



HAL
open science

Analyse de l'univers concurrentiel des sites de vente en ligne : une approche par le Web Analytics

Maria Mercanti-Guérin, Laurent Flores

► To cite this version:

Maria Mercanti-Guérin, Laurent Flores. Analyse de l'univers concurrentiel des sites de vente en ligne : une approche par le Web Analytics. *Vie et Sciences de l'Entreprise*, 2012, 191 - 192 (2), pp.96. hal-02055028

HAL Id: hal-02055028

<https://hal.science/hal-02055028>

Submitted on 2 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Analyse de l'univers concurrentiel des sites de vente en ligne :
une approche par le *Web Analytics***

**Maria Mercanti-Guérin
Maître de conférences
Conservatoire National des Arts et Métiers
Mail : maria.mercanti@gmail.com
Téléphone : 06 32 23 00 62**

**Laurent Florès
Université Paris II & INSEEC
CRMMETRIX**

Analyse de l'univers concurrentiel des sites de vente en ligne : une approche par le Web Analytics

Résumé

Les analyses concurrentielles sont un composant essentiel du marketing stratégique. Néanmoins, de nombreux sites de ventes en ligne échouent à définir leurs concurrents et préfèrent fonctionner à l'intuition. Les analyses récentes conduites grâce aux données logs l'ont été dans une optique de gain de nouveaux consommateurs et d'amélioration des taux de conversion. Très peu d'analyses de la concurrence utilisent le *Web Analytics* qui est un outil d'analyse des comportements des visiteurs sur les sites Web. Dans cet article, nous identifions les principales étapes permettant de mener une analyse concurrentielle en ligne efficace. Le *Web Analytics* enrichit la compréhension du trafic des sites, l'identification des concurrents, les frontières des marchés et apporte un nouvel éclairage à la littérature sur les analyses concurrentielles.

Mots clé : commerce en ligne, analyse concurrentielle, *Web Analytics*, ensemble évoqué

Abstract

Competitor analysis is an essential component of strategic marketing. Instead, many eBusiness Websites fail to define their competitors and tend to operate intuitively. Recently, firms have begun to conduct analysis of amounts of log data in order to gain new customers and improve their conversion rate but very few competitive analysis are conducted using Web Analytic services (Was) which is a process of analyzing the behavior of visitors to a Web Site. In this paper, we describe the main steps of an efficient online competitive analysis. Web Analytics services give rich insight into website traffic, competitors identification, market boundaries and new insights for literature on competitive analysis.

Key words: e-commerce, competitive analysis, *Web Analytics*, evoked set

Analyse de l'univers concurrentiel des sites de vente en ligne : une approche par le *Web Analytics*

Introduction

La mesure d'audience du Web popularisée sous le terme de *Web Analytics* représente une réelle opportunité pour les praticiens du marketing d'analyser les comportements des consommateurs et d'améliorer leurs offres sur Internet. Le *Web Analytics* repose sur l'utilisation de deux technologies. La première technologie se fonde sur les fichiers logs (dits aussi fichiers journaux) qui retracent l'ensemble des requêtes et réponses localisées sur un serveur Web. La deuxième technologie s'appuie sur des marqueurs ou tags implémentés sur chaque page importante d'un site. Ces technologies permettent de recueillir des données qui retracent la navigation des internautes : type de requête effectuée, site Web consulté, type et nombre de pages vues, durée moyenne de visite, nombre de clics...Or, certains analystes pointent du doigt l'énorme quantité de données disponibles au regard du peu d'analyses effectuées pour rendre ces données réellement exploitables. Pour Kaushik (p. 5, 2010), le *Web Analytics* ne peut « induire une optimisation continue de l'interaction en ligne » s'il ne s'accompagne pas d'une analyse de sa concurrence. Pourtant, cette intelligence concurrentielle est peu mobilisée par les acteurs du Web. Notre problématique est de montrer qu'il est possible d'utiliser les apports de la recherche portant sur la définition de la concurrence pour proposer aux praticiens du Web une démarche cohérente d'analyse de ces données. En ce sens, notre article représente une transposition d'un certain nombre de concepts (intensité concurrentielle, substituabilité, segmentation...) à la description et compréhension des mécaniques concurrentielles existant sur le Web. L'importance de notre question de recherche a trait à la structure même de la concurrence des sites de ventes en ligne et au manque de clarification conceptuelle de cette concurrence. Elle permet de répondre aux questions suivantes : quel est mon marché ? Quels sont mes concurrents les plus dangereux ? Quelle est l'intensité concurrentielle observable sur mes différents segments ? A travers une grille élaborée sur la base de la littérature, nous illustrons les apports des outils du *Web Analytics* dans l'analyse de la concurrence en prenant comme illustration deux sites importants de ventes en ligne du marché des produits culturels : Amazon et Fnac.

