

Patrick Nicholson
Professeur des Universités
Université Lille Nord de France
MERCUR - SKEMA
Avenue Willy Brandt
59777 Lille
Tél : 06 84 24 52 75
patrick.nicholson@skema.edu

Pauline de Pechpeyrou
Maître de conférences
Université Paris Ouest Nanterre La Défense
CEROS
200 avenue de la République
92001 Nanterre Cedex France
Tél : 06 24 42 46 71
pauline.depechpeyrou@u-paris10.fr

Online Privacy Notices: What Do They Tell Us About Their Intentions?

Des chartes de respect de la vie privée, pavés ou pavées de bonnes intentions.

Pour qui ? Pour dire quoi ?

Abstract:

In the e-commerce context, relationship marketing implies some level of disclosure of customer's privacy. Beyond legal constraints one should expect transparency and comprehensiveness from privacy notices as to reduce perceived level of risk. A lexical analysis of the privacy notices of the first 15 websites in terms of traffic in France points out a great heterogeneity. If privacy notices aim to greater consumers' understanding and trust it appears they should be amended so to actually help consumers make their choice across companies and help them to control their level of disclosure.

Résumé :

Toute démarche de marketing relationnel suppose un niveau plus ou moins élevé de dévoilement de la part du client d'un certain nombre de données personnelles. En dehors de considérations juridiques, la clarté avec laquelle s'expriment les entreprises, au travers de leur charte de respect de la vie privée, semble essentielle à l'établissement d'une relation de confiance. L'étude présentée dans cet article met en évidence le fait que la plus grande partie des chartes présentées par les 15 principaux acteurs du commerce en ligne ne correspond pas à l'intention de transparence que peuvent attendre les clients ou prospects. Une plus grande homogénéité du discours de ces acteurs semble faire partie des conditions nécessaires au développement à long terme du commerce en ligne et du marketing relationnel.

Des chartes de respect de la vie privée, pavés ou pavées de bonnes intentions.

Pour qui ? Pour dire quoi ?

Le rythme de croissance des échanges en vente à distance et plus particulièrement celui du commerce électronique ne se dément pas. Au 1^{er} semestre 2011, le chiffre d'affaires de l'ensemble des sites de ventes en ligne a progressé de 20% par rapport au même semestre de 2010 pour atteindre 17,5 milliards d'euros (FEVAD¹). Historiquement, le développement de cette activité passe par la construction de relations fortes avec les clients et la mise en œuvre d'opérations commerciales ciblées en « *up-* et *cross-selling* », permettant de fidéliser la clientèle (Peppers *et al.*, 1999).

Si les techniques de segmentation et de personnalisation sont intimement liées à cette démarche, l'évolution des outils informatiques modifie l'échelle à laquelle la collecte et l'exploitation de données personnelles sont désormais possibles (Dolnicar et Jordaan, 2007). Ainsi, des outils très élaborés de « web analytics » et de « data-mining » permettent d'observer la navigation des internautes et de l'exploiter à des fins de ciblage. Les produits consultés et abandonnés peuvent par exemple faire l'objet de relances via des publicités sur d'autres sites (« retargeting ») ou d'offres promotionnelles via e-mail (« remessaging »). Selon une étude américaine, la relance par e-mail d'un panier abandonné est considérée comme un service rendu pour 49% des clients et un moyen de concrétiser leur achat (Freedman, 2008). Il apparaît que 77% des personnes interrogées dans cette enquête déclarent acheter, au moins de temps en temps, des produits additionnels à leur achat principal, grâce aux recommandations personnalisées effectuées par les sites marchands.

Bien que ces techniques de « profiling » puissent être créatrices de valeur pour les clients en ligne, elles s'appuient sur des méthodes souvent intrusives de collecte d'informations personnelles (cookies, tags...). On peut y voir la source d'un niveau de plus en plus élevé de préoccupation des internautes à cet égard. Une enquête TNS Sofres menée pour Microsoft (mai 2010) sur la vie privée des Français révèle que 75% des répondants se déclarent inquiets du fait que les informations qu'ils peuvent laisser sur Internet puissent être vues et utilisées à leur insu. Par ailleurs, cette inquiétude est amplifiée par un sentiment de manque de contrôle ; 71% des répondants estiment qu'il leur est difficile de maîtriser les informations rendues accessibles et se montrent de plus en plus méfiants à l'égard des entreprises de commerce en ligne.

¹ <http://www.fevad.com/espace-presse/bilan-e-commerce-au-1er-semester-2011#topContent>

La confiance de l'internaute vis-à-vis des sites marchands qu'il fréquente est un élément-clé de sa propension à se dévoiler et à accepter les offres personnalisées du site (Eastlick *et al.*, 2006). Le niveau de confiance des internautes dépend de nombreux facteurs. Chouk et Perrien (2004) proposent une typologie des leviers de la confiance vis-à-vis d'un site marchand autour de quatre groupes de facteurs : les facteurs liés au site, les facteurs liés au marchand, les facteurs liés aux tiers et enfin les facteurs liés au client. La recherche présentée dans cet article s'intéresse aux facteurs liés au site, au travers de l'examen approfondi des chartes de respect de la vie privée des 15 principaux sites marchands français en termes d'audience (Médiamétrie//NetRatings / mai 2011).

Dans un contexte européen de forte législation du droit de l'internaute, force est de s'interroger sur la finalité réelle de ces chartes. Servent-elles à protéger le site marchand contre les réclamations potentielles des internautes ou à renseigner de manière claire et didactique le visiteur animé d'une intention d'achat mais soucieux d'en connaître les conditions ?

Deux approches semblent pouvoir répondre de façon complémentaire à cette question. La première consiste à faire une analyse descriptive de ces chartes et de leur structure. Cette analyse se fonde sur les travaux relatifs à la complétude des chartes (Culnan, 1999 ; Sheehan, 2005 ; Gauzente, 2008) ainsi que sur une analyse lexicale menée sur les textes collectés. Afin de compléter cette analyse descriptive et d'aller au-delà d'approches vérifiant la présence ou l'absence de certaines rubriques, nous nous interrogerons sur les intentions de leurs énonciateurs (Gavard-Perret et Moscarola, 1998) en prenant appui sur les travaux menés par Pollach (2007) sur les chartes anglo-saxonnes.

L'analyse du contenu des chartes des principaux sites marchands français répondra aux objectifs de recherche suivants :

- Repérer dans quelle mesure les sites marchands communiquent de manière claire, univoque et exhaustive sur leur politique en matière de recueil et d'utilisation de données personnelles.
- Mesurer le niveau de complétude de ces chartes.
- Révéler par l'analyse des énonciations le niveau de transparence réelle et les intentions d'information des énonciateurs.

Les perspectives ouvertes par ce travail sont nombreuses, tant vis-à-vis des contraintes législatives et du devoir d'information pesant sur les commerçants en ligne, que de l'attitude et des comportements de leurs clients face à cette problématique.

LE RESPECT DE LA VIE PRIVÉE

La politique de respect de la vie privée d'une entreprise (*privacy policy*) ou de respect des données personnelles présentée dans sa charte est définie comme « *l'ensemble des actions mises en œuvre afin d'informer les internautes / clients des mesures prises pour traiter et protéger les données communiquées volontairement ou non par les internautes lors de leur visite sur le site* » (Gauzente, 2008). Cette politique fait appel tant à des facteurs liés au site (rédaction et mise en œuvre d'une charte) qu'à des facteurs liés aux tiers dont les labels sont utilisés comme garants de la confiance qu'on peut leur accorder (L@belsite, Fianet, FEVAD, etc.).

Il apparaît que dans les pays européens, plus réglementés, les entreprises mettent davantage en avant leur charte que leurs labels et se montrent moins exigeantes en demande d'informations aux consommateurs (Gurau *et al.*, 2003). Par ailleurs, l'étude de Palmer *et al.* (2000) montre que lorsque le site possède déjà un fort capital de marque, il a tendance à promouvoir son capital-marque en réduisant la prééminence des labels (donc en augmentant le nombre de clics nécessaires pour les atteindre) et en mettant en avant sa propre charte de respect de la vie privée (en réduisant le nombre de clics).

Plusieurs recherches montrent l'intérêt de renouveler régulièrement l'étude des réducteurs du risque perçu associé à la collecte et à l'usage des données personnelles. Tout d'abord, le niveau de préoccupation des consommateurs évolue au fur et à mesure que les technologies s'affinent : selon le baromètre Harris-Equifax, la part des consommateurs « non préoccupés » est passée de 22% contre 10% entre 1999 et 2003. De même, en France, la proportion des consommateurs préoccupés par la collecte de données personnelles a augmenté sensiblement (baromètre ETO/Market Audit, 2010). Ensuite, les entreprises modifient leur politique du respect de la vie privée sous l'impulsion de nouvelles législations (Milne *et al.*, 2006 ; Gauzente, 2008). Ainsi, aux Etats-Unis, les entreprises de commerce en ligne ont revu leur politique après les deux lois promulguées en 2002 sur les chartes de respect de la vie privée. Elles en ont accru la longueur mais ne semblent pas en avoir amélioré la lisibilité, bien au contraire. En France, Gauzente (2008) observe également que les sites marchands se dotent de chartes plus complètes, même si leur structure semble s'être stabilisée.

