableau ci-dessous présente la source et le type des échelles de mesure sélectionnées ainsi que le nombre d'items respectifs. Le tableau complet avec les items se trouve à l'annexe n°3.

Figure n°3 : Récapitulatif des échelles de mesure utilisées

Variable	#Items de mesure	Type d'échelle	Source(s) : Auteur + journal + classification UNIFI des revues de haut niveau
Confiance envers l'entreprise en ligne	3	Echelle de Likert à 7 points	Palmatier, Scheer, Houston, Evans & Gopalakrishna (2007) *International Journal of Research in Marketing (class. 2)*
Intention de divulguer des données	4	Echelle sémantique différentielle mesurée sur 7 points	Malhotra, Kim & Agarwal (2004) *Information Systems Research (class. 3)*
Bénéfices perçus de la personnalisation suite à la divulgation des données	4	Echelle de Likert à 7 points	Treiblmaier & Pollach (2007) *International Conference on Information System Proceedings (class. 2)*
Risques perçus de la personnalisation suite à la divulgation des données	3	Echelle de Likert à 7 points	Dinev, Bellotto, Hart, Russo, Serra & Colautti (2006) *European Journal of Information Systems (class. 3)*
Contrôle	4	Echelle de Likert à 7 points	4 items de Xu, Dinev, Smith & Hart (2008) *International Conference on Information Systems (class. 2)*
Transparence	6	Echelle de Likert à 7 points	4 items de Awad & Krishnan (2006) *MIS Quarterly (class. 3)* + 2 items de Portes (2018)
Réalisme des scénarios	3	Echelle de Likert à 7 points	Source personnelle
Inquiétudes générales en matière de vie privée	4	Echelle de Likert à 7 points	Dinev, Bellotto, Hart, Russo, Serra & Colautti (2006) *European Journal of Information Systems (class. 3)*
Sexe	2	Echelle nominale	Source personnelle
Age	7	Echelle de proportion	Source personnelle
Niveau d'éducation	4	Echelle ordinale	Source personnelle
Profession	11	Echelle nominale	Source personnelle
Fréquence d'utilisation d'internet	8	Echelle Ordinale	Source personnelle

1.3. Choix de l'échantillon

La cible n'est pas fortement restreinte étant donné que la personnalisation en ligne peut concerner toute personne, belge dans notre cas, naviguant sur internet. Il a cependant été nécessaire de trier la population selon sa connaissance ou non l'Avenir, afin de s'assurer qu'elle réponde à la question de la confiance en connaissance de cause. La population belge semble être adéquate car selon le modèle culturel de Hofstede (Hofstede Insights, 2018), la population belge accorderait relativement beaucoup d'importance à la protection individuelle, ne tolèrerait pas les situations ambigües et prendrait tous les éléments en considération lors de la prise de décision afin d'atteindre le meilleur compromis. Les données de Febelfin (2017) et SPF Economie (2018) confirment qu'un peu plus de la moitié des belges est préoccupée par la vie privée et la protection des informations personnelles, n'ayant pas nécessairement recours à des mesures de protection mais attachant une grande importance à la transparence.

1.4. Le pré-test

Le pré-test est une étape indispensable pour la réalisation d'un bon questionnaire, avant de l'utiliser sur le terrain. Ceci a été effectué auprès de 5 personnes afin d'identifier d'éventuels problèmes liés à la conception du questionnaire, plus particulièrement au niveau de la compréhension des questions par les répondants, et d'évaluer la durée du questionnaire. Ensuite, nous avons également voulu s'assurer que les répondants perçoivent davantage de contrôle et de transparence dans les scénarios avec de la transparence et du contrôle élevé, que dans ceux ne contenant pas ces éléments. Ceci a été réalisé auprès d'une vingtaine de personnes, afin d'avoir minimum 5 réponses valides par scénario. A l'issue de ces pré-tests, quelques difficultés sont ressorties, qui nous ont amené à améliorer le questionnaire. Celles-ci concernaient principalement des ambiguïtés par rapport aux questions liées aux scénarios, car les répondants ne lisaient pas assez attentivement ou ne mémorisaient pas le scénario montré. Ils avaient également tendance à ne pas comprendre qu'il y avait plusieurs scénarios et que donc certaines questions étaient moins pertinentes en fonction de l'unique scénario qui leur avait été montré. Dès lors, nous avons davantage insisté sur le besoin d'être attentif à la mise en situation et intégré plus de textes explicatifs à chaque question relative à cela. Nous avons également corrigé des fautes d'orthographes, supprimé des répétitions d'information et agrandi les images des scénarios.

Quant aux analyses statistiques de ce pré-test, les scénarios ont été considérés comme réalistes et les individus semblent avoir perçu les différents niveaux de transparence et de contrôle. En effet, suite à une comparaison de moyennes ANOVA, il est apparu que la moyenne de la variable « transparence » est bien significativement supérieure dans les scénarios où la transparence est élevée (p-valeur = 0,002 ; moyenne transparence élevée = 4,7 > faible = 2,8), ce qui est également le cas pour le contrôle (p-valeur = 0,001 ; Moyenne contrôle élevé = 4,5 > faible = 1,9). Ces résultats sont repris à l'annexe n° 4 et la version finale du questionnaire se trouve à l'annexe n°5.

1.5. Matrice « hypothèses-questions »

Le tableau récapitulatif suivant présente clairement que les questions du questionnaire sont toutes nécessaires car elles correspondent à toutes les hypothèses que nous avons construites ou sont nécessaires pour nos analyses. La première question ne correspond à aucune hypothèse car c'est une question filtre.

Figure n°4 : Matrice hypothèses - questions

	H1	H2	H3	H4	H5	H6	Manipulation checks	Analyse de l'échantillon
Q1	Question filtre							
Q2				x				
Q3	x	x						
Q4	x	x						
Q5	x							
Q6		x						
Q7					x		x	
Q8						x	x	
Q9							x	
Q10			x					x
Q11								x
Q12								x
Q13								x
Q14								x
Q15								x

1.6. Distribution du questionnaire

Le questionnaire a été créé en ligne via le logiciel Qualtrics et a été distribué par e-mail et sur les réseaux sociaux pendant 4 jours au mois de juillet 2018, jusqu'à ce que le nombre de répondants nécessaires soit atteint. Afin de se rapprocher le plus possible d'un échantillon représentatif et potentiellement permettre la généralisation des résultats, nous avons tenté de diffuser au mieux le questionnaire, c'est-à-dire à des profils diversifiées en termes d'âge, sexe, profession et éducation. En ce qui concerne les scénarios, ils ont été distribués de manière aléatoire par le logiciel, formant quatre sous-groupes de répondants. Il avait été décidé qu'au moins 50 réponses par scénarios seraient nécessaires.

L'enquête en ligne a été retenue comme méthode de collecte de données pour plusieurs raisons. Premièrement, la diffusion s'est déroulée uniquement en ligne car ce n'était pas utile de toucher des personnes qui ne disposent pas d'une connexion internet. Ensuite, le plus gros avantage de cette méthode de sondage est qu'elle permet d'atteindre un grand nombre de personnes pouvant générer des résultats représentatifs, et ce, efficacement, à faible coût et assez rapidement (Malhotra & Birks, 2017). Par ailleurs, le questionnaire limite les biais inhérents à la présence d'un enquêteur, offre du contrôle sur l'ordre des questions ainsi qu'un certain niveau d'anonymat, motivant les répondants à répondre avec plus d'honnêteté. Cette technologie permet également de superviser automatiquement des erreurs, d'encoder immédiatement les réponses et d'intégrer une aide visuelle comme les scénarios (Lambin et de Moerloose, 2008). En outre, elle facilite la tâche des répondants en proposant des questions à choix multiple. Enfin, grâce à l'utilisation d'internet pour ce questionnaire, nous avons pu bénéficier d'un effet boule de neige, selon lequel nos contacts ont à leur tour partagé le questionnaire avec leurs connaissances.

Toutefois, le questionnaire en ligne présente des limites qui doivent être prises en compte lors de l'analyse des données. Les inconvénients de cette méthode comprennent le manque de contrôle de l'identité des répondants, l'échantillonnage non aléatoire, le contact impersonnel et la peur quant à la réelle confidentialité de l'information. Par ailleurs, la récolte des données via un sondage en ligne pose généralement plus de problème au niveau de la représentativité de l'échantillon que toutes les autres méthodes, surtout si la diffusion de celui-ci se fait auprès du réseau de connaissances (Lambin & de Moerloose, 2008).

2. Analyse des résultats

Une fois les données collectées, l'étape suivante est celle du processus d'analyse, en commençant par la préparation des données, la description de l'échantillon et les opérations préliminaires aux tests (Lambin & de Moerloose, 2008). Nous décrirons ces différentes étapes de préparation et vérification, suivis des tests d'hypothèses effectués à l'aide du logiciel SPSS dans ce chapitre. Notons que nous travaillerons avec un seuil de signification de 5% et que les détails de sorties se trouvent systématiquement dans les annexes.

2.1. Préparation des données

Pour que nous puissions traiter les données au moyen du logiciel SPPS, nous avons tout d'abord éliminé les questionnaires incomplets. Nous avons pu obtenir 439 réponses au total, mais étant donné que la première question était une question filtre et que certaines personnes ont été découragées par le scénario, nous avons finalement pu retenir 233 réponses complètes et exploitables. Ensuite, nous avons modifié la présentation des données récoltées en les transformant en variables numériques afin de créer une banque de données structurées. Les lignes de notre base de données correspondent aux répondants et les colonnes aux variables reprises dans chacune des questions. En outre, les catégories contenant un nombre d'effectifs insuffisant, c'est-à-dire inférieur à 5, ont été regroupées de manière logique. Dès lors, les 2 réponses « -18 ans » et « Enseignement secondaire » sont reprises dans la catégorie « 18-24 ans » et « Enseignement secondaire ». Enfin, nous avons regroupé les catégories concernant la profession selon leur contexte de travail et s'ils gagnent de l'argent ou non. Ainsi nous avons une catégorie « étudiant », une « femme/homme au foyer/ stagiaire / retraité / chercheur(se) d'emploi », une autre pour les « cadre / employé(e) / enseignant(e) » et enfin une dernière catégorie comprenant « indépendant(e) / profession libérale ». Les trois réponses libres à cette question, à savoir « stagiaire », « fonctionnaire, responsable de classes de vertes » et « administrateur de sociétés » ont également été logiquement assignées à ces catégories. Chacune de ces réponses a été traduite par un code. Le plan de codage se trouve à l'annexe n°6.

2.2. Description de l'échantillon

Notre échantillon se compose de 233 répondants dont 59% de femmes et 41% d'hommes, relativement soucieux en ce qui concerne la protection de la vie privée. Les répondants font partie de toutes les catégories d'âge sélectionnées (de -18 ans à +65 ans), la majorité appartenant aux catégories d'âge « 18-24 ans » (44%) et « 45-54 ans) (21%). Plus de la moitié de notre échantillon (66%) détient un diplôme d'enseignement supérieur universitaire et le type de profession est assez varié. Ainsi, nous avons des individus de chaque catégorie, les prédominantes étant celles contenant les employés, cadres et enseignants (40 %) ainsi que les étudiants (34%). Par ailleurs, 99% de nos répondants utilisent internet quotidiennement. Les données de notre échantillon en ce qui concerne l'utilisation d'internet sont très intéressantes car elles supposent que nos répondants ont une certaine expertise en ce qui concerne les potentiels risques de confidentialité en ligne, sont familiers avec l'internet et la divulgation en ligne et ont fort probablement déjà été confrontés à la personnalisation en ligne. On peut donc supposer qu'ils ont fourni des données pertinentes et exploitables pour notre analyse. Un aperçu complet du profil sociodémographique de notre échantillon se trouve à l'annexe n°7.

Il est nécessaire de prendre un certain recul par rapport à notre échantillon et évaluer s'il est représentatif de la population cible, à savoir toute personne belge utilisant internet. Selon les données de Statbel (2018), parmi les internautes belges, 88% des femmes et 86 % des hommes utilisent, tous les jours ou presque tous les jours, internet et 10% et 11% une fois par semaine au moins. Ceci est assez proche des résultats de notre échantillon car les répondants utilisent presque à l'unanimité quotidiennement internet. Un haut niveau d'utilisation de l'internet semblerait aller de pair avec un haut niveau d'inquiétudes au sujet de la vie privée selon les chercheurs, ces préoccupations étant une caractéristique de la population belge comme nous l'avons mentionné plus haut dans ce travail.