Les résultats de cette recherche montrent qu'il est possible, en mobilisant les outils du *Web Analytics*, de définir l'environnement marché par les éléments suivants :

- La détermination du marché par les requêtes
- Le nombre d'alternatives en fonction des mots clés du marché
- La segmentation du marché par mots clés
- La pression concurrentielle pour chaque mot clé représentatif

Par ailleurs, le processus de catégorisation des sites en fonction des avantages produits recherchés (évaluation multi-attributs) et l'ensemble de considération trouvent dans le *Web Analytics* une réponse à travers divers indicateurs dont la catégorisation des sites *via* les avantages recherchés par les internautes. Enfin, des outils dits de positionnement social évaluent la e-réputation des différents sites et la viralité de leur image telle qu'elle est perçue par les consommateurs.

Dans une première partie, nous présenterons les deux types de littérature ayant servi à la rédaction de cet article. Le premier type de littérature est d'ordre managérial. Il propose un état du *Web Analytics* à travers trois dimensions : son industrie, ses outils ainsi que ses apports et limites. Le deuxième type de littérature est d'ordre conceptuel. Il repose sur une synthèse des modèles d'analyses concurrentielles appliqués aux sites de commerce en ligne. Dans une deuxième partie, nous mènerons une analyse concurrentielle de deux sites de ventes en ligne : Amazon et La Fnac. Nous utiliserons, pour cela, les outils fournis par le *Web Analytics*. Dans une dernière partie, nous reviendrons sur les implications théoriques et opérationnelles de cette recherche.

1. Fondements managériaux et théoriques de la recherche

Cette recherche porte sur deux champs de la littérature distincts. Le premier champ est un champ managérial centré sur trois aspects du *Web Analytics* : l'industrie du *Web Analytics* et son évolution, les différents outils utilisés, les apports et limites de ce type d'analyses. Le deuxième champ est un champ plus ancien puisqu'il mobilise des concepts centraux de la réflexion marketing comme la délimitation d'un univers concurrentiel ou la notion de concurrent. Il s'est peu à peu enrichi des études empiriques réalisées sur la consommation en ligne.

1.1. L'analyse de l'univers concurrentiel des sites par le *Web Analytics* : industrie, outils, apports et limites

Selon Park, Kim et Koh (2010), l'augmentation considérable des données logs a permis d'analyser les comportements et les usages des clients et d'intégrer ces informations à l'élaboration des stratégies des différents acteurs de la vente en ligne. L'utilisation intensive de ces données explique l'essor de l'industrie du *Web Analytics*. Cette dernière peut être divisée en trois catégories : les outils du *Web Analytics* (logiciels d'extraction, d'analyse et de reporting des données), les services dits de *Benchmark* (qui permettent de définir et suivre sa concurrence), les fournisseurs de services de *Web Analytics* qui gèrent pour le compte de leurs clients l'implémentation et l'administration des outils. Les outils utilisés pour ce type d'analyses sont qualifiés par les praticiens d'outils « d'intelligence compétitive ». Leur objectif est de fournir une analyse de l'environnement permettant d'établir un avantage concurrentiel (choix d'une stratégie) pour l'entreprise. Ils reposent, à la fois sur de la veille stratégique, et sur la capacité à exploiter intelligemment les données pertinentes.