REVUE DE LITTERATURE

Une revue de la littérature consacrée au respect de la vie privée fait apparaître trois principaux axes de recherche :

- La confiance comme facteur-clé de succès du marketing relationnel
- Les préoccupations des consommateurs à l'heure des nouvelles technologies
- Les chartes et leurs objectifs

La confiance comme facteur-clé de succès du marketing relationnel

Le marketing relationnel a été défini comme « *l'ensemble des activités marketing destinées à établir, développer et maintenir des relations réussies avec les clients* » (Morgan et Hunt, 1994). C'est le cas des opérations de *cross-* et de *up-selling* qui permettent à l'entreprise de développer son chiffre d'affaires et aux clients fidèles de recevoir des offres ciblées, soutenues potentiellement par des actions promotionnelles. La coopération entre partenaires les aide à atteindre leurs objectifs communs. Elle conduit à créer plus de valeur que ne pourrait le faire chacun des partenaires séparément (Palmatier *et al.*, 2006).

Toutefois, un ciblage ou une personnalisation réussie requiert que le consommateur ait suffisamment confiance en l'entreprise pour accepter de se dévoiler. La confiance est définie comme la « *croissance dans la fiabilité et l'intégrité du partenaire de l'échange* » (Morgan et Hunt 1994, p. 23). La confiance vis-à-vis d'un site marchand a été structurée autour de quatre dimensions (Chouk et Perrien, 2005) : compétence et expertise du marchand (dimension crédibilité), respect de ses engagements (dimension intégrité), motivation d'aider le consommateur en cas de problème (dimension orientation à résoudre les problèmes) et attention portée à ses intérêts (dimension bienveillance).

Les préoccupations des consommateurs à l'heure des nouvelles technologies

Les nouvelles technologies de l'information et de la communication, si elles permettent une plus grande interactivité et un meilleur ciblage des messages et des offres, suscitent un sentiment croissant d'intrusion chez les consommateurs. 41% des consommateurs français (contre 35% en 2007) considèrent ainsi que les marques/enseignes se situent à un niveau élevé d'intrusion dans leur vie privée (Baromètre de l'intrusion 2010, ETO/Market Audit).

La notion de vie privée est le plus souvent appréhendée sous la forme d'un construit multi-dimensionnel (Burgoon *et al.*, 1989 ; DeCew, 1997). On peut la considérer selon trois angles principaux :

- le respect informationnel, défini comme « *la revendication des individus à déterminer eux-mêmes quand, comment et dans quelle mesure des informations les concernant seront communiquées à d'autres* » (Westin, 1967),
- le respect physique, qui est le degré auquel une personne est physiquement accessible à autrui,
- le respect de l'expression, qui protège l'individu des formes d'expression de son identité au travers de son discours et de son activité.

De ces approches, c'est l'approche informationnelle proposée par Westin qui prédomine dans les travaux consacrés au respect de la vie privée sur Internet (Buchanan *et al.*, 2007 ; Lancelot-Miltgen et Gauzente, 2006). Cette approche a été fortement relayée par les travaux de Hoffman *et al.* (1999) qui ont été les premiers chercheurs, dans le domaine du e-commerce et e-marketing, à souligner l'importance de la notion de contrôle sur les données personnelles.

Le sentiment de perte de contrôle quant à la collecte et à l'exploitation des données personnelles est croissant chez les internautes français. Lors de l'édition 2010 du Baromètre de l'intrusion (ETO/Market Audit), 76% des répondants se disent dérangés par le fait que des informations de tous types concernant chaque Français soient enregistrées dans de nombreux fichiers. De plus, 70% des répondants considèrent qu'ils ne sont pas suffisamment informés de leurs droits à la protection des données personnelles (contre 61% en 2008). De nombreux auteurs s'accordent ainsi sur le fait que pour réduire ces craintes et réussir dans leur démarche de marketing relationnel, les entreprises doivent chercher en priorité à rassurer leurs prospects et clients, notamment en ce qui concerne la collecte et l'exploitation des données personnelles (Eastlick *et al.*, 2006 ; Taylor *et al.*, 2009 ; Joinson *et al.*, 2010).

Les chartes et leurs objectifs

Puisque la manière dont une entreprise communique autour de ses pratiques a un impact sur la confiance et la perception du niveau de risque encouru par ses clients, il est important de comprendre pourquoi les chartes ne répondent pas forcément au besoin de réassurance des internautes quant à la collecte et à l'analyse de leurs données personnelles (Pollach, 2007). Nous détaillerons d'abord les rubriques potentiellement présentes dans ces chartes, puis leur impact sur le consommateur, et enfin les objectifs qu'elles poursuivent réellement.

Les différentes rubriques des chartes

Selon le modèle proposé par Smith *et al.* (1996) et testé empiriquement par Stewart et Segars (2002), la préoccupation vis-à-vis du respect informationnel des données personnelles recouvre quatre dimensions : la collecte de données personnelles (*Collection*), l'accès non autorisé (*Improper Access*), les erreurs (*Errors*) et les usages secondaires non autorisés (*Unauthorized Secondary Use*). Ces dimensions ont fait l'objet du développement d'une échelle de mesure (*Concern for Information Privacy Instrument*, Smith *et al.*, 1996).

Les travaux de Culnan (1999) distinguent cinq rubriques dans l'analyse d'une charte de respect de la vie privée :

- La notification qui informe le client des types de données collectées et de leur usage.
- Le choix qui informe de la communication éventuelle des données à des tiers.
- L'accès qui informe du droit à la modification des données collectées par le client.
- La sécurité qui informe le client des conditions de stockage et de transmission.
- Enfin, le contact qui informe le client des possibilités de contact sur les données collectées.

En France, la FEVAD rappelle à ses adhérents les obligations légales en matière de protection de la vie privée. Tout traitement de données à caractère personnel fait l'objet d'un droit d'accès, de rectification, d'opposition et d'autorisation à la cession pour communications électroniques. Lorsque la stratégie d'opt-in est retenue, le consentement est défini comme « *une manifestation de volonté libre, spécifique et éclairée* ».

Si la plupart des entreprises possèdent sur leur site une charte de respect de la vie privée, très peu d'entre elles fournissent les informations relatives à chacun de ces champs (Pollach, 2007 ; Li et Zhang, 2009 ; Li, 2009). L'examen des pratiques des 100 entreprises du classement Fortune révèle que 94% d'entre elles ont effectivement rédigé une charte de respect des données personnelles mais que la plupart ne suivent que partiellement chacun des quatre principes (notification, choix, accès et sécurité) (Li et Zhang, 2009 ; Li, 2009). Ainsi, seuls 19% des entreprises mentionnent qu'elles ont mené une démarche afin de sécuriser les informations lors de leur transmission et de leur stockage. De manière similaire, Pollach (2007) identifie les sites n'apportant pas de réponse satisfaisante aux questions que peuvent se poser les internautes autour de cinq rubriques : la collecte de données, la collecte de données

auprès de tiers, le stockage des données, le partage des données avec des tiers, la communication marketing. C'est essentiellement le stockage des données qui apparaît comme élément critique, aussi bien dans l'échantillon global (75% de sites n'apportant pas de réponse satisfaisante) que dans l'échantillon de sites jouissant d'un label de qualité (68,4%) (Pollach, 2007). Une étude menée par le cabinet Ernst & Young en 2011 révèle que si 59% des firmes déclarent que leur priorité de financement sur l'année à venir sera d'augmenter leur budget pour renforcer la sécurité sur le *cloud computing* (stockage à distance des informations), elles ne sont que 51% à disposer d'une stratégie documentée à cet effet (L'Atelier BNP-Paribas, 31 octobre 2011).

L'impact des chartes sur les croyances et comportements des internautes

Qui a déjà lu, ne serait-ce qu'un tiers, de la documentation de Word ou d'Excel ? L'une des premières questions que l'on peut se poser vis-à-vis des chartes concerne logiquement la propension des consommateurs à les lire ! Les études de Milne et Culnan (2001, 2004) et de Jensen *et al.* (2005) montrent qu'en réalité les consommateurs ne lisent que rarement les chartes et s'arrêtent généralement à la visualisation d'un label. Ainsi, sur un échantillon de 2 468 internautes, seuls 31% d'entre eux les lisent parfois, 13% souvent et 5% toujours.

Toutefois, ces pourcentages augmentent en condition de risque perçu élevé (informations personnelles réclamées, concrétisation du paiement, etc.). La lecture des chartes apparaît être positivement liée à la préoccupation du consommateur vis-à-vis du respect de la vie privée, au degré perçu de clarté de la charte et à la confiance accordée à la charte. Le caractère compréhensif de la charte influence d'ailleurs positivement à la fois l'envie de la lire et la confiance que l'internaute lui accorde.