Cependant, si nous jetons un coup d'œil aux statistiques descriptives de cette variable, nous constatons que les craintes en matière de vie privée de notre échantillon sont très élevées car plus de 65% de nos répondants considèrent être inquiets, voire très inquiets. Ceci est nettement supérieur à la moyenne belge. Ensuite, nous avons une répartition en fonction du sexe assez équitable, bien que cette différence entre les hommes et les femmes soit encore plus faible au niveau de la population belge en général (Statbel, 2017). Concernant l'âge de nos répondants, nous constatons que les différences, en termes d'utilisation d'internet au quotidien en fonction des tranches d'âge en Belgique, ne varient pas fortement. Ensuite, le niveau d'instruction est plus élevé dans notre échantillon que pour la population belge, 36% ayant un diplôme du secondaire supérieur, 17% et 12% un diplôme supérieur non universitaire et universitaire respectivement (Statbel, 2017, p.118). Quant à la profession, les employés seraient en effet la catégorie la plus courante en Belgique.

Les chiffres de notre échantillon diffèrent donc sur certains points par rapport à la population belge, à savoir des niveaux d'éducation, d'utilisation d'internet et d'inquiétudes supérieurs. Les différences sont dues au fait que pour la dispersion de notre questionnaire, nous avons eu recours à un échantillon de commodité, qui est une technique d'échantillonnage non probabiliste, car nos répondants font principalement partie de nos réseaux de proches (Malhotra & Birks, 2017). Dans tous les cas, il est important de connaître l'opinion en matière de collecte et d'utilisation des données de cette échantillon afin d'évaluer l'impact de ce phénomène sur les internautes du futur. En effet, sachant que le nombre d'internautes augmente constamment, de plus en plus d'individus se trouveront dans la situation de ceux faisant partie de notre échantillon dans les années à venir. Nous devons cependant garder à l'esprit que nos répondants semblent être plus inquiets que la normale.

2.3 Comparaison des échantillons des sous-groupes

Comme déjà mentionné, le logiciel Qualtrics a formé aléatoirement 4 sous-groupes de répondants afin qu'ils ne soient exposés qu'à un seul design. Il est à présent nécessaire de vérifier qu'il n'y ait pas de différences significatives entre les profils sociodémographiques de ces sous-groupes (transparence élevée vs faible et contrôle élevé vs faible). Autrement dit, nous devons vérifier si les variables sont dépendantes entre elles ou dues au hasard. Pour ce faire, nous avons réalisé des tests d'indépendance Khi-carré entre la variable « Conditions » et les variables « Genre », « Age », « Education » et « Profession ». Les hypothèses posées sont les suivantes :

H0 : la variation entre les différents échantillons est due au hasard.
(Indépendance des variables)
H1 : la variation entre les différents échantillons n'est pas due au hasard
(Dépendance des variables)

La condition d'application exigeant que le nombre de répondants par catégorie soit égal ou supérieur à 5 est remplie. Pour cette vérification prenons en tant qu'illustration, les résultats du test de la variable « Genre ».

Figure n°5 : Tableau croisé pour condition - sexe

		Sexe		
		Femme	Homme	Total
Condition	Condition 1	36	18	54
	Condition 2	29	33	62
	Condition 3	33	23	56
	Condition 4	39	22	61
Total		137	96	233

Ce tableau croisé présente la répartition des hommes et des femmes pour chaque condition ou scénario. Ensuite, les résultats du test du khi-carré dans le tableau ci-dessous indiquent que la valeur du khi-carré obtenu est de 5,745 avec 3 degrés de liberté associés et une p-valeur de 0,125. Cette dernière étant supérieure au seuil alpha de 0,05, nous ne pouvons pas rejeter l'hypothèse H0 d'indépendance. Cela signifie que les 4 sous-groupes ont bien été formés au hasard en termes de genre.

Figure n°6 : Test du khi-carré pour conditions - sexe

	Valeur	ddl	Signification asymptotique (bilatérale)
khi-carré de Pearson	5,745[a]	3	,125
Rapport de vraisemblance	5,719	3	,126
N d'observations valides	233		

Les 4 groupes sont également équivalents pour les autres variables. Nous pouvons donc en conclure que les analyses futures ne seront pas biaisées par la représentativité de ces sous-groupes. Le détail de ces résultats est disponible à l'annexe n°8.

2.4 Opérations préliminaires

2.4.1. Validité et fiabilité des échelles de mesure

Avant d'étudier les relations entre les variables, il est nécessaire de mener des analyses préliminaires sur les échelles de mesures de nos variables afin de nous assurer de la validité et fiabilité de nos échelles de mesure. La validité correspond au fait de mesurer la bonne dimensionnalité de nos variables, tandis que la fiabilité détermine s'il existe une cohérence au niveau de ce que l'on mesure via une variable. Les échelles à vérifier sont celles mesurant la confiance (« Conf »), l'intention de divulgation des données explicites (« Divulg1 ») et implicites (« Divulg 2 ») les bénéfices perçus (« Benef »), les risques perçus (« Risques »), le contrôle (« Ctl»), la transparence (« Transpa »), le réalisme (« Real ») et les inquiétudes en matière de vie privée (« Concerns »).

Premièrement, nous avons effectué des **analyses factorielles** en composantes principales qui nous ont permis de vérifier que les items sont corrélés et peuvent être regroupés entre eux, c'est-à-dire résumés en un ou plusieurs facteurs. Afin d'illustrer cette analyse, nous présentons ci-dessous en détail l'analyse factorielle de la variable « Risques ».

Pour vérifier si nos données sont factorisables, nous commençons par observer les résultats du test de sphéricité de Bartlett. Ceci nous permet de tester la corrélation, ce qui est une condition nécessaire pour réaliser l'analyse factorielle. Nous posons les hypothèses suivantes :

H0 : il n'y a pas de corrélation inter-items
H1 : il y a une corrélation inter-items

Par ailleurs, il est nécessaire d'observer l'indice de Kaiser-Meyer-Olkin (KMO) dont la valeur varie entre 0 et 1. Un indice élevé indique que les corrélations entre les items sont de bonne qualité, rendant l'analyse factorielle appropriée. Nous vérifierons donc que nos valeurs se rapprochent de 1, sachant que Kaiser recommande une valeur KMO de plus de 0,5 et qu'une valeur de 0,7 commence à être considérée comme bon (Field, 2013, p.1097).

Figure n°7 : Indice KMO et test de Bartlett pour la variable « risques»

Indice de Kaiser-Meyer-Olkin pour la mesure de la qualité d'échantillonnage.		,740
Test de sphéricité de Bartlett	Khi-carré approx.	403,647
	ddl	3
	Signification	,000

Les résultats obtenus pour la variable « Risques» présentés dans le tableau ci-dessus indiquent que la p-valeur du test de Bartlett est inférieure à 0,05, l'hypothèse nulle est rejetée et il existe donc une relation entre les items permettant l'analyse factorielle. Quant à l'indice KMO, il est de 0,740 et donc considérée comme bon.

Les conditions d'application étant remplies, nous pouvons, à présent, commencer l'analyse factorielle en tant que telle. Ainsi, si nous observons le tableau « variance totale expliquée » se trouvant ci-dessous, nous remarquons qu'une seule composante a une valeur propre supérieure à 1. Celle-ci permet d'exprimer environ 82 % de la variance totale. Le tracé d'effondrement se trouvant à l'annexe n°9 permet également de vérifier le nombre de composantes. Selon Cattel, quand la pente devient horizontale, le rapport coût-bénéfice n'est pas bon et il ne faut donc pas ajouter de composante.

Figure n°8 : Variance totale expliquée pour la variable « risques»

	Valeurs propres initiales			Sommes extraites du carré des chargements		
Composante	Total	% de la variance	% cumulé	Total	% de la variance	% cumulé
1	2,460	82,009	82,009	2,460	82,009	82,00
2	,319	10,647	92,656			
3	,220	7,344	100,000			

Méthode d'extraction : Analyse en composantes principales.

Ensuite, nous observons les valeurs des communautés (*communalities*) du tableau « qualités de représentation ». La partie extraction de cette table indique pour chaque item la part de variance expliquée par la solution, ici une composante, et doit idéalement être supérieur à 0,5 (Field, 2013, p.1097).

Dans le cas de la variable « risques», nous constatons sur le tableau ci-dessous que tous les items sont en effet supérieurs à ce seuil.
Finalement, le tableau « matrice des composantes » présente les *loadings*, c'est-à-dire les corrélations linaires entre les items originales en ligne et les composantes en colonnes. Ceux-ci doivent idéalement être supérieurs à 0,4, bien que certains chercheurs opte pour le seuil minimum de 0,3 (Field, 2013, p.1093). Nous observons que les *loadings* de chaque item de la variable « risques» sont élevés et supérieurs à 0,4. Sur base de tout cela, il apparait donc que les items de cette variable sont regroupés en un facteur.

Figure n°9 : Qualités de représentation pour la variable « risques »

	Initiales	Extraction
Risques1	1,000	,785
Risques2	1,000	,827
Risques3	1,000	,849

Figure n°10 : Matrice des composantes pour la variable « risques »

	Composante 1
Risques1	,886
Risques2	,909
Risques3	,921

Quant à l'analyse en composantes principales, nous pouvons conclure que toutes nos variables sont unidimensionnelles, ce qui corrèle avec les dimensions déterminées à priori grâce à la recherche théorique. Les résultats de ces autres variables sont détaillés à l'annexe n°9 et repris sommairement dans le tableau récapitulatif ci-dessous.

Deuxièmement, afin de vérifier si nos échelles de mesures sont fiables, nous avons mené des **analyses de l'Alpha de Cronbach** pour chaque facteur afin de vérifier si les items mis ensemble, sous une échelle, mesurent la même chose. Les résultats sont également représentés dans le tableau ci-dessous et à l'annexe n°9. Nous considérons qu'un facteur est fiable lorsque l'Alpha de Cronbach est supérieur à 0,7.

Le tableau ci-dessous présente l'Alpha de Cronbach pour la variable « Risques», qui est bien supérieur à ce seuil, ce qui confirme que l'échelle de mesure pour cette variable est fiable.

Figure n°11 : Test de fiabilité - Alpha de Cronbach

Alpha de Cronbach	Nombre d'éléments
,889	3

Quant aux autres variables, tous les résultats sont également supérieurs à 0,7, ce qui témoigne d'une bonne fiabilité de tous les items.

Sur base de ces analyses préliminaires, nous avons créé de nouvelles variables appelées « *summated scale* » pour chaque facteur, représentant la moyenne de leurs items. Ces nouvelles variables seront utilisées dans les analyses futures. Le tableau suivant fournit les résultats de ces opérations préliminaires pour toutes les variables

Figure n° 12 : Tableau récapitulatif des résultats des analyses factorielles et de fiabilité

Variable	#facteurs	Nom du facteur	#items	Alpha de Cronbach
Confiance	1	Conf	3	0,928
Intention de divulguer des données personnelles explicites	1	Divulg1	4	0,964
Intention de divulguer des données personnelles implicites	1	Divulg2	4	0,980
Bénéfices perçus de la personnalisation	1	Benef	4	0,811
Risques perçus de la personnalisation	1	Risques	3	0,889
Contrôle	1	Ctl	4	0,942
Transparence	1	Transpa	6	0,857
Réalisme des scénarios	1	Real	3	0,869
Inquiétudes en matière de vie privée	1	Concerns	4	0,901

2.4.2. Vérifications des manipulations expérimentales

Cette étape est nécessaire afin d'analyser l'influence des manipulations expérimentales sur les participants et ainsi s'assurer de leur exactitude. Il s'agit donc de vérifier que les répondants ont perçu un niveau élevé de transparence dans les deux premiers scénarios, et faible dans les deux derniers ainsi qu'un niveau élevé de contrôle dans les scénarios 1 et 3 et faible dans les autres. Pour ce faire, des tests ANOVA ont été effectués, permettant de vérifier que les moyennes des *manipulation checks* varient significativement d'un scénario à l'autre. Les conditions d'applications, notamment l'indépendance des données, la distribution normale des données et l'homogénéité des variances ont bien été vérifiées (voir annexe n°11).

Prenons comme exemple la manipulation de la transparence. Nous avons comme hypothèses :

H0: La moyenne de la transparence ne diffère pas en fonction des conditions
H1 : La moyenne de la transparence diffère en fonction des conditions

Sur base des résultats de l'analyse ANOVA affichés dans la figure n°12 ci-dessous et de la figure n°13 reprenant la comparaison des moyennes pour la transparence, il apparait que nos manipulations ont fonctionné. En effet, la p-valeur (= 0,002) est inférieure à 0,05 et nous observons que dans les scénarios où les participants étaient exposés à beaucoup de transparence, la moyenne du groupe est de 4,26 et est supérieure à la moyenne de la transparence faible, notamment 3,7.