L'intégration des données issues du Web comme nouveau mode d'analyse des consommateurs et des marchés est soulignée par Belveaux et Flores (2010). Ces données peuvent servir à construire des modèles prédictifs de l'évolution des marchés. Elles sont également au fondement de l'élaboration des stratégies de e-business centrées sur les taux de conversion et la monétarisation des sites de vente en ligne. Enfin, elles servent de façon encore très marginale à mieux comprendre son marché et ses concurrents. Ainsi, leur principale utilisation repose sur le « monitoring » des performances commerciales des sites et très rarement sur l'analyse de la concurrence qui se fait, paradoxalement, de façon plus intuitive.

Tableau 1
Principales utilisations des outils de *Web Analytics*, adapté de Malo et Warren (2009)

Le trafic et l'acquisition
Les indicateurs clés de performance
Les tableaux de bord
Le parcours et la transformation
Les indicateurs de qualité de visite
Les indicateurs de popularité de contenu
Les indicateurs de qualité du trafic entrant
Les indicateurs de l'assistance en ligne

L'industrie du *Web Analytics* est une jeune industrie qui a pris son essor dans les années 2000. Elle repose sur deux technologies (les fichiers journaux ou logs) et les marqueurs ou tags et sur deux types de données : les données hors-site qui permettent de retracer les navigations des visiteurs avant leur entrée sur le site et les données sur site qui permettent de suivre le consommateur jusqu'à l'achat appelé conversion. Ces données de parcours sont fournies par une multitude d'outils dont le modèle économique a été bouleversé par l'entrée des moteurs de recherche sur le marché. Google Analytics ou Yahoo ! Web Analytics ont un accès gratuit. Leur technologie fondée sur l'implémentation de tags sur les sites a permis une réelle démocratisation des données. Ces outils gratuits poussent les éditeurs payants (Omniture, Xiti...) à sophistiquer leurs solutions. L'arrivée des réseaux sociaux a constitué une véritable révolution, celle de l'analyse des données issues des plateformes sociales popularisées par les outils gratuits comme Google Analytics. Parallèlement aux moteurs de recherche, les acteurs du e-commerce comme Amazon proposent leurs propres solutions. Elles sont fondées sur des panels d'internautes qui acceptent que l'on étudie leur navigation grâce à l'activation de cookies sur leur ordinateur. Cette voie de développement du *Web Analytics* présente l'avantage d'être parfaitement accessible mais peut être limitée par la faible représentativité des panels constitués.

L'avenir des outils de *Web Analytics* passe par une spécialisation accrue de leurs offres. Mochibot permet de suivre des fichiers Flash, Piwik représente une solution intéressante en open-source.

Le tableau suivant synthétise les apports et les limites du *Web Analytics*.

Tableau 2
Apports et limites du *Web Analytics*

Apports	Limites
Analyse de données à la fois qualitatives (taux de rebond, avis des consommateurs) et quantitatives des parcours d'achat	Forte variabilité des résultats due aux différents outils et technologies utilisées
Outils visuels (graphique des flux de visiteurs, visualisation des parcours)	Analyses essentiellement descriptives

Outils intégratifs s'adaptant à la multiplicité des formes du Web (blogs, réseaux sociaux, sites de e-commerce, sites de contenu...)	Grande richesse des données fournies qui rend difficile leur exploitation
Outils de tests permettant de retenir les solutions les plus optimales (tests d'ergonomie, de concepts ou de prix de produit)	Lourdeur de l'implémentation de certaines technologies comme la technologie des marqueurs qui nécessite le recours à des développeurs
	Croisement des données du Web Analytics et des informations d'autres canaux de distribution encore impossible

Parallèlement à une large diffusion de ces outils sous forme payante, gratuite ou libre, le *Web Analytics* tente de mieux définir son intégration aux réflexions stratégiques des directions marketing. Ces réflexions sont essentiellement cross-canales et tentent de définir des croisements et synergies d'informations entre les différents canaux de distribution (achats en magasin versus en ligne). Dans cette optique, le *Web Analytics* se fonde sur six axes définis par Kaushik (2010) : la récupération des données, l'analyse des résultats, les expérimentations et tests, l'analyse des antécédents des comportements de navigation, l'intelligence concurrentielle, la connaissance exploitable. C'est sur l'intelligence concurrentielle que portent nos recherches.