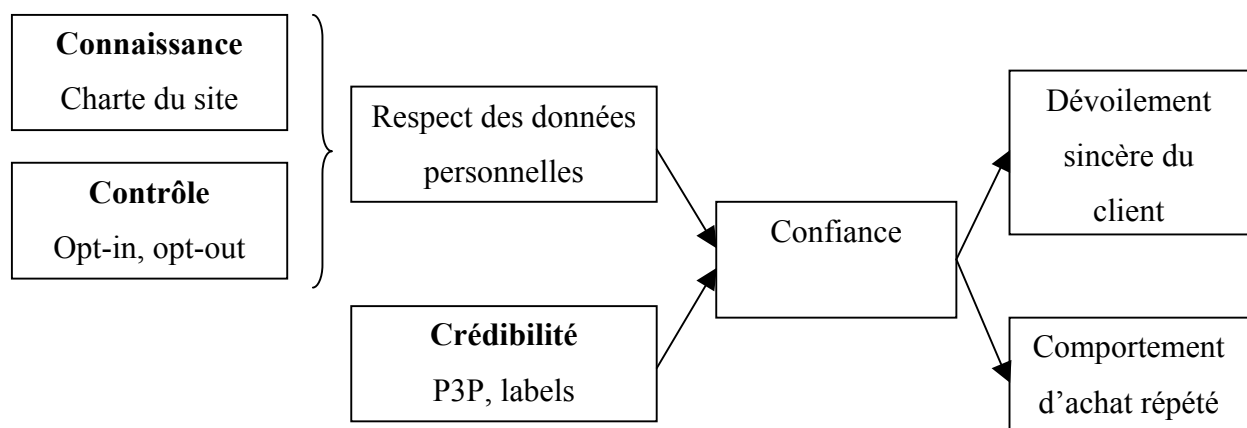
Les raisons de non-lecture invoquées par les répondants sont l'ennui associé à la lecture, la longueur des chartes, les caractères d'impression trop petits, leur similarité, mais surtout leur difficile compréhension du fait d'un vocabulaire juridique peu accessible (Culnan et Milne, 2001). Les verbatims sont d'ailleurs éloquentes : « *Revenez sur terre les gars, aucune personne sensée ne veut lire ce bordel* », « *J'aimerais qu'ils simplifient pour aller aux points essentiels et que la version longue et détaillée ne soit présente qu'en option, au cas où un idiot à la recherche d'argent facile essaye de les poursuivre* », « *Je n'ai pas une licence de droit* », « *Dites-moi juste ce que j'ai besoin de savoir sans tout le surplus légal* », « *Un vrai charabia*

et la nécessité d'avoir recours à un avocat de Philadelphie pour décrypter l'absurdité du jargon.... »².

Le premier impact des chartes concerne l'équité perçue du vendeur (Mollick et Mykytyn Jr, 2009). La demande de consentement, la limitation du partage des données personnelles et des usages secondaires renforcent la perception d'équité par le client. Dans ces conditions, celui-ci a une propension plus élevée à se dévoiler de manière sincère (Lauer et Deng, 2007 ; Hui *et al.*, 2007). La propension à communiquer des données personnelles est d'ailleurs davantage influencée par la présence d'une charte que par celle d'un label (Hui *et al.*, 2007).

Les chartes influencent les comportements des internautes par le biais d'une confiance accrue vis-à-vis du site marchand (Bart *et al.*, 2005 ; Lauer et Deng, 2007 ; Pollach, 2007 - Figure 1). Elles sont d'autant plus importantes aux yeux des internautes qu'ils perçoivent une situation de risque élevé sur la nature de l'information délivrée (Culnan et Milne, 2001), ce qui est le cas par exemple lors d'un achat réalisé sur un site de voyages (Bart *et al.*, 2005).

Figure 1. L'effet des chartes sur les réponses des internautes (adapté de Lauer et Deng, 2007 ; Pollach, 2007)



Selon l'étude menée par Li et Zhang (2009), la confiance et la consultation des chartes avant l'achat renforcent la propension d'achat en ligne via une diminution de la préoccupation vis-à-

² Verbatims originaux recueillis par Culnan et Milne (2001) :

“Get real, no one in their right mind WANTS to read that mumbo jumbo verbiage”

“...I wish they would just simplify the basic points and have a long wordy version as an option and to cover their bums in the case of some stupid money grubber suing them.”

“I don't have a law degree.”

“Just tell me what I need to know without the legal mumbo jumbo.”

“The gobbledegook and having to have a Philadelphia lawyer to go through all the ‘whereas’ nonsense. Too much covering of the proverbial butt.”

vis du respect de la vie privée. Paradoxalement, la vérification et la lecture des chartes sont liées à des niveaux préalables élevés de préoccupation, eux-mêmes liés à des niveaux élevés d'achat en ligne. Toutefois, le sens de la causalité entre lecture et préoccupation n'est pas clairement établi puisque Milne et Culnan (2004) démontrent, pour leur part, que la lecture d'une charte est la conséquence de la préoccupation du consommateur vis-à-vis du respect de sa vie privée. Pan et Zinkhan (2006) montrent qu'en accord avec la théorie du contrat social, un site web avec une politique claire de respect de la vie privée signale aux visiteurs qu'ils peuvent « avoir confiance ».

Les objectifs poursuivis par les chartes

La seconde question que l'on est en droit de se poser est de savoir si les chartes sont rédigées pour répondre aux contraintes du législateur ou aux craintes des consommateurs. Les acheteurs en ligne considèrent que les messages de respect de la vie privée sont plus faciles à lire et à comprendre lorsqu'ils sont courts et directs (Culnan et Milne, 2001 ; Pan et Zinkhan, 2006). L'expérimentation menée par Pan et Zinkhan (2006) permet de comparer l'efficacité de deux types de chartes : l'un exhaustif et légal, l'autre plus court et direct. En accord avec la théorie de l'information, les consommateurs sont d'autant plus attentifs à un message que celui-ci demande peu d'efforts et de temps. Une étude longitudinale de 312 chartes montre qu'au contraire de ce principe les chartes se sont rallongées et ont perdu en clarté (Milne *et al.*, 2006).

Deux raisons peuvent être avancées pour expliquer le manque de clarté des chartes (Milne et Culnan, 2004). Tout d'abord, les chartes, comme de nombreux labels ou avertissements sur un produit, ont souvent été mises en place pour répondre à des exigences juridiques. En conséquence, les organisations rédigent des notices exhaustives mais pas nécessairement accessibles et informatives pour les consommateurs. Ensuite, les chartes sont des documents qui peuvent être rédigés en format hypertexte. Alors qu'il existe donc des opportunités pour promouvoir une meilleure compréhension des notices au travers de liens renvoyant à la rubrique intéressant l'internaute, ces possibilités sont malheureusement sous-exploitées par les sites.

Au-delà de leur longueur et du vocabulaire juridique employé, il est possible d'avoir un regard critique sur les intentions de leurs auteurs. Pollach (2007) distingue quatre stratégies syntaxiques qui lui semblent être exploitées pour créer de l'ambiguïté chez l'internaute et réduire ainsi le niveau de risque perçu :

- L'atténuation (*mitigation*), par le recours à des adverbes de fréquence indiquant que le risque présente une probabilité faible d'occurrence.
- La revalorisation (*enhancement*) qui consiste à mettre en avant les nombreux bénéfices associés à l'exploitation des données personnelles et de l'historique de navigation.
- L'obscurcissement (*obfuscation*), qui, au travers de tournures complexes, rend confus pour l'internaute les modalités de collecte et les dangers liés à l'exploitation.
- L'omission des agents (*omission*) qui masque l'auteur de la collecte et de l'exploitation des données personnelles.

Tableau 1. Analyse des figures de rhétorique des chartes (Pollach, 2007)

Stratégie	Objectif	Illustration
Atténuation	Réduire la fréquence	<i>Occasionnellement, de temps à autre, par moments</i>
Revalorisation	Mettre en avant les qualités	<i>Partage de données : digne de confiance, précautionneux</i>
		<i>Spam : vous intéressant</i>
Obscurcissement	Se couvrir	<i>Peut, peut-être, sauf si, sur une base limitée, nous nous réservons le droit de, incluant mais ne se limitant pas à</i>
	Masquer la causalité	<i>Si vous nous autorisez, seulement si vous nous l'autorisez, avec votre consentement/permission</i>
Omission	Supprimer les agents	<i>Le partage, est partagé (plutôt que « nous partageons ») Vous recevez (plutôt que « nous envoyons »)</i>

Il apparaît que les chartes des entreprises ayant des labels de certification ne sont pas moins ambiguës que celles des entreprises n'en ayant pas, ce qui suggère que la soumission à un label peut avoir un impact sur le contenu, mais pas sur le langage (Pollach, 2007).