Figure n° 13 : Vérification de la manipulation de la transparence - Test ANOVA

ManipulationCheckTranspa

	Somme des carrés	ddl	Carré moyen	F	Sig.
Intergroupes	17,935	1	17,935	10,009	,002
Intragroupes	413,943	231	1,792		
Total	431,878	232			

Figure n° 14 : La comparaison des moyennes pour la variable transparence

ManipulationCheckTranspa

	N	Moyenne	Ecart type	Erreur standard	Intervalle de confiance à 95 % pour la moyenne		Minimum	Maximum
					Borne inférieure	Borne supérieure		
Faible	117	3,7009	1,41193	,13053	3,4423	3,9594	1,00	7,00
Elevé	116	4,2557	1,26040	,11703	4,0239	4,4876	1,00	7,00
Total	233	3,9771	1,36438	,08938	3,8010	4,1532	1,00	7,00

Similairement, la moyenne de 3,4 du groupe « contrôle élevé » est supérieure à la moyenne 2,7 du groupe « contrôle faible » et la p-valeur est de 0,003. Les différences de transparence et contrôle parmi les différents scénarios ont donc bien été perçues par les répondants. Concernant le réalisme, le test ANOVA n'est pas significatif et les moyennes sont assez proches, ce qui est normal pour cette variable qui devrait être assez identique dans les 4 conditions. Les détails de ces tests se trouvent à l'annexe n° 10.

2.5 Analyse des hypothèses

Nous allons à présent procéder aux tests des 6 hypothèses de notre modèle de recherche. Celui-ci contient des régressions linéaires et des modérations. Les tests des hypothèses sous-jacentes, nécessaires au bon déroulement des tests de régression, ont été effectués. Il s'agit de la vérification de l'indépendance des résidus, de la normalité des résidus, du respect de la condition d'homoscédasticité et dans certains cas, la colinéarité. Ces conditions d'applications sont relativement bien respectées et présentées en annexe ainsi que toutes les sorties des analyses des tests d'hypothèses (annexe n°12).
Notons que le B mentionné dans cette partie fait référence au coefficient de régression non-standardisé (beta non-standardisé) et le R^2 est le coefficient de détermination ajusté, c'est-à-dire qu'il s'ajuste en fonction du nombre de variables indépendantes du modèle.

2.5.1. Hypothèses du *Privacy Calculus*

Hypothèse 1 : Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation).

H1.1 Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **explicites** (nom, âge, adresse...).

H1.2 Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **implicites** (Géolocalisation, parcours de navigation, adresse IP…).

Hypothèse 2 : Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation).

H2.1 Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles **explicites** (nom, âge, adresse...).

H2.2 Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles **implicites** (Géolocalisation, parcours de navigation, adresse IP…).

Pour ces hypothèses du *privacy calculus*, nous avons effectué **deux régressions linéaires multiple**, étant donné que deux variables, notamment les bénéfices perçus et les risques perçus expliquent la variable dépendante « Divulg 1 » (l'intention de divulgation de données personnelles explicites) ou « Divulg 2 » (l'intention de divulgation de données implicites). Les deux hypothèses sont vérifiées. L'ensemble de ces résultats se trouve à l'annexe n°11 (A) et (B).

Ainsi, dans le cas de la régression linéaire multiple pour les données implicites (**H1.1 et H2.1**), la p-valeur du modèle ainsi que celles des variables sont significatives (p-valeur = 0,00 → NRH0), l'effet des bénéfices est positif (B = 0,516) et celui des risques négatif (B=-,265) sur l'intention de divulguer ces données. Ainsi l'équation de la régression linéaire multiple est :

Divulg1 = 2,471 + 0,516 (bénéfices) – 0,265 (risques) + erreur

Sur base des valeurs absolues du coefficient de régression standardisé, nous pouvons remarquer que l'importance des bénéfices perçus sur la variable dépendante est presque deux fois plus importante que celle des risques.

En outre, le R^2, qui correspond à la proportion de variation d'une variable expliquée par l'autre variable et mesure donc la force de l'association (Malhotra, & Birks, 2017), est de 0,242. Ainsi, cela signifie que les variables « Bénéfices» et « Risques» expliquent 24,2% de la variation de la variable « Divulg 1 ».

Les hypothèses sont donc **vérifiées**. Nous pouvons affirmer que l'intention de divulguer des données personnelles explicites est négativement influencée par les risques perçus et positivement par les bénéfices perçus. Ainsi, plus un individu perçoit des bénéfices de la personnalisation, plus il aura l'intention de divulguer des données personnelles explicites. Puis, plus il perçoit des risques de la personnalisation, moins il aura l'intention de divulguer ses données personnelles explicites.

Similairement, pour la régression linéaire multiple en ce qui concerne les données personnelles implicites (**H1.2 et H2.2**), la p-valeur du modèle est significative (p-valeur = 0,00 → NRH0), l'effet des bénéfices est positif (p-valeur = 0,00, B = 0,415) et celui des risques négatif (p-valeur = 0,001, B=-,232) sur l'intention de divulguer ces données. Dès lors, l'équation de régression linéaire multiple est la suivante :

Divulg 2 = 2,036 + 0,415 (bénéfices) – 0,232 (risques) + erreur

Nous pouvons également supposer, à partir des valeurs absolues du coefficient de régression standardisé, que la variable « bénéfices» a une relation plus forte avec la variable indépendante que celle des risques. Par ailleurs, le R^2 est de 0,201, ce qui signifie que « Bénéfices » et « Risques» expliquent 20,1% de la variation de la variable « Divulg 2 ». Les hypothèses sont donc **vérifiées**.

Nous pouvons affirmer que l'intention de divulguer des données personnelles implicites est négativement influencée par les risques perçus et positivement influencée par les bénéfices perçus. Ainsi, plus un individu perçoit des bénéfices de la personnalisation, plus il aura l'intention de

divulguer des données personnelles implicites. Puis, plus il perçoit des risques de la personnalisation, moins il aura l'intention de divulguer ses données personnelles implicites.

En conclusion, nous pouvons affirmer que l'intention de divulguer des données personnelles est significativement influencée par les bénéfices et risques perçus de la personnalisation, peu importe le type de données. Par ailleurs, les bénéfices perçus de la personnalisation semblent impacter plus fortement la prise de décision de divulguer de l'information personnelle, que les risques perçus.

En ce qui concerne le type de données, via un **test T pour échantillons appariés**, nous avons remarqué qu'il existe une différence significative entre les moyennes de «Divulg 1 » (intention de divulguer des données explicites) et «Divulg 2 » (intention de divulguer des données implicites) (p-valeur = 0,00 → NRH0). Par ailleurs, la moyenne pour les données explicites (= 3,21) est supérieure à celle des données explicites (= 2,54). Il semblerait donc que les individus sont plus disposés à divulguer des données explicites de type nom, âge, adresse que des données implicites telles que l'historique de navigation la géolocalisation etc.

2.5.2. Hypothèse : Inquiétudes en matière de vie privée – risques perçus

Hypothèse 3 : Les inquiétudes en matière de vie privée ont une influence positive sur les risques perçus de la personnalisation.

Pour analyser cette hypothèse, nous avons effectué une **régression linéaire simple**. Cela nous a permis d'affirmer que la variable « Risques » est significativement et positivement influencée par la variable « Concerns » (Inquiétudes en matière de vie privée).

En effet, la p-valeur est significative (p-valeur = 0,00 → NRH0) et l'effet est positif (B = 0,324). Le R^2 est de 0,081, ce qui signifie que « Concerns » explique 8,1% de la variation de la variable « risques ». L'équation linéaire simple est la suivante :

$$\text{Risques} = 2{,}934 + 0{,}324\ (\text{Concerns}) + \text{erreur}$$

Cette hypothèse est **validée**. Nous pouvons conclure que plus un individu sera préoccupé par la confidentialité de ses données, plus il percevra des risques en matière de vie privée. Le détail des résultats est disponible à l'annexe n°11 (B).

2.5.3. Hypothèse : Modération via confiance

Hypothèse 4 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données sera moins forte si le consommateur a confiance dans l'entreprise en ligne (modération).

H 4.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur a confiance dans l'entreprise en ligne.

H 4.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur a confiance dans l'entreprise en ligne.

Ces hypothèses concernent la modération, ce qui signifie que la variable modératrice va influencer l'effet de la variable indépendante sur la variable dépendante. Il est dès lors supposé que la confiance dans l'entreprise (variable « Conf ») modère l'effet entre les risques perçus (« Risques ») et l'intention de divulguer des données personnelles. Afin de tester les deux sous-hypothèses, nous avons testé deux modérations, avec la variable dépendante qui change entre «Divulg 1 » (intention de divulguer des données personnelles explicites) et « Divulg 2 » (intention de divulguer des données personnelles implicites). Nous avons effectué cela via une Macro PROCESS sur base du modèle 1.

Les résultats du test **H4.1**, donc avec la variable dépendante concernant les données explicites, n'indiquent pas d'effet significatif pour l'interaction (p–valeur = 0,4873). Ceci ne nous permet pas d'affirmer que l'effet des risques sur la divulgation varie en fonction de la confiance. Cette hypothèse est **rejetée**. Cependant, nous constatons que l'effet simple de la confiance sur l'intention de divulgation des données personnelles explicites est significatif (p-valeur = 0,0495). L'ensemble des résultats est présenté à l'annexe n°11 (E).

De manière similaire, pour **H4.2**, l'interaction n'est pas significative (p-valeur = 0,1879), contrairement à l'effet simple de la confiance qui est significatif (p-valeur = 0,0365). Toutefois, cela ne nous permet pas d'affirmer que l'effet négatif des risques sur l'intention de divulgation des données personnelles implicites varie en fonction de la confiance. L'hypothèse H4.2 est donc **rejetée**. L'ensemble des résultats est présenté à l'annexe n°11 (F).

2.5.4. Hypothèse : Modération via le contrôle

Hypothèse 5 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données sera moins forte si le consommateur perçoit du contrôle sur ses données (modération).

> **H 5.1 :** L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur perçoit du contrôle sur ses données.
>
> **H 5.2 :** L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur perçoit du contrôle sur ses données.

Les hypothèses suivantes concernent également la modération. Il est dès lors supposé que la variable « control » modère l'effet entre les « risques » et l'intention de divulgation. Afin de tester les deux sous-hypothèses, nous avons testé deux modérations, avec la variable dépendante qui change entre « Divulg 1 » (intention de divulguer des données personnelles explicites) et « Divulg 2 » (intention de divulguer des données personnelles implicites). Nous avons effectué cela via une Macro PROCESS sur base du modèle 1.

Pour **H5.1**, les résultats du test n'indiquent pas d'effet significatif pour cette interaction ni pour l'effet simple du contrôle (p –valeur = 0,4958 et 0,6100 respectivement). Ceci ne nous permet pas d'affirmer que l'effet des risques sur l'intention de divulguer des données personnelles explicites varie en

fonction du contrôle. Cette hypothèse est **rejetée**. L'ensemble des résultats est présenté à l'annexe n°11 (I).

Ceci est également le cas pour **H5.2**, car ni l'effet de l'interaction ni l'effet simple du contrôle n'est significatif (p –valeur = 0,1671 et 0,7096 respectivement). Nous ne pouvons donc pas affirmer que l'effet des risques sur l'intention de divulguer des données personnelles implicites varie en fonction de la transparence. Cette hypothèse est **rejetée**. L'ensemble des résultats est présenté à l'annexe n°11 (J).
En résumé, il semblerait que le fait d'avoir un certain contrôle sur ses données ne modère pas la relation négative entre les risques perçus et l'intention de divulgation des données personnelles.

2.5.5. Hypothèse : Modération via la transparence

Hypothèse 6 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données (modération).

H 6.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données.

H 6.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données.

Les hypothèses suivantes concernent également la modération. Il est dès lors supposé que la transparence (variable « Transpa ») modère l'effet entre les « risques» et l'intention de divulgation des données personnelles. Afin de tester les deux sous-hypothèses, nous avons testé deux modérations, avec la variable dépendante qui change entre « Divulg 1 » (intention de divulguer des données personnelles explicites) et « Divulg 2 » (intention de divulguer des données personnelles implicites). Nous avons effectué cela via une Macro PROCESS sur base du modèle 1.