1.2. Quelle modélisation de la concurrence pour les sites de ventes en ligne ?

La concurrence est un concept central multi-dimensionnel. Elle repose sur trois concepts : la similarité, la substituabilité et la mesure de l'intensité concurrentielle. L'ensemble de considérations en est une des premières opérationnalisations. La mesure de la concurrence des sites de ventes en ligne mobilise, en outre, les stratégies de recherche d'information des internautes.

1.2.1. Substituabilité, similarité et intensité concurrentielle

L'étude de la concurrence en marketing est fondée sur trois concepts clés : la substituabilité, la similarité et l'intensité de concurrence reposant sur l'ensemble de considération.

La substituabilité peut se définir comme le fait d'être « interchangeable dans une situation d'usage particulier » (Shocker et al., 1990), c'est à dire la capacité d'une marque à se substituer à une autre¹. Elle varie selon la situation d'usage envisagée par le consommateur (Aurier, 1993).

¹ La substituabilité perçue des produits a un caractère évaluatif et repose sur le concept de préférences ou d'ensemble évoqué (Chandon et Strazzieri, 1986). La substituabilité peut donc être opérationnalisée, soit selon le niveau de préférence (intensité de l'attitude), soit selon l'ensemble évoqué.

La similarité est, par contre, stable d'une situation à l'autre. La similarité repose sur la comparaison des perceptions des caractéristiques possédées par chaque produit. L'opérationnalisation de la similarité se fait en évaluant le nombre de caractéristiques distinctives et communes (Tversky, 1977). Elle ne prend pas en compte la situation d'usage et est donc stable dans le temps contrairement à la substituabilité (Aurier, 1993). L'importance des marques prototypiques dans les calculs de similarité permet également d'introduire la notion d'asymétrie dans la similarité entre marques (la marque B est plus similaire à A que A à B). Cette asymétrie aboutit à considérer que les marques d'un même marché sont des sous-ensembles d'attributs de la marque prototypique.

La mesure de l'intensité de concurrence est fondée sur la notion d'ensemble de considération (Howard, 1963). Du fait de la limitation des capacités cognitives et d'un processus de choix de type conjonctif (Pras et Summer, 1975), il est intéressant d'étudier chez le consommateur les considérations simultanées de deux alternatives (*combined set*) (Wilson, 1981). Les considérations simultanées sont directement issues du courant de recherche sur l'ensemble de considération. Chandon et Strazzeri (1986) décrivent la séquence type d'interrogation permettant de délimiter un marché et d'évaluer la concurrence intra-marché : saillance des marques (marques qui viennent à l'esprit), ensemble de rejet (marques que l'on n'achèterait certainement pas), ensemble de considération (marques que l'on envisagerait sérieusement d'acheter). Le nombre de fois où deux marques sont citées ensemble correspond au nombre de considérations simultanées. Il permet d'identifier les cas de sur-concurrence et de sous-concurrence et donc de délimiter des champs de concurrence. Par ailleurs, les volumes des champs de concurrence (volume net des considérations) permettent d'avoir une vision « de l'ampleur des divers champs de concurrence constitutifs du marché » (Chandon et Strazzeri, 1986).

1.2.2. Comment la concurrence entre sites de ventes en ligne s'exprime-t-elle ? Synthèse de la littérature

La littérature s'est d'abord interrogée sur les différences de comportement du consommateur entre un site marchand et un magasin. Helme-Guizon (2001) souligne que le mode de fréquentation d'un site en ligne est, soit un comportement dirigé vers un objectif précis (ce qui va nécessiter une stratégie spécifique de recherche d'information), soit un comportement de type expérientiel.