METHODOLOGIE DE LA RECHERCHE

Le matériel analysé

Prenant appui sur les travaux conceptuels sur les chartes qui viennent d'être présentés, notre recherche concerne l'analyse des chartes présentes sur les sites des 15 entreprises de e-commerce jouissant de la plus grande audience en France (Médiamétrie//NetRatings/mai 2011). Ce choix ne correspond donc pas à un réel souci de représentativité entendu au sens statistique du terme (on dénombre 81 900 sites marchands actifs en France en 2010) mais relève de l'analyse des cas que l'on peut *a priori* considérer comme les plus probablement représentatifs de bonnes pratiques, ne serait-ce que par leur haut degré de visibilité. Ces 15 sites regroupent par ailleurs des entreprises ayant des modèles économiques assez différents,

allant de « *pure players* » à des entreprises ayant opté pour une stratégie multi-canal, ce qui permet de prendre en compte une grande variété de situations (cf. Annexe 1). Les chartes étudiées ont été collectées durant l'été 2011 et représentent un total de près de 35 pages de texte au format A4 et 8543 mots.

La mesure de la complétude des chartes

Parmi les différentes mesures présentes dans la littérature (Tableau 2) nous avons fait le choix de nous baser sur une analyse de la complétude des chartes. Pour ce faire, nous avons élaboré une grille reprenant 6 dimensions évoquées dans les travaux de Gauzente (2008), Pollach (2007) et Sheehan (2005) (cf. Annexe 2).

Tableau 2. Indicateurs utilisés dans la revue de la littérature.

Concept	Auteur	Opérationnalisation
Intensité de la demande d'information	Gauzente (2008)	<ul style="list-style-type: none"> • Nombre d'informations demandées • Caractère obligatoire ou non de l'indication des informations (champs obligatoires).
Saillance des réducteurs du risque perçu	Gauzente (2008)	Nombre de clics nécessaires pour visualiser les labels et les chartes de respect de la vie privée en partant de la page d'accueil.
Complétude des chartes de respect de la vie privée	Gauzente (2008) Pollach (2007) Sheehan (2005)	Présence de six rubriques spécifiques : <ul style="list-style-type: none"> • La notification • Le choix • L'accès aux données • La sécurité • La durée de conservation des données • Le contact Sheehan (2005) distingue 2 à 4 items pour chacune de ces rubriques (voir Annexe 2).
	Parayitam <i>et al.</i> (2008)	La taille de la charte, allant de : 0 = absence de charte 5 = explication exhaustive de la politique de respect de la vie privée et communication de contacts en la matière
Lisibilité de la charte	Sheehan (2005)	Indice Flesch de facilité de lecture ³ Indice Flesch-Kincaid – niveau d'études requis Indice Gunning-Fog de facilité de lecture
	Milne <i>et al.</i> (2006)	Indice Flesch-Kincaid – niveau d'études requis Indice Dale-Chall Indice Fog Indice Smog

³ L'annexe 3 présente la définition des indices de lisibilité. L'indice de Flesch peut être calculé à l'aide de logiciels tels que TextStat 3.0 (Allorge, 2004) qui permet une approximation du nombre de syllabes par mot.

Cette analyse de complétude a été menée à partir d'une double lecture et codification indépendante menée par les auteurs, puis une confrontation selon la méthodologie utilisée par Pollach (2007). Une première typologie des textes a donc été élaborée sur la base du calcul d'un score de complétude selon les items retenus. Il en résulte une typologie en 3 groupes selon le niveau de complétude des chartes étudiées.

L'analyse de la complétude d'une charte relève d'une lecture approfondie et d'une grande concentration qui ne correspond sans doute pas aux conditions réelles de lecture par un internaute. On peut supposer que celui-ci jugera de ce niveau de complétude par la présence de certains mots ou expressions correspondant à son besoin d'information. Il est néanmoins apparu pertinent de réaliser cette analyse avant d'avoir recours à une analyse lexicale des textes présentés.

L'analyse lexicale des textes

L'analyse lexicale menée sur les 15 chartes comporte deux niveaux : le premier descriptif, le deuxième interprétatif.

Si les textes présentés ont une réelle intention d'information des consommateurs, leur niveau de lisibilité doit être important. On doit à Flesch (1949) le mode de calcul d'un indice permettant d'apprécier la lisibilité d'un texte, prenant en compte la longueur des phrases et le nombre de syllabes par mot. Cet indicateur est le plus fréquemment utilisé dans la pratique. Il a notamment été employé par Cases *et al.* (2005) pour étudier la lisibilité des e-mails commerciaux (à la fois leur objet et leur introduction).

Cette première analyse descriptive a été complétée par quatre indicateurs proposés par Ganassali (1998) : la longueur, la richesse, la répétition et la banalité. Cette vision descriptive des textes permettra peut-être, au cours du temps, d'observer en France la surprenante augmentation de la longueur des textes révélée par la littérature (Milne *et al.*, 2006). Enfin, une analyse lexicale des textes menée sous Sphinx Lexica a permis d'identifier les mots, verbes, adjectif et segments répétés les plus fréquemment utilisés. Dans cette perspective, un tableau lexical croisé, limité aux mots les plus fréquents, a donné lieu à une seconde typologie qui, comparée à la première, a pour objectif de révéler les mots, s'ils existent, les plus révélateurs du niveau de complétude des chartes.

Au-delà d'une telle analyse descriptive, il est possible d'adopter un second niveau de lecture inspiré de la considération suivante : « *Toute communication peut être abordée sous l'angle de son contenu et de ses traits constitutifs (l'énoncé), ou sous l'angle, plus complexe et plus avancé sur la voie de l'interprétation, des adaptations personnelles que chaque individu fait subir, consciemment ou non, à l'outil linguistique (l'énonciation) [...] Dans bien des cas, les énoncés sont identiques ou comparables, alors que l'énonciation diffère de façon importante, accordant plus ou moins de place à l'émetteur ou au récepteur, reliant de manière spécifique l'énoncé à l'énonciateur ou à l'énonciataire* » (Gavard-Perret et Moscarola, 1998).

D'évidence, bien des mots, alors qu'ils font partie du vocabulaire courant prennent un sens différent du sens commun, lorsqu'ils sont remis dans leur contexte. De ce point de vue, il ne s'agit plus de considérer ce qui est dit mais la façon dont les choses sont dites. Cette voie d'analyse correspond à celle empruntée par Pollach (2007) dont nous tenterons une application aux cas des chartes étudiées. Notre objectif dépassera alors la simple analyse des champs juridiques couverts par les chartes ainsi que leur niveau de compréhension pour porter un regard critique sur leurs intentions.

RESULTATS OBTENUS

Analyse lexicale descriptive

Avant d'entrer dans le détail de leur analyse, les textes rassemblés diffèrent sensiblement les uns des autres par leur taille tout comme par leur accessibilité (Cf. annexe 1). Il est par ailleurs frappant que très peu de sites indiquent une date de mise à jour des informations qu'ils délivrent. Confirmant la teneur des verbatims collectés par Culnan et Milne (2001) lors de leur étude qualitative, l'ensemble des textes analysés présente une faible facilité de lecture (voir modes de calcul en Annexes 4 et 5). Les valeurs de l'indice de lisibilité de Flesch permettent de tous les situer au niveau le plus élevé correspondant à celui de revues scientifiques ou académiques (Tableau 3).

Tableau 3 : Indices de lisibilité de Flesch des 15 chartes étudiées

SITES / VARIABLES	EBAY	AMAZON	CDISCOUNT	PRICE MINISTER	FNAC	GROUPON	VOYAGES-SNCF	LA REDOUTE	VENTE PRIVEE	PIXMANIA	CARREFOUR	3SUISSES	RUE DU COMMERCE	BRANDALLEY	KIABI
Nombre de paragraphes	90	94	9	46	103	12	44	23	10	16	93	12	43	11	11
Nombre de mots	3218	3921	177	485	2105	349	1416	645	550	688	1926	672	468	379	337
Nombre de phrases	166	220	7	34	158	17	90	27	35	29	140	25	70	19	14
Nombre de mots par phrase	19,3	17,8	25,2	14,2	13,3	20,5	15,7	23,8	15,7	23,7	13,7	26,8	6,6	19,9	24
Nombre de syllabes par mot (approx.)	1,97	2,06	1,99	2,08	1,99	2,16	2,08	2,00	1,93	2,06	1,96	1,90	1,97	2,17	2,05
Indice de Flesch	20,4	14,6	12,9	16,1	24,7	3,22	14,9	13,1	27,8	8,1	27,2	18,4	33	3,3	8,6

Les résultats issus de l'analyse du niveau de complétude des chartes (cf. annexe 2) permettent de constater que seule celle d'eBay recouvre de façon exhaustive les 12 indicateurs utilisés pour le mesurer. Si elles font très majoritairement mention des conditions dans lesquelles les données collectées peuvent être transmises à des tiers ou fournissent des explications sur l'utilisation des cookies, elles n'abordent que plus rarement les conditions dans lesquelles un client peut supprimer les données le concernant, ni leurs conditions de stockage. Le niveau d'information délivré sur la sécurité des données peut d'ailleurs paraître inquiétant au regard d'événements récents témoignant de failles dans les dispositifs de sécurité informatique d'acteurs majeurs du marché. Enfin, il n'existe pas de relation linéaire évidente entre le niveau de lisibilité des chartes et leur niveau de complétude ($r=0,45$), le texte le plus accessible selon l'indice de Flesch (Charte du site « Rue du commerce ») faisant partie de ceux couvrant le plus mal les indicateurs de complétude.