Pour **H6.1**, les résultats du test n'indiquent pas d'effet significatif pour cette interaction ni pour l'effet simple de la transparence (p–valeur = 0,8168 et 0,6686 respectivement). Ceci ne nous permet pas d'affirmer que l'effet des risques sur l'intention de divulgation des données personnelles explicites varie en fonction de la transparence. Cette hypothèse est **rejetée**. L'ensemble des résultats est présenté à l'annexe n°11 (I).

Ceci est également le cas pour **H6.2**, car ni l'effet de l'interaction ni l'effet simple de la transparence n'est significatif (p –valeur = 0,7656 et 0,5330 respectivement). Nous ne pouvons donc pas affirmer que l'effet des risques sur l'intention de divulgation des données personnelles implicites varie en fonction de la transparence. Cette hypothèse est **rejetée**. L'ensemble des résultats est présenté à l'annexe n°11 (J).

Dès lors, la transparence ne modère pas la relation entre risques perçus et intention de divulguer des données personnelles.

Afin d'éventuellement trouver des éléments de réponses aux effets non-significatifs de la modération du contrôle et de la transparence, nous avons effectué des tests supplémentaires qui se trouvent à l'annexe n°12. Premièrement, nous avons vérifié à l'aide d'un test ANOVA si l'existait un effet simple de ces variables dichotomiques sur les variables du *privacy calculus*, notamment sur les bénéfices perçus, les risques perçus et sur les deux types d'intention de divulgation. Cependant, aucun effet simple n'a été relevé (voir annexe n°12 (A) et (B)). Ensuite, nous avons testé si ces deux variables avaient une influence significative uniquement en cas d'effet combiné. Pour cela, une analyse de variance univariée a été effectuée mais n'a également pas été concluante (voir annexe n°12 (C)). Enfin, avons contrôlé si ces effets de modérations n'étaient pas significatifs dû au fait que les variables de transparence et contrôle s'influencent entre elles. Pour cela, nous avons échangé leurs valeurs et réalisé des tests ANOVA mais n'avons également pas obtenu de résultat significatif (voir annexe n°12 (D)). Il n'est donc pas possible de relever un quelconque effet significatif pour ces variables dans notre étude.

Figure n° 15 : Récapitulatif des résultats des tests d'hypothèses

	Hypothèses	Conclusion
H1	Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation). • H1.1 : Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **explicites.** • H1.2 : Les bénéfices perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **implicites.**	✓ ✓ ✓
H2	Les risques perçus de la personnalisation ont une influence négative sur l'intention de partager des données personnelles en ligne (et donc, de bénéficier de la personnalisation). • H2.1 : Les risques perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **explicites** • H2.2 : Les risques perçus de la personnalisation ont une influence positive sur l'intention de partager des données personnelles **implicites**	✓ ✓ ✓
H3	Les inquiétudes en matière de vie privée ont une influence positive sur les risques perçus de la personnalisation	✓
H4	L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur a confiance dans l'entreprise en ligne. • H4.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur a confiance dans l'entreprise en ligne. • H 4.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur a confiance dans l'entreprise en ligne.	✗ ✗ ✗
H5	L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur perçoit du contrôle sur ces données. • H 5.1 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur perçoit du contrôle sur ses données. • H 5.2 : L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur perçoit du contrôle sur ses données.	✗ ✗ ✗

H6	L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données. • **H 6.1 :** L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles explicites sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données. • **H 6.2 :** L'influence négative des risques perçus de la personnalisation sur l'intention de partager des données personnelles implicites sera moins forte si le consommateur perçoit de la transparence en ce qui concerne le traitement de ses données.	✗ ✗ ✗

3. Discussion des résultats

Dans le chapitre précédent, nous avons testé statistiquement toutes nos hypothèses afin de vérifier les suppositions faites sur base de la littérature. Dès lors, dans ce qui suit, nous allons tout d'abord synthétiser et tenter de comprendre les résultats obtenus. Ensuite, nous présenterons les implications managériales ainsi qu'un recul critique par rapport au RGPD et mentionneront enfin les limites de notre étude.

3.1. Interprétation des résultats

Suite à l'analyse statistique de nos résultats dans la section précédente, validant 3 de nos hypothèses et rejetant les 2 autres, nous nous devons d'interpréter ces résultats en les comparant aux résultats obtenus lors d'études précédentes.

Concentrons–nous tout d'abord sur le modèle du *privacy calculus*, à savoir les relations entre les bénéfices et risques perçus et l'intention de divulguer de l'information personnelle. A partir des résultats statistiques, nous pouvons valider la base de notre modèle de recherche. En effet, les résultats confirment bien que tant les bénéfices perçus que les risques perçus sont pris en compte lors de la prise de décision de divulguer de l'information personnelle en vue de bénéficier de la personnalisation. La personnalisation en elle-même motive donc la divulgation de l'information. Par ailleurs, les effets des variables sont bien opposés et les bénéfices semblent l'emporter sur les risques. Cette idée de calcul entre ces deux éléments contraires est donc en accord avec le concept de *privacy calculus* dans un contexte lié à la protection de la vie privée de Culnan et Bies (2003) et les résultats obtenus sont similaires à ceux des études de Chellappa & Sin (2005), Xu et al. (2011) et Xu et al. (2010). Cela confirme également que le concept de vie privée n'est pas absolu ; les consommateurs sont prêts à céder un peu de leur vie privée si des bénéfices, même non-monétaires comme c'est régulièrement le cas avec la personnalisation, sont offerts en échange. En ce qui concerne le type de données, nous observons que les individus partagent tout type de données personnelles, ne restreignant pas les possibilités de la personnalisation. Cependant, les répondants ont exprimé une préférence pour des données explicites telles que l'âge, l'adresse e-mail, le nom etc. La littérature n'est pas unanime sur ce point mais l'étude de Metzger et Phelps (cités dans Kobsa, 2007) avait également montré que les participants sont plus disposés à fournir des données démographiques basiques (âge, sexe, éducation) que l'information à propos du comportement réel en ligne ou des données plus sensibles. Cela peut être dû au fait que, comme le soulignaient (Karwatzki et al., 2017), les individus réalisent clairement quelles données seront utilisées car c'est eux qui les ont introduites, tandis que pour des données

implicites, l'enregistrement et l'utilisation de ces données sont plus flous. Par ailleurs, les chercheurs dont Xu et al. (2011), ont régulièrement souligné que l'approche de la personnalisation la plus invasive est celle où la localisation est connue et utilisée, une information faisant partie des données implicites que le consommateur peut accepter de partager avec l'entreprise.

Ensuite, les résultats soutiennent l'hypothèse selon laquelle les inquiétudes en matière de vie privée impactent les risques perçus de la personnalisation. Ceci est en accord avec les études de Malhotra, Kim et Agarwal (2004) et Fortes et Rita (2016). Les traits de personnalités jouent donc bien un rôle dans ce processus.

En ce qui concerne la modération, les résultats étaient moins concluants. Etonnement, aucunes de nos variables modératrices, à savoir la confiance, la transparence et le contrôle, n'ont montré d'effet significatif sur la relation entre les risques perçus et l'intention de divulgation. Cela signifie que l'impact des risques perçus sur l'intention de divulguer des données personnelles ne varie pas en fonction du niveau de confiance, de transparence ou de contrôle. Ceci va à l'encontre des idées reçues de la littérature existante, bien que le rôle exact de ces variables n'a pas été clairement défini. Cependant, nous avons pu identifier des raisons qui pourraient expliquer ces résultats.

Concernant la confiance, une explication probable est que, l'entreprise choisie pour l'expérimentation n'était pas optimale. En effet, sur base du tableau de fréquence de ce facteur, nous observons un manque de variance dans les réponses des individus aux questions liées à la confiance car la majorité a opté pour une réponse neutre (annexe n°11(E)). Bien qu'il y ait une question filtre, nous pouvons supposer que les individus ne sont pas assez familiers avec la marque que pour avoir une opinion prononcée en ce qui concerne la fiabilité, l'expertise et la bienveillance de celle-ci. Par ailleurs, Norberg, Horne et Horne (2007) mentionnaient qu'il se peut que la confiance ne montre pas d'effet significatif car elle est surtout susceptible de fonctionner dans des conditions extrêmes, lorsque des données sont très sensibles par exemple. Ceci n'est pas nécessairement le cas avec une entreprise d'actualité telle que l'Avenir. Concernant l'effet simple de la confiance sur l'intention de divulgation qui est apparu dans les résultats de la modération, nous pouvons penser que cette variable peut avoir un effet direct, indépendamment des variables de risques et bénéfices. Ainsi, Morosan & DeFranco (2015) suggèrent de prendre en considération, d'une part la valeur perçue de la divulgation, obtenue à partir de l'évaluation des risques et bénéfices perçus, et d'autre part, la confiance dans l'entreprise en ligne comme facteurs ayant un impact sur l'intention de divulgation de l'information. Enfin, selon Norberg, Horne et Horne (2007) et Milne et Boza (1999), la confiance aurait un impact sur le comportement réel, tandis que lorsqu'on interroge une personne sur ses intentions de divulgation, l'évaluation du risque influence considérablement la réponse.

Enfin, nous pouvons tenter de comprendre les effets non-significatifs de la transparence et du contrôle en se basant sur les résultats des études d'Awad & Krishnan (2006), de Norberg, Horne et Horne (2007) et sur la segmentation de la vie privée renommée de Westin (cité dans Awad & Krishnan, 2006). Dès lors, nous attribuons ce manque d'effet à la composition de notre échantillon.

En effet, comme nous l'avons déjà mentionné, un niveau d'éducation élevé en combinaison avec une utilisation d'internet élevée, et ce dans un pays où le besoin de contrôle, l'individualisme et la masculinité règnent comme valeurs culturelles, sont souvent le reflet d'un haut degré d'inquiétudes en matière de vie privée (Kobsa, 2007 ; Sheehan, 2002 ; Veltri, Krasnova & Elgarah, 2011). Notre échantillon aurait donc une tendance à accorder beaucoup d'importance à la protection de la vie privée, ce qui est également reflété dans le tableau des fréquences cette variable (annexe n°13). Par conséquent, il se pourrait que les individus ayant répondu à notre questionnaire soient des soi-disant « fondamentalistes de la protection de la vie privée» selon la classification de Westin (cité dans Awad & Krishnan, 2006) à propos des différentes attitudes possibles en matière de confidentialité. Les individus appartenant à ce groupe, représentant un peu plus du quart de la population (Kobsa, 2007), seraient extrêmement préoccupés par l'utilisation de leurs données et ne seraient pas disposés à divulguer des données personnelles à des sites web en général, contrairement aux pragmatistes et non-préoccupés. Tant la personnalisation que des mesures de protection de la vie privée mises en place ne changeront donc pas l'intention de divulgation de ces personnes particulièrement sensibles à la protection de leur intimité. Ainsi, Awad et Krishnan (2006) avaient également considéré que le paradoxe de la personnalisation et vie privée fait référence au fait que les personnes qui accordent le plus d'importance aux caractéristiques de protection de la vie privée, notamment à la transparence, sont peu disposées à être profilé en ligne à des fins de personnalisation.

Par ailleurs, une explication peut être que dans un contexte de confidentialité, les intentions revendiquées ne sont pas toujours le reflet des comportements, en raison de facteurs tels que le traitement d'heuristiques ou la routinisation du comportement influençant l'intention et le comportement très différemment (Norberg, Horne & Horne, 2007 ; Milne & Boza, 1999). Nous ne devons donc pas nous attendre à retrouver les mêmes résultats dans une situation réelle.

En résumé, ce manque de significativité implique que s'il existe bien un effet, il faudrait un échantillon beaucoup plus grand et varié en termes de traits de personnalités pour le détecter ainsi qu'une mise en situation réelle.

3.2. Recommandations managériales

Après mure réflexion et sur base des idées reçues de la littérature existante et des résultats de notre étude, nous souhaitons désormais formuler quelques recommandations à l'intention des entreprises, spécifiquement les spécialistes marketing. Nous finirons par proposer quelques suggestions aux consommateurs.

3.2.1. Allier confidentialité et *Big Data* en vue de la personnalisation en ligne

Notre modèle de recherche fournit de l'information quant à l'importance d'éléments fournis par l'entreprise, tels que les bénéfices de la personnalisation et des mécanismes de transparence et contrôle fournis, mais également de l'importance des caractéristiques et convictions du consommateur, à savoir les inquiétudes et risques perçus en matière de vie privée et la confiance, dans la prise de décision de participer ou non à la personnalisation. Ceci peut donc aider les spécialistes de marketing à prévoir quels facteurs influencent l'intention de divulgation.