Les 12 items de complétude ont été utilisés pour faire apparaître, par le biais du calcul d'un score global, trois niveaux de complétude, du plus élevé (Groupe 1) au plus faible (Groupe 3).

Groupe 1 : Ebay - Carrefour - Amazon - Fnac - Voyages-sncf.com

Groupe 2 : La Redoute - 3 Suisses - Priceminister - Vente-privee

Groupe 3 : Kiabi - Brandalley - Rue du Commerce - Pixmania - Cdiscount- Groupon

La longueur des textes présentés semble favoriser la couverture de l'ensemble des thèmes qu'un consommateur est en droit d'attendre d'une chartre concernant le respect de sa vie privée ($F=22,666$; $p=0,01$). Bien que ce résultat ne soit pas surprenant, il manifeste la difficulté à rendre compatibles deux aspects contradictoires des chartes : leur complétude et leur longueur, cette dernière ne pouvant qu'avoir des effets négatifs sur la propension des internautes à les lire (Milne et Culnan, 2004).

De même, bien que la liaison ne soit pas statistiquement significative, la lisibilité mesurée selon l'indice de Flesch évolue dans le sens contraire à la complétude. Plus une charte est longue, plus elle a de chances de couvrir correctement le champ des dispositions législatives mais plus elle risque de contenir des mots complexes.

En suivant la méthodologie proposée par Ganassali (1998) et pour donner à l'analyse réalisée un caractère synthétique au regard du nombre potentiellement très élevé des indicateurs pouvant servir à analyser un texte, nous ne considérerons que 4 indicateurs pour tenter de faire correspondre la typologie relative à la complétude des chartes étudiées et leur contenu lexical. Ces 4 facteurs sont la longueur, la richesse, la répétition et la banalité⁴. Longueur et richesse sont logiquement corrélées et il n'est pas surprenant d'observer un lien significatif entre le niveau de complétude et la richesse du vocabulaire utilisé. En ce qui concerne la banalité, bien qu'elle soit plus grande lorsque les niveaux de complétude sont faibles, la relation entre ces deux variables n'est statistiquement pas significative. Par contre la répétition permet de discriminer les niveaux de complétude. Les chartes qui correspondent au plus faible niveau de complétude ont un niveau de répétition égal à celles qui ont les meilleurs scores sur cette variable. La sobriété du discours n'est donc pas révélatrice d'une réelle volonté d'éclairer le lecteur. Ce sont les chartes d'un niveau moyen de complétude qui présentent le plus faible niveau de répétition et qui semblent donc présenter le meilleur compromis entre longueur, richesse et répétition.

Il n'en reste pas moins qu'aucune charte ne possède à la fois les qualités de concision et de complétude, ce qui laisse penser que tant les sociétés concernées que les organismes professionnels qui leur proposent des guides de bonne conduite ont encore du chemin à parcourir pour rendre leurs chartes accessibles, crédibles et utiles vis-à-vis des attentes de leurs clients et prospects. La grande hétérogénéité des indicateurs utilisés confirme quant à

⁴ Le mode de calcul de ces indicateurs fournis par Sphinx Lexica est détaillé dans Ganassali S. (1998) : Les mots qui touchent : le marketing textuel appliqué à la chanson française, *Actes des Journées d'Analyse de Données Textuelles*, Nice (<http://lexicometrica.univ-patris3.fr/jadt1998/ganasall.htm>)

elle que le niveau de complétude ne permet pas, à lui seul, de porter un jugement sur la qualité ou la valeur pour les clients des chartes étudiées.

Ces résultats conduisent assez naturellement à tenter d'identifier des formes de discours sur la base du vocabulaire employé sur les différents sites étudiés. En posant comme filtre à la sélection du vocabulaire utilisé le fait de voir un mot utilisé au moins 8 fois (soit a priori au moins par la moitié des sites étudiés) et qui soit présent au moins dans 3 chartes (ce qui évite les mots trop exclusifs), le corpus étudié se réduit à 189 mots (en dehors des « mots outils »). On dispose alors d'un tableau de contingence lexical croisant les mots et les sociétés étudiées. Une typologie en trois groupes basée sur les coordonnées factorielles issues d'une analyse factorielle des correspondances de ce tableau fait apparaître les résultats suivants :

Groupe 1 : Amazon- Ebay

Groupe 2 : Fnac - Rue du Commerce - Kiabi - 3 Suisses - Carrefour - La Redoute

Groupe 3 : Groupon - Venteprivée - Brandalley - Pixmania – Voyages-sncf.com - Priceminister - Cdiscount.

Les registres lexicaux utilisés par ces différents groupes permettent de les qualifier (cf. Annexe 5). Le groupe 1 peut être qualifié « d'orientés services », mettant en avant les avantages du partage d'informations, le groupe 2 « de techniciens orientés marketing » qui utilisent plus fréquemment les mots « cookies », « sécurité » mais aussi « recevoir » et « offres ». Le groupe 3 peut être qualifié de « légalistes » faisant surtout référence à la loi et au droit. En croisant cette typologie avec les indicateurs utilisés précédemment, on s'aperçoit de liaisons significatives entre la longueur et la richesse des textes et le fait d'appartenir à l'un des deux premiers groupes. Ces deux groupes présentent également moins de répétition et moins de banalité. Pour autant, il n'existe pas de lien global significatif entre le niveau de complétude et ces registres lexicaux (en dehors du cas d'Ebay et d'Amazon).

En fait, trois stratégies semblent correspondre aux registres de vocabulaire employés. La plus simple consiste en une stratégie de type « parapluie » en arguant du respect de la législation (groupe 3). Dans cette optique il est possible de rédiger des notices courtes mais ne faisant que très peu acte de sens pédagogique et dont on peut douter de la réelle volonté d'éclairer le lecteur. A l'autre extrême (groupe 1) se trouverait une stratégie consistant à « cacher l'arbre sous la forêt » en mettant en avant les services et les avantages liés à l'utilisation de données personnelles. Ces chartes sont celles qui, sur la forme, semblent avoir été l'objet du plus grand

soin. Par contre leur souci d'exhaustivité les amène à utiliser des termes comme « illégal », « frauduleux », « tiers que nous ne contrôlons pas » ! ... qui peuvent augmenter le niveau de risque perçu, ce qui représente une limite à la volonté de transparence dont elles témoignent. Enfin la dernière stratégie (groupe 2) consisterait à mettre en avant des aspects techniques, comme l'emploi de cookies, tout en présentant les potentiels bénéfiques commerciaux de tels dispositifs.

Ces résultats conduisent à aller au-delà de la simple analyse du contenu des discours et d'essayer d'identifier les intentions de leurs auteurs. L'analyse des mots, des adjectifs, des verbes utilisés et surtout du contexte dans lequel ils sont utilisés le plus fréquemment permet de mieux apprécier les intentions exprimées au travers de ces textes.

Analyse lexicale interprétative

La linguistique critique étudie la façon dont les auteurs utilisent la grammaire et le vocabulaire pour construire leur propre version de la réalité, abusant ainsi de leur pouvoir en tant que fournisseur d'informations (Fowler *et al.*, 1979). On retrouve dans cette posture celle adoptée par Gavard-Perret et Moscarola (1998) qui proposent : « *considéré non plus dans son contenu sémantique mais pour sa capacité à traduire des mécanismes intérieurs spécifiques à l'énonciateur ou des intentions dépendantes des conditions de l'énonciation, le mot n'a plus seulement valeur de surface* ».

En prenant appui sur les apports de Pollach (2005, 2007) et son application de la linguistique critique à l'analyse des chartes de respect de la vie privée, il est relativement aisé d'en démontrer toute l'ambiguïté. Une analyse rapide des mots, des verbes et adjectifs utilisés dans la rédaction des chartes est plutôt rassurante : on y trouve un registre de vocabulaire conforme aux dispositions en vigueur et de nature à minimiser le risque perçu par un utilisateur potentiel. Malgré cette apparence superficielle, on peut pourtant retrouver dans les chartes analysées certains mots ou façons d'exprimer les choses qui peuvent laisser perplexe un lecteur attentif. Nous avons repris les quatre modèles linguistiques proposés par Pollach (2007) et les avons illustrés par des segments de phrases identifiés dans les chartes étudiées (Tableau 4).

Tableau 4. Illustration des modèles linguistiques

Modèle	Objectif	Segments identifiés
Atténuation	Réduire la fréquence	pouvons envoyer de temps en temps des offres à certaines catégories de clients éventuellement de les communiquer à des tiers
Revalorisation	Mettre en avant les qualités	nous tenons nos promesses , nous ne changerons jamais de manière substantielle nos pratiques dans un sens qui réduirait la protection des informations clients afin de respecter les obligations légales mettions tout en œuvre pour assurer le respect de votre vie privée
		vous adresser des informations ciblées en fonction de vos centres intérêts ces données sont nécessaires pour vous adresser notamment des newsletters et vous permettre de bénéficier de certains services offerts par ce site nous permettent d'améliorer et personnaliser les services que nous vous proposons et les informations que nous vous adressons sont destinées à mieux Vous connaître et ainsi à améliorer les services qui Vous sont proposés, tout en en assurant la sécurité nos services pour mieux vous satisfaire données personnelles pour assurer la livraison des commandes par ses prestataires
Obscurcissement	Se couvrir	Conformément au droit français , ces informations pourront être dévoilées par XXX à toute personne physique, morale ou entité administrative qui aurait la faculté de l'exiger. XXX peut être amené à utiliser le système de cookies XXX peut être amenée à transmettre aux partenaires des données vous serez donc susceptible de recevoir des offres des sociétés du groupe sauf stipulation contraire dans ce règlement sauf bien entendu autre accord du client il se peut que nous utilisions à cet effet des cookies. XXX n'a aucun accès et ne peut exercer aucun contrôle sur les cookies tiers.
	Masquer la causalité	Nous ne vendons ou ne louons pas vos données personnelles à des tiers à des fins de marketing sans votre consentement formel À l'avenir, sous réserve du consentement que Vous avez fourni dans Votre Compte
Omission	Supprimer les agents	vous rencontrerez peut-être des cookies de tiers sur certaines pages elles peuvent être transmises aux sociétés filiales du groupe les coordonnées de l'acheteur sont communiquées au vendeur pour assurer la bonne livraison du produit vendu

Il conviendrait d'élargir le nombre de modèles linguistiques étudiés, en incluant notamment l'imprécision de certaines expressions et le doute que peut suggérer au lecteur la conjugaison de certains verbes. On serait alors proche d'une forme de circonlocution consistant à « dire de manière indirecte ou imprécise qui témoigne d'une difficulté à dire ». L'utilisation des termes « nécessaires », « susceptibles », « pourront » ou « pourraient être », « éventuellement », « le cas échéant ». Ces mots sont souvent utilisés à proximité des mots « tiers », « partenaires », « du groupe » et concernent tant l'échange et le transfert des données que leur utilisation en dehors du cadre de l'activité propre aux entreprises concernées.

On peut comprendre que les services juridiques puissent chercher à se prémunir des conséquences de situations imprévues. Une analyse plus fine des formes rhétoriques devrait permettre d'identifier les modes d'expression pouvant avoir un effet contraire à la volonté de transparence des entreprises, destinée en première intention à limiter le niveau de risque perçu par leurs clients.

DISCUSSION ET IMPLICATIONS MANAGERIALES

L'analyse textuelle des chartes de respect de la vie privée constitue un moyen d'apprécier, *a priori*, les intentions de leurs rédacteurs. On ne peut être que frappé par leur manque d'homogénéité. Le recours à un vocabulaire juridique et les références aux lois et décrets en vigueur semblent indiquer qu'elles sont le plus souvent élaborées par un service juridique sans réelle concertation avec les services dédiés à la gestion de la relation clients. Un premier entretien auprès d'un expert laisse à penser que cette situation est très courante.

La lecture d'une charte requiert chez l'internaute un effort, du temps et une volonté de prendre conscience de l'importance du message. Le recours à un vocabulaire trop juridique et à des phrases relativement longues et complexes conduit à des taux de lecture relativement faibles, de l'ordre d'un tiers des internautes (Culnan et Milne, 2001 ; Milne et Culnan, 2004 ; Jensen *et al.*, 2005). Le débat qu'il est possible d'initier sur la base de ce travail n'est pas sans rappeler celui entretenu à propos des labels nutritionnels. Comme le soulignent les discussions tenues lors du forum de l'*European Food Information Council* (EUFIC) en 2005 : le « système nutritionnel » est perçu comme complexe et technique alors que l'acte de manger est vu comme simple et immédiat⁵. Plus les messages sont simples et clairs, comme « 1% de matières grasses », plus ils sont assimilés. Les trois barrières identifiées concernant les labels

⁵ <http://www.eufic.org/article/fr/nutrition/etiquetage-aliments-revendications/expid/forum-consumer-attitudes-information-food-labelling/>

nutritionnels : un manque d'autorité, de compréhension et de motivation, semblent transférables à la lecture des chartes de respect de la vie privée.

Les internautes sont également pragmatiques dans leur démarche de dévoilement de données personnelles : ils comparent les risques perçus aux avantages liés à une meilleure connaissance par l'entreprise d'informations les concernant. Le rapport de la CNIL publié en 2009 sur la publicité ciblée considère que la publicité est le « carburant » de l'économie numérique et que cela explique le fait que les acteurs du e-commerce diversifient leurs services et leurs activités, afin de collecter toujours plus d'informations sur le comportement de leurs clients. Le fait de dévoiler des données personnelles afin de bénéficier de services (achat en ligne, promotions ciblées, commodité de navigation) devient une condition indispensable à un certain niveau de confort dans l'achat en ligne. Le « *privacy paradox* » établit ainsi que les individus dévoilent plus que ce qu'ils disent être prêts à dévoiler, par habitude ou dans l'attente de gratifications (Norberg *et al.*, 2007 ; Yap *et al.*, 2009).

Reste posée la question des modalités de contrôle par l'internaute et de la sécurité de ces informations. Dans notre échantillon, moins d'un tiers des sites marchands (eBay, Amazon, Fnac, Voyages-sncf.com, Carrefour) évoque dans leur rubrique « respect de la vie privée » la manière dont les données personnelles sont protégées lors de la transmission, mais surtout lors du stockage. Certains sites communiquent toutefois sur la sécurité lors de la transmission des données financières dans une rubrique séparée intitulée « Sécurité et Internet ». Le plus souvent, seul l'aspect sécurisé du paiement y est évoqué. Toutefois, si les protocoles sécurisés (ex : cryptage SSL) ont contribué à rassurer l'internaute quant à l'achat en ligne, la question de la sécurité de leurs données personnelles après l'achat reste cruciale. A l'heure où de nombreux sites adoptent le *cloud computing*, des scandales récents relancent la question de la sécurité. En avril 2011, Sony confirme sur son blog officiel un vol de données personnelles : les hackers responsables de la panne ont bien eu accès aux données personnelles des joueurs. Fin août 2011, les serveurs de la société « Orange » ont été piratés, rendant disponibles de nombreux fichiers sensibles. Un pirate dénommé HiddenTapz aurait ainsi profité d'une faille de sécurité pour récupérer de nombreuses informations privées commerciales⁶.

Ces actes de cybercriminalité soulèvent deux questions : à court terme, la gestion de crise ; à long terme, la mise en place d'éléments de réassurance de l'internaute en matière de sécurité de ses données personnelles. En reprenant les trois dimensions de la confiance identifiées par Gurviev et Korchia (2002), à savoir la présomption de compétence (maîtrise d'un savoir-

⁶ <http://www.silicon.fr/les-serveurs-dorange-fr-pirates-59240.html>

faire), la présomption d'honnêteté, qui consiste à tenir ce qui a été promis, et enfin la capacité de la marque à prendre en compte l'intérêt des consommateurs, il apparaît que la première, à savoir la compétence du site, apparaît essentielle dans la gestion des données personnelles. Elle repose sur l'attribution à la marque d'un degré d'expertise quant aux attentes fonctionnelles du consommateur. Des recherches futures devraient s'intéresser aux leviers pouvant être mobilisés dans les chartes pour rassurer l'internaute sur la sécurité des données personnelles stockées par le site marchand.

Ce premier travail offre de nombreuses perspectives pour des recherches futures. La première consistera à rencontrer les responsables de la rédaction de ces chartes en espérant que ce travail constitue une clé d'entrée. Il semble peu probable, malgré les préconisations des organismes professionnels, qu'une réflexion ait été menée sur les conditions de production de l'information et l'opportunité d'harmoniser les informations mises à disposition des internautes. Au-delà d'une réflexion sur la production d'informations, se pose le problème de leur accès et de leur utilisation par les consommateurs. Nul doute que les activités commerciales en ligne connaîtront encore de fortes évolutions technologiques. La perception d'un niveau de risque élevé en matière de dévoilement constitue encore un frein majeur à leur développement. Comme le confirment de nombreux travaux de recherche, ce sont justement ceux qui perçoivent ce risque qui lisent les notices ou chartes de respect de la vie privée. Il semblerait donc nécessaire de mieux explorer la façon dont les consommateurs trouvent une solution, au travers des chartes, à leur problème de perception de risque. On a pu constater que le fait de rédiger des chartes courtes favorisait leur taux de lecture (Milne et Culnan, 2004) mais que cette concision est difficilement conciliable avec un souci ou une obligation de complétude. Notre étude des textes accessibles aux internautes ne laisse pas vraiment penser qu'ils sont de nature à diminuer le niveau de risque perçu.

Par ailleurs, d'autres signaux inspirant la confiance peuvent être mobilisés par les consommateurs. Les sites étudiés présentent, là encore, une très grande disparité, notamment dans la façon dont ils mettent en avant les labels dont ils peuvent ou pourraient pourtant se prévaloir.