Ainsi, les conclusions de notre étude mettent en exergue la nécessité pour les entreprises d'adapter leur stratégie marketing afin de rester en phase avec les nouvelles attentes, c'est-à-dire appréhender la personnalisation et la confidentialité comme des problématiques indissociables. Dès lors, nous recommandons aux entreprises de **tant s'engager dans la création de valeur via la personnalisation, que de mettre en place des méthodes pour réduire les préoccupations des consommateurs en matière de vie privée.**

3.2.2. Importance des bénéfices de la personnalisation

A l'heure actuelle, une entreprise ne peut plus se permettre de négliger cet aspect, tant d'un point de vue consommateur que concurrentiel. Effectivement, l'information récoltée dans le cadre de notre étude souligne que les consommateurs regardent au-delà des résultats négatifs de la divulgation, accordant une valeur à l'ensemble des effets de la personnalisation. Etant donné le poids que nos répondants ont attribué aux bénéfices de la personnalisation dans le cadre du processus de décision de divulgation, nous insistons sur le fait qu'il est utile d'investir dans les outils et autres ressources afin de proposer de la personnalisation. Pour cela, nous suggérons les bonnes pratiques suivantes : la mise en avant du consommateur et au centre de toute réflexion, l'apport d'une contrepartie à l'utilisation des données personnelles et la communication à ce propos, la bonne utilisation des informations recueillies, c'est-à-dire la pertinence et la qualité du contenu, et la compréhension du consommateur. Ce dernier signifie que les entreprises doivent prendre conscience des avantages et coûts que les utilisateurs perçoivent de la personnalisation et tenter de connaître les attitudes de ses clients à l'égard de la vie privée et de la personnalisation. Ceci permet d'orienter au mieux ses actions et ainsi éviter d'inutilement faire des efforts, au cas où certains clients sont des fondamentalistes de la protection de la vie privée, par exemple, comme ce fut le cas dans notre étude. Awad et Krishnan (2006) estiment en effet que les entreprises devraient uniquement se concentrer sur les consommateurs qui sont disposés à participer à la personnalisation en ligne dès le début. Une façon pour les marketers de réaliser cela serait de regrouper les clients particulièrement sensibles à la protection de la vie privée comme un segment différent.

Ensuite, nous conseillons aux entreprises d'être vigilantes lorsqu'ils récoltent et utilisent les données personnelles à propos du comportement en ligne du consommateur ou sa géolocalisation, car il est apparu dans notre étude que les individus sont plus réticents à divulguer ces données. Afin de tenir compte de cela, il serait judicieux que les entreprises informent clairement et en temps réel qu'elles collectent ces données personnelles, à l'aide des cookies par exemple. Ceci est pour l'instant rarement le cas car ces données ont tendance à être recueillies selon le principe de non-adhésion. Par ailleurs, nous supposons que cette réticence signifie également que les individus perçoivent plus d'intrusion à la vie privée en cas d'utilisation de ces données. Ainsi, les messages personnalisés portant sur ces éléments devraient éviter de les mentionner explicitement dans la communication mais le faire plus subtilement.

3.2.3. Importance du respect de la vie privée

Par ailleurs, bien que nous déduisions de notre étude que les bénéfices expliquent mieux la prise de décision de divulgation que les risques, ceux-ci sont tout de même bien présents. Dès lors, il serait opportun pour les entreprises, d'entre autres, uniquement demander les données pertinentes et

nécessaires pour le bon fonctionnement de leurs activités et d'entièrement intégrer les pratiques de protection de la vie privée dans leurs activités, plutôt que d'uniquement se restreindre aux normes légales à respecter. La réduction des risques fait sans aucun doute partie des voies à suivre dans un environnement qui sera encore plus connecté dans le futur. Dans cette perspective, les effets que nous avons observés risquent d'ailleurs de changer rapidement, notamment de par l'expansion des données nécessitées par l'évolution de l'Internet des objets.

3.2.4. L'éducation et autre voies fructueuses pour motiver le partage de l'information

Bien que la littérature ait suggéré que la confiance, la transparence et le contrôle des données par les consommateurs pouvaient être un exemple de combinaison essentielle afin que les consommateurs perçoivent une certaine équité et se sentent confortables à l'idée de divulguer leurs données personnelles, notre étude ne nous permet pas de confirmer cela. Cependant, en vue des changements à prévoir afin d'être conforme au RGPD, c'est-à-dire des mesures proches de celles que nous avons abordées, il nous semble quand même utile de proposer des recommandations sur base de toute l'information recueillie à ce sujet. Ainsi, nous sommes convaincus que les entreprises doivent faire preuve de transparence et céder le contrôle des données aux consommateurs, afin que tant les entreprises que les individus soient responsable de la protection des données. Pour cela, nous suggérons de présenter ces paramètres de contrôle et information à propos des pratiques de confidentialité de sorte qu'ils ne soient pas trop flous ni trop détaillés mais compréhensible par tous, facilement accessibles et apparents au moment où les consommateurs en ont le plus besoin et s'intégrant bien à la conception de la page web. Il s'agit donc de faire plus qu'uniquement publier des déclarations de confidentialité. Toutefois, il est primordiale de trouver un juste équilibre afin d'éviter de faire émerger davantage de préoccupations du côté des clients.

Par ailleurs, comme de nombreux chercheurs l'ont souligné, plutôt qu'uniquement tenter de réduire les risques, il semble que les entreprises pourraient également développer des actions afin d'établir ou renforcer les relations de confiance et ainsi indirectement impacter la perception des risques. Ceci serait même plus performant et facile à exécuter que les efforts visant à réduire les préoccupations (Karwatzki et al., 2017).

En outre, en lien avec la transparence, la littérature indique que les individus savent très peu de choses sur le fonctionnement de la personnalisation et le traitement des données, formant un obstacle à l'utilisation des services personnalisés. Ainsi, nous conseillons aux pouvoirs publics d'investir davantage dans l'éducation des individus afin de mieux les informer quant à leurs droits d'expression, de garder certaines informations secrètes, de pouvoir modifier des données etc., le RGPD étant une potentielle initiative pour solutionner cela.

Enfin, nous estimons que dans ce contexte de personnalisation et protection de la vie privée, le principe clé est la proactivité et non pas la réactivité, étant donné les lourdes conséquences que peuvent entrainer des failles de confidentialité ou ciblages non-adéquats. Ainsi, les entreprises gagnantes seront celles qui arriveront à atteindre un équilibre entre personnalisation et respect de la vie privée, ces deux éléments étant indispensables à la bonne expérience en ligne du consommateur. Nous

souhaitons finir cette section en rappelant que les entreprises et autres institutions doivent garder à l'esprit que les individus ne se comportent pas nécessairement de la façon qu'ils disent.

En ce qui concerne les consommateurs, nous leurs conseillons avant tout de continuer à faire confiance à des sites renommés. Ceci dit, au cours de nos recherches, nous avons pu remarquer que certaines entreprises interprètent les restrictions règlementaires de manière très différentes, apportant davantage de confusion au consommateur. Nous pensons par exemple aux bannières cookies, qui sont loin d'être uniformes et facilement interprétables par tout un chacun. Ainsi, nous recommandons aussi aux individus d'utiliser des mesures de protections telles qu'utiliser au maximum des pseudonymes ; le nom ou l'adresse e-mail n'étant régulièrement pas nécessaires à la réalisation des actions souhaitées et n'empêche pas entièrement la personnalisation, ainsi que maitriser leurs dispositifs et les sites web en gérant efficacement les paramètres. Par ailleurs, il est important qu'ils réalisent qu'il n'est pas possible de ne laisser aucune trace en ligne, que refuser les cookies ne signifie pas un accès limité au site web et qu'un service gratuit rime souvent avec exploitation des données. Nous ne disons en aucun cas qu'il ne faut pas révéler d'information en ligne, mais souhaitons proposer des éclairages sur la manière dont les individus peuvent profiter des bénéfices de la personnalisation tout en naviguant confortablement sur le web.

3.3. Limites de l'étude et voies futures de recherches

Comme pour toute recherche, notre étude comporte certaines limites qui doivent être prises en considération lors de la discussion des résultats et qui représentent autant de pistes d'amélioration et de recherches futures dans ce domaine.

Premièrement, en ce qui concerne l'échantillon, nous remarquons une certaine homogénéité au niveau de l'éducation, de l'utilisation d'internet, du trait de personnalité observé et fort probablement au niveau de la culture des répondants. Ces derniers sont sans doute plus expérimentés que l'internaute belge lambda et plus inquiets en matière de protection de la vie privée. Ceci est dû au fait que nous avons eu recours à une méthode d'échantillonnage non-probabiliste ; l'échantillonnage n'était pas entièrement aléatoire car il a été diffusé parmi notre réseau de proches, bien que l'effet boule de neige ait permis de minimiser ce biais. Il est donc nécessaire d'envisager que les conclusions pourraient être différentes pour d'autres répondants. Par ailleurs, des recherches futures devraient englober une plus grande diversité d'internautes pour valider et étendre les résultats de nos recherches.

Une autre limite de la méthode de collecte de nos données provient du fait que les sondages en ligne à propos de la protection de la vie privée ne constituent peut-être pas le meilleur moyen de mesurer l'attitude des personnes à l'égard de la protection de la vie privée. Des personnes extrêmement préoccupées par la question de la confidentialité pourraient avoir refusé de participer à notre étude.

Ensuite, bien que difficilement mesurable dans un contexte lié à la confidentialité, il serait utile d'étudier le comportement réel des consommateurs. Il semblerait, en effet, que lorsque les individus sont exposés à des décisions du monde réel, ils divulgueront plus d'information que ce que leurs intentions exprimées indiquent car les comportements reposeraient moins sur un processus cognitif,

étant plus fortement influencé par des éléments environnementaux, des états d'humeur ou autres (Norberg, Horne & Horne, 2007). Dès lors, nous ne pouvons pas supposer que l'intention de divulgation est entièrement liée au comportement réel.

Notons également que nos données ont été recueillies lors d'une période confuse de changement sur le plan des données privées en ligne, en commençant par le scandale lié à l'exploitation des données de Facebook suivi de la mise en place du RGPD. Nous imaginons que le processus de prise de décision de nos répondants puisse avoir été influencé par les messages alarmants sur la confidentialité des données, véhiculés par les médias, ce qui pourrait résulter en une évaluation des risques plus approfondie qu'en temps normal. Norberg, Horne et Horne (2007) affirmaient ainsi que lorsque l'on interroge une personne sur sa volonté de divulguer, on peut s'attendre à ce que les risques perçus soient prépondérants car l'individu aura tendance à se reposer sur des sources d'informations externes telles que les médias ou sur des stéréotypes.

Une cinquième limite porte sur le choix de l'entreprise utilisée dans notre expérimentation. Nous avons fait le choix de prendre la société éditrice l'Avenir comme base de nos scénarios, estimant que celle-ci provoquerait des niveaux distincts de confiance auprès des consommateurs. Cependant, il est apparu que tous les répondants semblent moyennement faire confiance à cette entreprise, les réponses n'étant pas réparties sur différents niveaux. Ceci pourrait expliquer le manque d'effet significatif de ce facteur. Ainsi, nous proposons qu'une future étude considère à nouveau le rôle de la confiance, de la même façon ou en tant que variable indépendante, ainsi qu'en analysant également son impact sur les bénéfices de la personnalisation et en utilisant éventuellement deux entreprises différentes pour les scénarios. Il est également important de rappeler que le comportement lié à la protection de la vie privée est un phénomène très contextuel. Les bénéfices perçus de la personnalisation ne l'emporteront pas dans toujours sur l'impact des risques perçus et les perceptions peuvent être différentes, voire entièrement opposées, dans des contextes différents de celui d'un site d'actualité, qui ne nécessite pas des données aussi sensibles que pour un contexte financier par exemple. La généralisation des résultats n'est donc pas sûre, il serait intéressant de répliquer cette recherche dans d'autres secteurs.