En ne considérant que les 15 premiers sites en termes d'audience, ce travail présente une limite sérieuse en termes de représentativité. Si ces sites peuvent être considérés comme des références en matière de succès commercial, une perspective naturelle à ce travail consistera à analyser la façon dont des sites moins connus traitent du problème du respect de la vie privée.

Références bibliographiques

- Bart Y., Shankar V., Sultan F. et Urban G.L. (2005), Are the drivers and role of online trust the same for all web sites and consumers? A large-scale exploratory empirical study, *Journal of Marketing*, 69, 4, 133-152.
- Buchanan T., Paine C., Joinson A.N. et Reips U-D. (2007), Development of measures of online privacy concern and protection for use on the Internet, *Journal of the American Society for Information Science and Technology*, 58, 2, 157-165.
- Burgoon J.K., Parrott R., LePoire B.A., Kelley D.L., Walther J.B. et Perry D. (1989), Maintaining and restoring privacy through communication in different types of relationship, *Journal of Social and Personal Relationships*, 6, 131-158.
- Cases A-S., Fournier C. et Dubois P-L. (2005), Lisibilité d'un e-mail : étude qualitative et approche expérimentale in vivo, *Journée nantaise sur le e-marketing*, Nantes
- Chouk I. et Perrien J. (2004), Les facteurs expliquant la confiance du consommateur lors d'un achat sur un site marchand : une étude exploratoire, *Décisions Marketing*, 35, 75-86.
- Chouk I. et Perrien J. (2005), La confiance du consommateur vis-à-vis d'un marchand Internet : proposition d'une échelle de mesure, *Revue Française du Marketing*, 5, 205, 1-15.
- Culnan M.J. (1999), Privacy and the Top 100 Web Sites: Report to the Federal Trade Commission [http://www.securitymanagement.com/archive/library/elec_privacy.pdf]
- Culnan M.J. et Milne G.R. (2001), The Culnan-Milne survey on consumers & online privacy notices: Summary of responses, rapport disponible au lien suivant : <http://www.ftc.gov/bcp/workshops/glb/supporting/culnan-milne.pdf> [dernier accès : 22 août 2011]
- DeCew J. (1997), *In pursuit of privacy: Law, ethics, and the rise of technology*, Ithaca, NY: Cornell University Press.
- Eastlick M.A., Lotz S.L. et Warrington P. (2006), Understanding online B-to-C relationships: An integrated model of privacy concerns, trust, and commitment, *Journal of Business Research*, 59, 8, 877-886.
- Flesch R. (1949), *The Art of Readable Writing*, Harper & Row.
- Fowler R., Hodge B., Kress G. et Trew T. (1979), *Language and Control*, Routledge, Londres.
- Freedman L. (2008), Personalized product recommendations through the merchant and consumer lens, *the e-tailing group – Mybuys*.

- Ganassali S. (1998), Les mots qui touchent : le marketing textuel appliqué à la chanson française, *Actes des Journées d'Analyse de Données Textuelles*, Nice.
- Gauzente C. (2008), Politiques et pratiques des sites marchands en matière de respect de la vie privée : analyse du dosage entre chartes et labels et de leur évolution, *Revue Management et Avenir*, 3, 17, 44-65.
- Gavard-Perret M-L. et Moscarola J. (1998), Enoncé ou énonciation ? Deux objets différents de l'analyse lexicale en marketing, *Recherche et Applications en Marketing*, 13, 2, 31-47.
- Gurau C., Ranchhod A. et Gauzente C. (2003), "To Legislate or not to Legislate": A comparative study of privacy/personalisation factors affecting French, UK and US web sites, *Journal of Consumer Marketing*, 20, 7, 652-664.
- Hoffman D., Novak T. et Peralta M. (1999), Information privacy in the marketplace: implications for the commercial uses of anonymity on the web, *The Information Society*, 15, 2, 129-140.
- Hui K-L., Teo H.H. et Lee S-Y.T. (2007), The value of privacy assurance: An exploratory field experiment, *MIS Quarterly*, 31, 1, 19-33.
- Jensen C., Potts C. et Jensen C. (2005), Privacy practices of Internet users: self-reports versus observed behavior, *International Journal of Human-Computer Studies*, 63, 203-227.
- Joinson A.N., Reips U-D., Buchanan T. et Schofield C.B.P. (2010), Privacy, Trust, and Self-Disclosure Online, *Human-Computer Interaction*, 25, 1, 1-24.
- Lancelot-Miltgen C. et Gauzente C. (2006). Vie privée et partage de données personnelles en ligne : une approche typologique, *Actes du 22^{ème} Congrès International de l'Association Française du Marketing (AFM)*, Nantes.
- Lauer T.W. et Deng X. (2007), Building online trust through privacy practices, *International Journal of Information Security*, 6, 5, 323-331
- Li S. et Zhang C. (2009), An empirical investigation of information privacy concern and its antecedents and consequences, *Northeast Decision Sciences Institute Proceedings*
- Li S. (2009), An analysis of online privacy policies of Fortune 100 companies, in *Online Consumer Protection: Theories of Human Relativism*, eds. Kuanchin Chen et Adam Fadlalla, 269-283 [dernier accès : 22 août 2011].
- Milne G.R. et Culnan M.J. (2004), Strategies for reducing online privacy risks: why consumers read (or don't read) online privacy notices, *Journal of Interactive Marketing*, 18, 3, 15-29.

- Milne G. R., Culnan M. J. et Greene H. (2006), A longitudinal assessment of online privacy notice readability, *Journal of Public Policy and Marketing*, 25, 2, 238-249.
- Mollick J.S. et Mykytyn Jr P.P. (2009), An empirical investigation on the effects of privacy policies on perceived fairness of online vendors, *Journal of Internet Commerce*, 8, 1/2, 88-112.
- Morgan R.M. et Hunt S.D. (1994), The commitment-trust theory of relationship marketing, *Journal of Marketing*, 58, 3, 20-38.
- Norberg P.A., Horne D.R. et Horne D.A. (2007), The privacy paradox: Personal information disclosure intentions versus behaviors, *Journal of Consumer Affairs*, 41, 1, 100-126.
- Palmatier R.W., Dant R.P., Grewal D. et Evans K.R. (2006), Factors influencing the effectiveness of relationship marketing: A meta-analysis, *Journal of Marketing*, 70, 4, 136-153.
- Palmer J.W., Bailey J.P. et Faraj S. (2000), The role of intermediaries in the development of trust on the WWW: The use and prominence of trusted third parties and privacy statements, *Journal of Computer-Mediated Communication*, 5, 3, doi: 10.1111/j.1083-6101.2000.tb00342.x [<http://jcmc.indiana.edu/vol5/issue3/palmer.html>].
- Pan Y. et Zinkhan G.M. (2006), Exploring the impact of online privacy disclosures on consumer trust, *Journal of Retailing*, 82, 4, 331–338.
- Parayitam S., Desai K.J. et Desai M.J. (2008), E-commerce policies, customer privacy and customer confidence, *Vilakshan: The XIMB Journal of Management*, 5, 2, 45-58.
- Peppers D., Rogers M. et Dorf B. (1999), Is your company ready for one-to-one marketing?, *Harvard Business Review*, 77, 1, 151-160.
- Pollach I. (2007), What's wrong with online privacy policies?, *Communications of the ACM*, 50, 9, 103-108
- Pollach I. (2005), A typology of communicative strategies in online privacy policies: Ethics, power and informed consent, *Journal of Business Ethics*, 62, 3, 221-235.
- Smith H.J., Milberg S.J. et Burke S.J. (1996), Information privacy: Measuring individuals' concerns about organizational practices, *MIS Quarterly*, 20, 2, 167-196.
- Stewart K.A. et Segars A.H. (2002), An empirical examination of the concern for information privacy instrument, *Information Systems Research*, 12, 1, 36-49.
- Taylor D.G., Davis D.F. et Jillapalli R. (2009), Privacy concern and online personalization: the moderating effects of information control and compensation, *Electronic Commerce Research*, 9, 3, 203-223.
- Westin A. (1967), *Privacy and Freedom*, Athenaeum, New York.

Annexe 1. Liste et emplacement des chartes étudiées

Rang	Enseigne	Visiteurs uniques moyens par mois en 2010	Rubrique(s)	Sous-rubrique	Labels
1	eBay	10 444 000	Règlement sur le respect de la vie privée		Non
2	Amazon	9 899 000	Vos informations personnelles		Non
			Conditions générales de vente	Article 11 : Protection des données personnelles	
3	Cdiscount	8 671 000	Conditions générales de vente		Fianet Fevad
4	Priceminister	8 469 000	Ethique & Sécurité	Ethique et Vie privée	Société Générale
5	Fnac	7 778 000	Infos légales	Protection des données Deux liens hypertexte vers : <ul style="list-style-type: none"> Voici la procédure à suivre pour que votre navigateur refuse les cookies : ici Liste des Partenaires : cliquer ici 	Non
6	Groupon	7 494 000	Protection des données personnelles		Non
			Conditions Générales de Commercialisation	Article 16 – Informatique et libertés	
7	Voyages-sncf.com	7 211 000	Charte de confidentialité		Non
			Conditions Générales de Vente et d'Utilisation du Site Voyages-sncf.com	Article 14. Protection des données personnelles	
8	La Redoute	6 995 000	Données personnelles		Non
			Conditions Générales de Vente	Article 13 – Informations nominatives	
9	Vente Privée	6 079 000	Conditions Générales de Vente du site www.vente-privee.com (24/06/2011)	Article 15 – Données personnelles	Non
			Devenir membre – Consulter notre politique de confidentialité.	La protection des données personnelles	
10	Pixmania	6 069 000	Conditions générales de vente	10 – Protection des données personnelles Lien vers les conditions d'utilisation	Fianet Paypal
11	Carrefour	5 849 000	Mentions légales		Non
12	3 Suisses	5 372 000	Protection des données		Fevad

			Conditions générales de vente	<ul style="list-style-type: none"> • Informations nominatives et protection des données personnelles • Cookies 	Cadenas
13	Rue du Commerce	4 986 000	Vie Privée		AFAQ ISO 14001 et 9001
14	Brandalley	4 735 000	Conditions générales de vente	Art. 8. Informatique et Libertés	Paypal
15	Kiabi	4 101 000	Conditions de vente	"Informatique et Libertés" Mise en place des "cookies"	Fevad Fianet

Annexe 2. Détail du score de complétude des différentes chartes

		EBAY	AMAZON	CDISCOUNT	PRICE MINISTER	FNAC	GROUPON	VOYAGES-SNCF	LA REDOUTE	VENTE PRIVEE	PIXMANIA	CARREFOUR	3SUISSES	RUE DU COMMERCE	BRANDALLEY	KIABI	TOTAL
Notification	Fournir une notice sur les informations collectées.	X	X		X	X			X			X				X	7
	Fournir une notice sur la manière dont les informations sont utilisées en interne.	X	X	X		X	X	X	X	X		X	X				10
	Fournir une notice sur la manière dont les informations sont divulguées à des tiers.	X	X	X	X	X	X	X	X	X		X	X		X		12
	Fournir une notice sur la présence de cookies.	X	X		X	X		X	X		X	X	X	X	X	X	12
Choix	Proposer un choix sur les communications futures	X	X		X	X		X	X	X	X	X	X			X	11
	Proposer un choix sur la transmission des informations personnelles à des tiers.	X	X	X	X					X		X		X	X		8
Accès	Fournir au consommateur les informations sur ses accès possibles à ses données personnelles	X	X		X	X		X	X	X	X	X	X	X			11
	Permettre aux visiteurs de supprimer des informations personnelles qu'ils ont préalablement données	X															1
Sécurité	Fournir des informations sur les procédures de sécurité mise en œuvre une fois que des données personnelles sont collectées.	X	X					X				X					4
	Fournir des informations sur les procédures de sécurité mise en œuvre pendant la transmission de données personnelles.	X	X			X		X				X					5
Durée	Communiquer la durée de conservation des données	X				X											2
Contact	Fournir l'adresse d'un contact pour s'informer ou modifier leurs données personnelles.	X				X	X	X	X	X		X	X	X	X	X	11
	TOTAL	12	9	3	6	9	3	8	7	6	3	10	6	4	4	4	94

Annexe 3. Définition et interprétation des indices de lisibilité

Milne *et al.* (2006) ont recours à différents indices pour évaluer la lisibilité des chartes de respect de la vie privée :

- Flesch-Kincaid Grade Level = $0.39 \times \text{longueur moyenne des phrases} + 11.8 \times \text{nombre moyen de syllabes par mot} - 15.59$
- Dale-Chall = $0.0496 \times \text{longueur moyenne des phrases} + 0.1579 \times \text{pourcentage de mots difficiles} + 3.635$
- Fog Index = $0.4 \times [(\text{nombre moyen de mots par phrase}) + (N)]$, où N est le nombre de mots de trois syllabes ou plus sur un échantillon de 100 mots.
- Smog Index = Racine carrée de $(N) + 3$, où N est le nombre de mots de trois syllabes ou plus sur un échantillon de 30 phrases.

Le test de lisibilité de Flesch (et son dérivatif le test Flesch-Kincaid) basant son évaluation sur le nombre moyen de syllabes par mot et de mots par phrase, il peut être calculé à l'aide des statistiques communiquées par Word.

Test de lisibilité Flesch

Ce test évalue le texte sur une échelle de 100 points. Plus la note est élevée, plus le document est facile à comprendre. Pour un fichier standard, la note doit être comprise entre 60 et 70.

La formule utilisée pour le résultat du test de lisibilité Flesch est la suivante :

$$206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$$

où :

ASL = longueur de phrase moyenne (nombre de mots divisé par le nombre de phrases)

ASW = nombre moyen de syllabes par mot (nombre de syllabes divisé par le nombre de mots)

Test du niveau de qualité Flesch-Kincaid

Ce test évalue le texte en fonction des niveaux scolaires américains. Ainsi, une note de 8.0 signifie qu'un élève de huitième, ce qui correspond à la deuxième année du collège en France, peut comprendre le document. Une note comprise entre 7.0 et 8.0 est généralement considérée satisfaisante.

La formule du test Niveau de qualité Flesch-Kincaid est la suivante :

$$(.39 \times \text{ASL}) + (11.8 \times \text{ASW}) - 15.59$$

où :

ASL = longueur de phrase moyenne (nombre de mots divisé par le nombre de phrases)

ASW = nombre moyen de syllabes par mot (nombre de syllabes divisé par le nombre de mots)

Annexe 4. Calcul du coefficient de Flesch sous TextStat 3.0 (charte des 3 Suisses)⁷

TextStat 3.0

File Clipboard Language Help

File to stat: C:\Users\Pauline\Desktop\Chartes téléchargées 23 août 2011\3sui

File type: Text file HTML file

Results file type: Text file HTML file

File for results: output.txt

Words separators: Defaults

Sentences separators: Defaults

Ignore capital letters

Ignore accent letters

DOS text (ASCII)

Repetitions counting

List word repetitions on the: last words

List of excluded words from repetition counting:

Defaults

Statistics on file: C:\Users\Pauline\Desktop\Chartes téléchargées 23 août 2011\3suisses23.txt

Word separators: " :;!?,\"/> Defaults

Sentences separators: ".!?"

Number of paragraphs: 5

Number of words: 186

Number of sentences: 7

Number of printable characters (including spaces): 1119

Number of spaces: 175

Number of tabulations: 0

Number of Carriage Return: 5

Number of Line Feed: 5

Number of non-printable characters (others than the above): 0

Number of words per sentence: 26.57

Number of syllables per word (approximate): 1.96

Flesch index: 14.30

Start of list:

CNIL	1
Ces	1
Clientèle	1
En	1
Informatique	1
Les	1
Libertés	1
Maroc	1
Nam	1
Par	1
Pour	1
Relations	1
Service	1
Tunisie	1

File : C:\Users\Pauline\Desktop\Chartes téléchargées 23 août 2011\3suisses23.txt, Size : 1640 bytes

⁷ La difficulté dans le calcul de l'indice de Flesch résidant dans le calcul du nombre de syllabes par mot, une approche alternative a été proposée par Cases *et al.* (2005) fondée sur le nombre de voyelles au sein d'un mot. En effet, une syllabe étant définie comme un sommet syllabique, il suffit en français de compter les syllabes dénombrant les voyelles qui formeront les sommets syllabiques (cf. <http://encyclopedie.snyke.com/articles/syllabe.html>).

Annexe 5 : Liste des mots caractéristiques des 3 groupes de textes étudiés⁸.

CLASSE 1/ 3
EFFECTIF: 72

RK	DISTANCE	IDENT.	RK	DISTANCE	IDENT.
1	0.08578	services	2	0.11637	Tiers
3	0.15062	utiliser	4	0.18379	personnelles
5	0.18533	cas	6	0.21760	automatiquement
7	0.24047	informations	8	0.26626	mail
9	0.27355	exemple	10	0.28343	personnes

CLASSE 2/ 3
EFFECTIF: 63

RK	DISTANCE	IDENT.	RK	DISTANCE	IDENT.
1	0.14360	cookies	2	0.16166	pouvez
3	0.26489	Sécurité	4	0.31105	recevoir
5	0.34087	être	6	0.34390	offres
7	0.34619	souhaitez	8	0.34772	accès
9	0.34872	compte	10	0.35506	site

CLASSE 3/ 3
EFFECTIF: 54

RK	DISTANCE	IDENT.	RK	DISTANCE	IDENT.
1	0.29364	Conformément	2	0.35772	loi
3	0.38623	rectification	4	0.45136	adresse
5	0.46153	Libertés	6	0.48126	Informatique
7	0.50488	partenaires	8	0.50602	droit
9	0.58983	cadre	10	0.60690	notamment

⁸ Traitement réalisé à l'aide du logiciel Coheris-Spad.