Par ailleurs, nous estimons avoir opté pour une manière neutre de présenter l'information quant au contrôle et à la transparence dans notre expérimentation, car notre objectif n'était pas d'étudier l'impact du choix du message ou du design des paramètres de contrôle. Cependant, il est évident que des heuristiques, telles que la manière dont les choses sont demandées et présentées, pourraient également influencer l'intention comportementale du consommateur. Ainsi, la manière idéale de fournir du contrôle et de la transparence dans ce type de contexte demeure une question ouverte. Il serait intéressant de comparer l'effet de la formulation de l'information et de la présentation dans des recherches futures. Dans cette même idée, l'effet de dualité de la transparence pour les consommateurs, mentionné dans la partie théorique de ce travail, laisse croire qu'un standard minimum de transparence est attendu et qu'aller au-delà de ce seuil donnerait un effet contraire. Il serait donc opportun d'appréhender dans de futurs travaux cette idée de seuil idéal de transparence.
En outre, d'autres variables comme les valeurs culturelles ou le design du site web pourraient influencer les perceptions des individus. Pour des raisons de faisabilité, ces variables n'ont pas pu être

étudiées dans cette étude mais il est recommandé de les mesurer dans le cadre de recherches ultérieures.

Enfin, une dernière limite concerne notre étude statistique car nous avons eu recours à une approche classique pour valider notre modèle quantitatif, plutôt qu'à une méthode d'équations structurelles, cette dernière servant à appréhender des phénomènes complexes et avec plus de précision de par le fait qu'elle permet de mettre en relation des concepts non observables.

3.4. Recul critique par rapport au RGPD

Avant de conclure ce mémoire, nous souhaitons porter un regard critique sur l'actualité, estimant que nous ne pouvons parler de confidentialité et transparence sans aborder le nouveau Règlement Général sur la Protection des Données dans l'Union Européenne (RGPD) entré en vigueur ce 25 mai 2018.

En effet, la collecte des données personnelles que nous avons abordée dans cette étude comporte tout un nouvel ensemble de défis à l'ère du RGPD. Celui-ci a pour but de renforcer et d'homogénéiser la protection en matière de données personnelles et sensibles aux consommateurs au sein de l'UE. Les consommateurs se verront désormais demandés uniquement des données pertinentes, seront clairement informés de la collecte et de l'utilisation de leurs données (type, raison et durée), devront donner leur consentement explicite avant que ces données ne soient stockées ou utilisées et auront la possibilité granulaire de refuser certains éléments ou se retirer facilement par après, quand ils le souhaitent, selon le principe du droit à l'oubli (Bergen, 2018). La question que nombreuses organisations traitant des données des citoyens de l'UE se posent alors est : dans quelle mesure cette nouvelle réglementation impactera les stratégies marketing nécessitant des données personnelles, tels que pour les stratégies de personnalisation ?

Tandis que nombreux sont ceux qui craignent que cette nouvelle réglementation forme une contrainte aux possibilités de marketing personnalisé, d'autres affirment qu'ils ont à y gagner, réalisant que cela peut améliorer l'expérience globale d'une manière qui profite à tous, spécialistes du marketing et consommateurs. Bien que nous en sommes encore qu'aux premières étapes du RGPD et que beaucoup de confusion persiste, nous rejoignons ce deuxième point de vue selon lequel le projet de loi est une mesure positive pour la personnalisation. Sur base de ce que la littérature a révélé et des hypothèses validées de notre étude, nous considérons que ce nouvel environnement réglementaire offre le respect que les consommateurs méritent. En effet, nous estimons que le RGPD répond enfin aux préoccupations et attentes des consommateurs liées à la confidentialité, comme développées dans ce travail, et met fin à la pratique quelque peu opportuniste de maximisation des données, du fait que la transparence et le consentement sont des éléments clés du RGDP, tout comme ils devraient l'être pour le marketing personnalisé. Si cette méfiance et inégalité avaient gagné en importance ces dernières années, c'est dû à l'essor plus rapide des technologies que des législations et à un flou juridique, permettant aux entreprises de jouir d'une liberté presque totale en matière de collecte et d'utilisation des données (Hérault, 2014). Il ne faut pas se le cacher, les consommateurs ont été, pendant de nombreuses années, en grande partie à la merci de marques qui n'ont pas toujours employé les pratiques les plus éthiques lorsqu'il s'agissait d'utiliser les données. Ainsi, l'objectif du RGPD est de

rééquilibrer le rapport de force entre le consommateur et les marques, dont nous avons largement discuté dans ce travail, en plaçant la transparence au centre de leurs échanges et donnant à l'individu le contrôle de ses données. Cela permettra de recréer un environnement de confiance, rendant les consommateurs potentiellement plus aptes à communiquer certaines informations et impactera la création d'une relation durable et « gagnant-gagnant » (Portes, 2018).

En réponse à la question posée ci-dessus, le RGPD n'empêchera pas, par principe, de personnaliser son contenu, cela ne change pas. Ainsi, les résultats de notre étude, bien que ne permettant pas de tirer de conclusions à propos de l'impact de ces mécanismes sur la divulgation, n'indiquent pas non plus que la transparence ou le contrôle réduiront significativement la divulgation des données. Les entreprises auront toujours des données sur lesquelles capitaliser.

Toutefois, les entreprises devront être plus vigilantes, c'est-à-dire utiliser ces données utilisateur sur base de l'intérêt légitime ou demander l'approbation des consommateurs en amont pour récolter et exploiter leurs données, pour chaque utilisation et de manière non-ambiguë (DMA, 2018). L'importance de consentement avait déjà été soulignée dans notre revue littéraire par Norberg, Horne et Horne (2007), affirmant que le moyen le plus efficace est que les consommateurs fassent partie de la première ligne de défense pour protéger leur vie privée. Comme nous en avons discuté, toutes les entreprises ont intérêt à offrir des expériences personnalisées et le RGPD n'est qu'un cadre plus strict forçant les entreprises à avoir une politique de gestion des données plus respectueuse et à proposer de la protection de manière proactive, dès la conception (*privaycy by design*).

Certes, se mettre en conformité avec la loi impose des nouvelles contraintes organisationnelles, techniques et réglementaires car cela aura un impact sur la manière dont les entreprises communiquent, gardent trace des consommateurs et gèrent les bases de données. Les entreprises devront notamment auditer leurs pratiques, nettoyer leurs bases de données, redemander la permission de garder les données recueillies avant le RGDP et pouvoir accéder rapidement aux données du profil du consommateur s'il demande de modifier ou supprimer ses données. En vue de ce changement, les entreprises auront aussi besoin d'un certain niveau de formation juridique pour les aider dans la compréhension des règles, devront mettre en place un processus de violation des données et nommer un responsable de la conformité (Rizkallah, 2018). Ensuite, elles risquent d'avoir moins de données sur lesquelles travailler, mais les données dont elles disposeront seront de meilleure qualité et les prospects plus engagés. En effet, comme nous l'avons vu dans la littérature, les individus seront plus disposés à fournir des données personnelles s'il y a une contrepartie et peu de risques perçus; l'équité pourrait donc constituer une incitation à la divulgation. En d'autres mots, si les clients donnent leur consentement avec le RGPD, c'est qu'ils veulent réellement entamer une relation et recevoir de la valeur en échange. Même si les clients se retirent ou ne participent pas, cela peut être une bonne indication pour la marque car cela permet de comprendre que quelque chose ne va pas.

Autrement dit, toutes les opportunités développées auparavant sont plus importantes que tous ces coûts (Bergen, 218). Les entreprises n'ont donc pas le choix de devenir conformes. D'une part, car les pratiques antérieures devenaient de moins en moins efficaces. D'autre part, car c'est ce que les

consommateurs veulent. Si les entreprises ne respectent par ces nouvelles restrictions, le consommateur ira voir ailleurs.

En conclusion, dans le cadre de notre étude et du RGPD, le challenge actuel pour les entreprises sera de réfléchir par quels moyens plus éthiques ils pourront dépasser le paradoxe personnalisation-protection des données personnelles. Plutôt que de le considérer comme une contrainte réglementaire en plus, les entreprises devraient envisager ce nouveau règlement comme une opportunité pour mieux utiliser les outils CRM , innover en matière de confidentialité et personnalisation et ainsi regagner la confiance des consommateurs. Le RGPD suit notre idée mise en exergue tout au long de ce travail selon laquelle une bonne pratique en matière de protection de la vie privée, permet une bonne pratique de marketing. Bien sûre, comme tout changement, le RGPD nécessitera un temps d'adaptation et de compréhension.

CONCLUSION

Dans le cadre de notre mémoire, nous avons souhaité nous intéresser au paradoxe entre la personnalisation et la protection des données personnelles, un phénomène au cœur des débats actuels. Ainsi, l'objectif de cette étude était de comprendre les raisons pour lesquelles les consommateurs acceptent ou rejettent la participation à la personnalisation en ligne, malgré les craintes relatives à la protection des données. Plus précisément, il s'agissait de déterminer l'effet de la vulnérabilité des données perçue par les clients sur leur intention de divulgation en vue de bénéficier de la personnalisation, ainsi que le rôle des principaux mécanismes d'atténuation des risques. A notre connaissance, aucune étude n'avait jusqu'alors étudié l'influence de ces mécanismes d'une telle manière. Le but final était de pouvoir formuler des recommandations aux entreprises afin qu'elles puissent offrir des stratégies de personnalisation qui soient acceptables et avec lesquelles les consommateurs se sentent confortables. Pour ce faire, nous avons, d'une part, proposé un modèle de recherche fondé sur une revue de la littérature, d'autre part, mené une analyse quantitative afin de tester nos hypothèses.

A l'issue de la phase d'exploration, nous avons constaté que les individus n'ont plus le temps pour des milliers de messages marketings non pertinents ; ils s'attendent à mieux et surtout à ce que les entreprises comprennent leurs besoins et préférences. Ainsi, la personnalisation est désormais un attendu des internautes. Nous avons aussi mis en exergue que la personnalisation repose sur la collecte des données personnelles, obtenues à l'aide de technologies sophistiquées, et que c'est exactement cette pratique intrusive qui rend les consommateurs hésitants à utiliser les services personnalisés. En effet, certains consommateurs craignent et estiment que les entreprises usent et abusent dans certains cas de toute cette information mise à leur disposition, ce qui empiète sur leur désir de protéger leurs vies privées et renforce l'inégalité entre ces deux parties. Le respect de la vie privée et l'assurance d'équité sont donc également des attendus. Toutefois, nous avons aussi évoqué que ces craintes ne sont pas reflétées dans leurs comportement, les individus étant prêts à divulguer de l'information personnelle s'ils reçoivent une certaine valeur ajoutée en retour de leurs données, si des pratiques équitables d'information et de protection sont établies afin de réduire les risques ou si un lien de confiance a déjà été établi. Basé sur ces observations, nous avons identifié plusieurs facteurs susceptibles d'influencer l'intention de cette divulgation. Nous avons ainsi posé que les critères sur lesquels se fondent majoritairement la décision de dévoiler des données personnelles sont les bénéfices et risques perçus de la personnalisation, la force de ce dernier sur l'intention de divulgation pouvant être atténuée par la confiance, la transparence par rapport au traitement des données, le contrôle des données par le consommateur et influencé directement par les inquiétudes en matière de vie privée.

Une fois le modèle de recherche établit, nous avons effectué une étude quantitative afin de tester les hypothèses posées. Cela a été possible grâce à la diffusion d'un questionnaire en ligne, contenant 4 scénarios, faisant varier la transparence et le contrôle. Les résultats confirment l'existence du *privacy calculus* dans le contexte de la personnalisation. Ainsi, il apparait que non seulement, les bénéfices perçus de la personnalisation mais également les risques perçus en matière de vie privée ont une influence sur l'intention de divulguer des données personnelles et que les bénéfices l'emporteraient sur

les risques. Notre analyse soutient également l'hypothèse suggérant que les traits de personnalités, dans notre cas représentés par la tendance à être inquiets en ce qui concerne la protection de ses données personnelles, ont un rôle dans la prise de décision de divulgation. Les données récoltées mettent aussi en avant l'existence d'une préférence pour la divulgation des données explicites, en opposition aux données implicites. Cependant, les interactions entre la confiance, la transparence et le contrôle n'ont pas pu être confirmées, bien que l'effet simple de la confiance révèle un certain effet dans la direction attendue.

Ces résultats montrent bien l'importance pour les entreprises de trouver un équilibre entre respect de la vie privée et de la personnalisation, qui fera que les consommateurs se sentent confortables avec la divulgation. Nos résultats non-concluants portent également à croire qu'il existe bien une différence entre les intentions et les comportements dans ce contexte ainsi qu'un segment de fondamentalistes en matière de protection de la vie privée, une catégorie de la population qui est extrêmement préoccupée par l'utilisation des données personnelles et n'est généralement pas disposée à fournir ces informations aux sites web, même lorsque des mesures de protection de la vie privée sont mises en place. L'analyse de nos résultats ne nous permet donc pas de répondre à la question de recherche initiale. Malgré tout, des éléments de réponse ont été identifiés grâce à la littérature.

Sur cette base, et bien que nous sommes conscients que notre recherche présente certaines limites, plusieurs recommandations aux entreprises ont été formulées. Ainsi, il est nécessaire que l'entreprise veille à mettre l'utilisateur au centre de toute réflexion et surpasse le simple fait d'offrir de la personnalisation, en évitant toute communication intrusive et en tenant compte des préoccupations des consommateurs en matière de confidentialité. Par ailleurs, nous suggérons aux spécialistes de marketing, d'entre autres, rassurer les individus au moyen de pratiques équitables de traitement des données allant au-delà de la conformité, d' instaurer un environnement de confiance et de considérer le RGPD comme une opportunité afin de développer des tactiques plus créatives, pouvant encourager les consommateurs à partager de l'information. Enfin, nous proposons aux consommateurs d'être vigilants tout en gardant à l'esprit que leurs droits seront de plus en plus protégés et régulés.

Cependant la prudence s'impose quant à la généralisation de ces résultats car notre étude comporte plusieurs limites, tant au niveau de l'échantillon que de la méthode d'étude. Dès lors, ce mémoire pourrait être un point de départ pour des recherches futures. Des pistes intéressantes seraient d'incorporer d'autres variables à ce modèle, de l'appliquer à d'autres contextes ou encore d'étudier l'effet des comportements réel plutôt que les déclarés et surtout auprès d'un échantillon plus diversifié. Il serait également pertinent de mener des études afin de déceler la formulation et présentation idéale des mécanismes de protection.

Pour conclure, nous espérons sincèrement que nos résultats susciteront l'intérêt d'autre chercheurs du domaine afin d'apporter plus de réponses et des pistes d'actions concrètes à ce paradoxe complexe. Ceci est nécessaire car, en vue de l'importance de la récolte des données et de l'amélioration des technologies d'analyses, nous supposons que le défi éthique, technique et économique lié à la stratégie de personnalisation restera un enjeu majeur dans l'avenir proche.

BIBLIOGRAPHIE

Abraham, M., Mitchelmore, S., Collins, S. Maness, J., Kistulinec, M., Khodabandeh, S. Hoenig, D., & Visser, V. (2017). *Profiting from Personalization.* Récupéré de https://www.bcg.com/en-be/publications/2017/retail-marketing-sales-profiting-personalization.aspx

Accenture Interactive (2018). *Making it Personal. Pulse Check 2018.* Récupéré de https://www.accenture.com/t20161011T222718__w__/us-en/_acnmedia/PDF-34/Accenture-Pulse-Check-Dive-Key-Findings-Personalized-Experiences.pdf

Accenture Strategy (2017). *What today's consumers want isn't what you think*. Récupéré de https://www.accenture.com/t20171220T024439Z__w__/us-en/_acnmedia/PDF-68/Accenture-Global-Anthem-POV.pdf#zoom=50

Adjerid, I. Acquisti, A., Brandimarte, L., & Loewenstein, G. (2013). Sleights of Privacy: Framing, Disclosures, and the Limits of Transparency. *Symposium on Usable Privacy and Security (SOUPS 2013).* Newcastle, UK. Récupéré de https://cups.cs.cmu.edu/soups/2013/proceedings/a9_Adjerid.pdf

Adomavicius, G., & Tuzhilin, A. (2001). Using data mining methods to build customer profiles. *Communications of the ACM, 34*(2), 74-82. doi: 10.1109/2.901170

Aguirre, E., Mahr, D., Grewal, D., De Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing, 91*(1), 34-49. doi:10.1016/j.jretai.2014.09.005

Alekh, S. (2018). *Human Aspects and Perception of Privacy in Relation to Personalization.* Seminar on Privacy and Big Data, Aachen, Germany. Récupéré de https://arxiv.org/ftp/arxiv/papers/1805/1805.08280.pdf

Alhouti, S., Johnson, C. M., & D'Souza, G. (2016). The complex web of values: The impact on online privacy concerns and purchase behavior. *Journal of Electronic Commerce Research, 17*(1), 22.

Ansari, A., & Mela, C. F. (2003). E-customization. *Journal of Marketing Research, 40*(2), 131-145. doi:10.1509/jmkr.40.2.131.19224

Arora, N., Dreze, X., Ghose, A., Hess, J. D., Iyengar, R., Jing, B., Joshi, Y., Kumar, V., Lurie, N., Neslin, S., Sajeesh, S., Su, M., Syam, N., Thomas, J., & Zhang, Z. J. (2008). Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters, 19*(3/4), 305-321. doi:10.1007/s11002-008-9056-z

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quarterly, 30*(1), 13-28.

Baek, T. H., & Morimoto, M. (2012). Stay away from me: Examining the determinants of consumer avoidance of personalized advertising. *Journal of Advertising, 41*(1), 59. doi:10.2753/JOA0091-3367410105

Bansal, G., Zahedi, F., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems, 49*(2), 138-150. doi:10.1016/j.dss.2010.01.010

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review: Privacy concerns meta-analysis. *Journal of Communication, 67*(1), 26-53. doi:10.1111/jcom.12276

Beke, F.T., Eggers, F., & Verhoef, P.C. (2018). Consumer Informational Privacy: Current Knowledge and Research Directions. *Foundations and Trends® in Marketing, 11*(1), 1-71. http://dx.doi.org/10.1561/1700000057

Bergen, B. (2018). *GDPR : What all Marketers need to know*. Récupéré de https://www.salesforce.com/blog/2018/04/gdpr-what-marketers-need-to-know.html

Capgemini Consulting (2017). *Hyper-personnalisation vs segmention. A-t-on encore besoin de segmenter les clients à l'ère du Big Data ?* Récupéré de https://www.capgemini.com/consulting-fr/wp-content/uploads/sites/31/2017/08/hyperpersonnalisation_vs_segmentation_25-11.pdf

Chang, S. E., Changchien, S. W., & Huang, R. (2006). Assessing users' product-specific knowledge for personalization in electronic commerce. *Expert Systems with Applications, 30*(4), 682-693. doi:10.1016/j.eswa.2005.07.021

Chau, P. Y. K., Ho, K. K. W., Ho, S. Y., & Yao, Y. (2013). Examining the effects of malfunctioning personalized services on online users' distrust and behaviors. *Decision Support Systems, 56*, 180-191. doi:10.1016/j.dss.2013.05.023

Chellappa, R. K., & Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management, 6*(2-3), 181. doi:10.1007/s10799-005-5879-y

Chellappa, R. K., & Shivendu, S. (2007). An economic model of privacy: A property rights approach to regulatory choices for online personalization. *Journal of Management Information Systems, 24*(3), 193-225. doi:10.2753/MIS0742-1222240307

Chen, D., Hu, P. J., Kuo, Y., & Liang, T. (2010). A web-based personalized recommendation system for mobile phone selection: Design, implementation, and evaluation. *Expert Systems with Applications, 37*(12), 8201-8210. doi:10.1016/j.eswa.2010.05.066

Chung, T. S., Wedel, M., & Rust, R. T. (2016;;). Adaptive personalization using social networks. *Journal of the Academy of Marketing Science, 44*(1), 66-87. doi:10.1007/s11747-015-0441-x

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science, 10*(1), 104-115. doi:10.1287/orsc.10.1.104

Culnan, M. J., & Bies, R. J. (2003). Consumer privacy: Balancing economic and justice considerations. *Journal of Social Issues, 59*(2), 323-342. doi:10.1111/1540-4560.00067

De Filippi, P. (2016). Gouvernance algorithmique: Vie privée et autonomie individuelle à l'ère des big data. *Open data & Big data ; Nouveaux défis pour la vie privée*. Paris: Éditions Mare et Martin. Récupéré de https://hal.archives-ouvertes.fr/hal-01382010/document

Deloitte (2015). *The Deloitte Consumer Review Made-to-order : The rise of mass personalisation.* Récupéré de https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/consumer-business/ch-en-consumer-business-made-to-order-consumer-review.pdf

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., & Colautti, C. (2006). Privacy calculus model in e-commerce – a study of italy and the united states. *European Journal of Information Systems, 15*(4), 389-402. doi:10.1057/palgrave.ejis.3000590

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for E-commerce transactions. *Information Systems Research, 17*(1), 61-80. doi:10.1287/isre.1060.0080

Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems, 22*(3), 295-316. doi:10.1057/ejis.2012.23

DMA (2018). *GDPR for marketers: consent and legitimate interest*. Récupéré de https://dma.org.uk/uploads/misc/5ae1fbf5c60fd-gdpr-for-marketers---consent-and-legitimate-interest_5ae1fbf5c6066.pdf

Doney, P. M., & Cannon, J. P. (1997). An examination of the nature of trust in buyer-seller relationships. *Journal of Marketing, 61*(2), 35-51. doi:10.2307/1251829

Eirinaki, M. & Vazirgiannis, M. (2003). Web mining for web personalization. *ACM Transactions on Internet Technology*, 3(1), 1-27. doi: 10.1145/643477.643478

Evergage (2018). *2018 Trends in Personalization.* Récupéré de http://www.evergage.com/resources/ebooks/trends-in-personalization-survey-report/

Garcia-Rivadulla, S. (2016). Personalization vs. privacy: An inevitable trade-off? *IFLA Journal, 42*(3), 227-238. doi:10.1177/0340035216662890

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly, 27*(1), 51-90. doi:10.2307/30036519

Grenouilleau, H. (2015). *Les enjeux du marketing personnalisé pour votre entreprise.* Récupéré de https://www.alesiacom.com/blog/les-enjeux-du-marketing-personnalise-pour-votre-entreprise

Hackett, R. (2015). *Apple CEO Tim Cook's privacy letter is a huge shot at Google*. Récupéré de http://fortune.com/2015/09/29/apple-tim-cook-google/

Hargittai, E., & Marwick, A. (2016). "What can I really do?" explaining the privacy paradox with online apathy. *International Journal of Communication (Online), 10*, 3737-3757.http://web.a.ebscohost.com.proxy.bib.ucl.ac.be/ehost/pdfviewer/pdfviewer?vid=1&sid=eafbbbaa-3f1b-473b-aa8e-be03f9b4e3d9%40sessionmgr4010

Heath, A. (2017, 14 septembre). 'Big data' is the black gold of today. It's time for politicians to catch up. *The Telegraph.* Récupéré de https://www.telegraph.co.uk/news/2017/09/14/big-data-black-gold-today-time-politicians-catch/

Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., & Skiera, B. (2010). The impact of new media on customer relationships. *Journal of Service Research, 13*(3), 311-330. doi:10.1177/1094670510375460

Hérault, S., & Belvaux, B. (2014). Privacy paradox et adoption de technologies intrusives le cas de la géolocalisation mobile. *Décisions Marketing*, (74), 67-82.

Ho, S. Y., & Bodoff, D. (2014). The effects of web personalization on user attitude and behavior: An integration of the elaboration likelihood model and consumer search theory. *MIS Quarterly, 38*(2), 497.

Hofstede Insights (2018). *What about Belgium?* Récupéré de https://www.hofstede-insights.com/country/belgium/

Hoy, M. G., & Milne, G. (2010). Gender differences in privacy-related measures for young adult facebook users. *Journal of Interactive Advertising, 10*(2), 28-45. doi:10.1080/15252019.2010.10722168

Febelfin (2017). *La banque numérique et votre vie privée: une question de confiance réciproque.* Récupéré de http://files.febelfin.be/Banque_numerique-vie_privee/files/assets/common/downloads/publication.pdf

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics (4th Edition)*. London, United Kingdom : SAGE Publication Limited, London.

Fortes, N., & Rita, P. (2016). Privacy concerns and online purchasing behaviour: Towards an integrated model. European Research on Management and Business Economics, 22(3), 167. doi:10.1016/j.iedeen.2016.04.002

Guo, X., Zhang, X., & Sun, Y. (2016). The privacy–personalization paradox in mHealth services acceptance of different age groups. *Electronic Commerce Research and Applications, 16*, 55-65. doi:10.1016/j.elerap.2015.11.001

Karwatzki, S., Dytynko, O., Trenz, M., & Veit, D. (2017). Beyond the personalization-privacy paradox: Privacy valuation, transparency features, and service personalization. *Journal of Management Information Systems, 34*(2), 369. doi:10.1080/07421222.2017.1334467

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal, 25*(6), 607-635. doi:10.1111/isj.12062

Khelladi, I., Castellano, S., & Limongi, L. (2014). L'impact de la personnalisation basée sur le profil et la localisation sur le comportement du client dans un contexte de téléphonie mobile/impact of profile and location-based personalization on customer behavior in a mobile context. *Revue Française Du Marketing*, (248), 43.

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems, 44*(2), 544-564. doi:10.1016/j.dss.2007.07.001

Kobsa, A. (2007). Privacy-enhanced personalization. *Communications of the ACM, 50*(8), 24-33. doi:10.1145/1278201.1278202

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security, 64*, 122-134. doi:10.1016/j.cose.2015.07.002

Komiak, S.Y.X, & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly, 30*(4), 941-960

KPMG (2017). *Data: this time its personal.* Récupéré de https://home.kpmg.com/xx/en/home/insights/2017/01/data-this-time-its-personal.html

KPMG (2017). *Global retail trends 2017*. Récupéré de https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/06/retail-trends.pdf

KPMG (2017). *The truth about online consumers. 2017 Global Online Consumer Report*. Récupéré de https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/01/the-truth-about-online-consumers.pdf

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal of Information Technology, 25*(2), 109-125. doi:10.1057/jit.2010.6

Kuin, T. (2018). *How to start with digital marketing personalization*. Récupéré de https://www.accenture-insights.nl/en-us/articles/digital-market-personalization

Lambin, J., & de Moerloose, C. (2012). *Marketing stratégique et opérationnel: Du marketing à l'orientation-marché* (8e éd. [actualisée] ed.). Paris: Dunod.

Lanier, C. D., Jr, & Saini, A. (2008). Understanding consumer privacy: A review and future directions. *Academy of Marketing Science Review, 2008*, 1.

Lee, C. H., & Cranage, D. A. (2011). Personalisation–privacy paradox: The effects of personalisation and privacy assurance on customer responses to travel web sites. *Tourism Management, 32*(5), 987-994. doi:10.1016/j.tourman.2010.08.011

Lee, K. C., & Kwon, S. (2008). Online shopping recommendation mechanism and its influence on consumer decisions and behaviors: A causal map approach. *Expert Systems with Applications, 35*(4), 1567-1574. doi:10.1016/j.eswa.2007.08.109

Li, T., & Unger, T. (2012). Willing to pay for quality personalization? Trade-off between quality and privacy. *European Journal of Information Systems, 21*(6), 621-642. doi:10.1057/ejis.2012.13

Liang, T., Lai, H., & Ku, Y. (2006). Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems, 23*(3), 45-70. doi:10.2753/MIS0742-1222230303

Malhotra, N. K., & Birks, D. (2017). *Marketing research: An applied approach* (5è ed.). Londres: Pearson.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Inter net users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4), 336–355. doi:10.1287/isre.1040.0032

Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing, 81*(1), 36-58. doi:10.1509/jm.15.0497

Martin, K. D., & Murphy, P. E. (2017). The role of data privacy in marketing. *Journal of the Academy of Marketing Science, 45*(2), 135-155. doi:10.1007/s11747-016-0495-4

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20*(3), 709-734. doi:10.5465/AMR.1995.9508080335

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *Journal of Strategic Information Systems, 11*(3), 297-323. doi:10.1016/S0963-8687(02)00020-3

Milne, G. R., & Boza, M. (1999). Trust and concern in consumers' perceptions of marketing information management practices. *Journal of Interactive Marketing, 13*(1), 5-24. doi:10.1002/(SICI)1520-6653(199924)13:1<5::AID-DIR2>3.0.CO;2-9

Montgomery, A. L., & Smith, M. D. (2009). Prospects for personalization on the internet. *Journal of Interactive Marketing, 23*(2), 130-137. doi:10.1016/j.intmar.2009.02.001

Morosan, C., & DeFranco, A. (2015). Disclosing personal information via hotel apps: A privacy calculus perspective. International Journal of Hospitality Management, 47, 120-130. doi:10.1016/j.ijhm.2015.03.008

Mosteller, J., & Poddar, A. (2017). To share and protect: Using regulatory focus theory to examine the privacy paradox of consumers' social media engagement and online privacy protection behaviors. *Journal of Interactive Marketing, 39*, 27-38. doi:10.1016/j.intmar.2017.02.003

Murthi, B. P. S., & Sarkar, S. (2003). The role of the management sciences in research on personalization. *Management Science, 49*(10), 1344-1362. doi:10.1287/mnsc.49.10.1344.17313

Noble, S. M., & Phillips, J. (2004). Relationship hindrance: Why would consumers not want a relationship with a retailer? *Journal of Retailing, 80*(4), 289-303. doi:10.1016/j.jretai.2004.10.005

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *The Journal of Consumer Affairs, 41*(1), 100-126. doi:10.1111/j.1745-6606.2006.00070

Palmatier, R. W., Scheer, L. K., Houston, M. B., Evans, K. R., & Gopalakrishna, S. (2007). Use of relationship marketing programs in building customer–salesperson and customer–firm relationships: Differential influences on financial outcomes. *International Journal of Research in Marketing, 24*(3), 210-223. doi:10.1016/j.ijresmar.2006.12.006

Park, S., Matic, A., Garg, K., & Oliver, N. (2017). *ACM Transactions on the Web (TWEB), 12*(2). doi:10.1145/3143402

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce, 7*(3), 101-134.

Portes, A. (2018). *La transparence numérique: le rôle du client et conséquences sur la relation à la marque.* (Thèse de doctorat). Université de Montpellier, Montpellier.

Rizkallah, J. (2018, 21 février). The role of Marketing in GDPR. *Forbes*. Récupéré de https://www.forbes.com/sites/forbestechcouncil/2018/02/21/the-role-of-marketing-in-gdpr/#2bb11e173ca0

Ryan, 2017. *Are data sharing concerns still holding back true personalization?* Récupéré de *https://www.retailwire.com/discussion/are-data-sharing-concerns-still-holding-back-true-personalization/*

Rygielski, C., Wang, J., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society, 24*(4), 483-502. doi:10.1016/S0160-791X(02)00038-6

Sahni, N. S., Wheeler, S. C., & Chintagunta, P. (2018). Personalization in email marketing: The role of noninformative advertising content. *Marketing Science, 37*(2), 236-258. doi:10.1287/mksc.2017.1066

Salonen, V., & Karjaluoto, H. (2016). Web Personalization: The State of the Art and Future Avenues for Research and Practice. *Telematics and Informatics*, 33 (4), 1088-1104. doi:10.1016/j.tele.2016.03.004

Salesforce Research (2016). *State of the connected customer*. Récupéré de https://www.salesforce.com/form/pdf/state-of-the-connected-customer.jsp

SAS (2015). *Finding the right balance between personalization and privacy.* Récupéré de https://www.sas.com/content/dam/SAS/en_us/doc/research1/balance-between-personalization-privacy-107399.pdf

SAS (2015). *Personnalisation et confidentialité des données. Etude 2015 – Europe de l'Ouest*. Récupéré de https://www.sas.com/content/dam/SAS/es_es/doc/research1/personnalisation-et-confidentialite-des-donnees.pdf

Sheehan, K. B., & Hoy, M. G. (2000). Dimensions of privacy concern among online consumers. *Journal of Public Policy & Marketing, 19*(1), 62-73. doi:10.1509/jppm.19.1.62.16949

Smart Insights (2015). *Are you using all the different types of online personalisation?* Récupéré de https://www.smartinsights.com/ecommerce/web-personalisation/types-ecommerce-personalisation/

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, *20*(2), 167-196. doi:10.2307/249477

SPF Economie (2018). *Le baromètre de la société de l'information 2017*. Récupéré de https://economie.fgov.be/fr/publications/barometre-de-la-societe-de-3

Son, J., & Kim, S. S. (2008). Internet users' information privacy-protective responses: A taxonomy and a nomological model. *MIS Quarterly*, *32*(3), 503-529.

Sundar, S. S., & Marathe, S. S. (2010). Personalization versus customization: The importance of agency, privacy, and power usage. *Human Communication Research*, *36*(3), 298-322. doi:10.1111/j.1468-2958.2010.01377

Sutanto, J., Palme, E., Tan, C., & Phang, C. W. (2013). Addressing the personalization-privacy paradox: An empirical assessment from a field experiment on smartphone users. *MIS Quarterly*, *37*(4), 1141.

Statbel (2017). *Chiffres clés : Aperçu statistique de la Belgique*. Récupéré de https://statbel.fgov.be/sites/default/files/files/documents/FR_kerncijfers_2017_web.pdf

Statbel (2018). *2016-2017 TIC auprès des individus*. Récupéré de https://statbel.fgov.be/fr/themes/menages/utilisation-des-tic-aupres-des-menages#figures

Sterling, G. (2015). *Survey: 99 percent of consumers will share personal info for rewards, but want brands to ask for permission.* Récupéré de https://marketingland.com/survey-99-percent-of-consumers-will-share-personal-info-for-rewards-also-want-brands-to-ask-permission-130786

Sweet, K. (2016). *Intro to the 5 main types of personalization messages and experiences*. Récupéré de http://www.evergage.com/blog/intro-5-main-types-of-personalization-messages/

Tam, K. Y., & Ho, S. Y. (2005). Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Information Systems Research*, *16*(3), 271-291. doi:10.1287/isre.1050.0058

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research*, *51*(5), 546.

Treiblmaier, H. & Pollach, I. (2007). *Users' perceptions of benefits and costs of personalization*. Paper presented at International Conference on Information System Proceedings, Montreal, Québec. Récupéré de https://pdfs.semanticscholar.org/bb64/490a76717a23cd8fa82a0f68c811f744b8a7.pdf

Veltri, N., Krasnova, H., & Elgarah, W. (2011). *Online disclosure and privacy concerns: A study of Moroccan and American Facebook users*. Americas Conference on Information Systems (AMCIS) 2011 Procedings, Detroit, Etats-Unis. Récupéré de https://pdfs.semanticscholar.org/cc4b/7496a65680c70a070a4ba29d3cfe9235d0ef.pdf

Vesanen, J. (2007). What is personalization? A conceptual framework. *European Journal of Marketing*, *41*(5/6), 409-418. doi:10.1108/03090560710737534

Wander, E. (2018). *Infographic: How much privacy people will give up for personalized experiences.* Récupéré de http://www.adweek.com/digital/infographic-how-much-privacy-people-will-give-up-for-personalized-experiences/

White, T. B., Zahay, D. L., Thorbjørnsen, H., & Shavitt, S. (2008). Getting too personal: Reactance to highly personalized email solicitations. *Marketing Letters*, *19*(1), 39-50. doi:10.1007/s11002-007-9027-9

Wirtz, J., Lwin, M. O., & Williams, J. D. (2007). Causes and consequences of consumer online privacy concern. *International Journal of Service Industry Management*, *18*(4), 326-348. doi:10.1108/09564230710778128

Xu, H., Dinev, T., Smith, H. J. & Hart, P. (2008). *Examining the Formation of Individual's Privacy*. *Concerns*: *Toward an Integrative View. Paper presented at International Conference on Information Systems (ICIS 2008), Paris, France. Récupéré de* https://faculty.ist.psu.edu/xu/papers/conference/icis08a.pdf

Xu, H., Luo, X., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems, 51*(1), 42-52. doi:10.1016/j.dss.2010.11.017

Xu, H., Parks, R., Chu, C.H., and Zhang, X.L. (2010). *Information disclosure and online social networks: From the case of Facebook news feed controversy to a theoretical understanding*. Americas Conference on Information Systems Proceedings *(AMCIS 2010)*. Récupéré de https://pdfs.semanticscholar.org/20dd/d61394c9c5d4be32840c8824423cd191e8f8.pdf

Xu, H., Teo, H., Tan, B. Y., & Agarwal, R. (2009). The Role of Push--Pull Technology in Privacy Calculus: The Case of Location-Based Services. *Journal Of Management Information Systems*, *26*(3), 135-173.

Zhang, B., Wang, N., & Jin, H. (2014). *Privacy Concerns in Online Recommender Systems: Influences of* Control and User Data Input. Presented at 10th Symposium on Usable Privacy and Security (SOUPS), 159-173. Récupéré de *https://www.usenix.org/system/files/conference/soups2014/soups14-paper-zhang.pdf*

Zhao, L., Lu, Y., & Gupta, S. (2012). Disclosure intention of location-related information in location-based social network services. *International Journal of Electronic Commerce, 16*(4), 53-90. doi:10.2753/JEC1086-4415160403