Par ailleurs, trois types de variables modératrices sont à prendre en compte : les variables d'atmosphère (ergonomie, design), les variables situationnelles (rôle des communautés, échange social...), les variables individuelles (les motivations cognitives et affectives, le risque perçu, les variables psychographiques). Au final, l'évaluation de l'offre commerciale par le consommateur semble facilitée et le prix acquiert une visibilité accrue (Lynch et Ariely, 2000).

La modélisation du processus de choix d'un point de vente peut s'appliquer pour une grande part au choix d'un site en tant que canal de distribution (Engel, Kollat et Blackwell, 1968). Ce choix passe par plusieurs étapes : l'identification de l'ensemble évoqué des sites sur les caractéristiques perçues, le processus de catégorisation effectué selon la forme perçue, le processus cognitif d'évaluation multi-attributs en fonction des avantages produits recherchés. Filser (1985) parle

de concept de polarité (combinaison d'avantages recherchés). Par ailleurs, le modèle de concurrence des sites de ventes en ligne est constitué de variables mesurant l'intensité concurrentielle, la similarité entre sites, la substituabilité entre alternatives.

Tableau 3
Principales variables étudiées par la littérature permettant de mesurer la concurrence entre sites

Type de variables étudiées par la littérature	Modèle de concurrence
Aides à la sélection, agents de recommandation, moteurs de recherche Stratégie de sélection des offres par l'internaute Environnement marché Ensemble évoqué	Mesure de l'intensité concurrentielle (nombre d'alternatives possibles produit lors d'une requête) Substituabilité (alternatives proposées au consommateur en fonction de la situation d'usage souhaitée –types de recherche, mots clés)
Attributs produits Fonctionnalités des sites	Similarité (classement des sites en fonction du type de produit vendu et de leurs différentes fonctionnalités, émergence de sites de référence prototypiques d'une catégorie)

Parmi les avantages recherchés, la réduction du risque perçu apparaît centrale sur Internet. Ce souci de réduction des risques est corrélé positivement à la recherche d'information (Dandouau, 1999). Plus récemment Punj et Moore (2009) ont effectué une synthèse de la littérature sur l'ensemble de considération des sites de vente en ligne. Ce dernier serait constitué de cinq variables décrites dans la littérature comme des concepts clés permettant d'expliquer les différents champs de concurrence de ce marché (Punj et Moore, 2009) :

Variable 1 : Les aides à la sélection, agents de recommandation ou moteurs de recherche

La recherche d'information sous forme électronique est un axe central de développement de la vente en ligne. Les alternatives sont évaluées en fonction des filtres d'information correspondant ou non à la recherche de l'internaute.

Variable 2 : Les stratégies de sélection des offres

Les stratégies de sélection des offres permettant d'évaluer les différentes alternatives du marché reposent sur des critères de sélection qui peuvent être des critères d'exclusion ou des critères additionnels. Les critères d'exclusion peuvent amener à considérer assez peu d'alternatives et à éliminer de façon prématurée des alternatives attractives.

Variable 3 : Les attributs produits

Les attributs produits peuvent se scinder en deux types : l'information digitale et l'information non digitale. Les attributs digitaux sont des attributs produits qui sont communiqués directement par l'environnement digital (prix, taille...).

Variable 4 : Les fonctionnalités des sites

Il apparaît que les sites ayant un certain nombre de fonctionnalités ont un avantage concurrentiel important permettant de transformer un visiteur en acheteur.

Variable 5 :L'environnement marché

Deux variables environnementales influencent la recherche d'information et la formation de l'ensemble de considération : le nombre d'alternatives (produits possibles) et le temps disponible. Le nombre d'alternatives dépend de la stratégie de recherche de l'internaute. Si celui-ci décide de multiplier les critères de sélection, l'ensemble de considération peut se voir considérablement réduit. Par ailleurs, il peut être amené à intensifier les recherches (nombre d'itérations de recherche) afin de définir les alternatives possibles.

Ainsi, Punj et Moore (2009) classent les recherches sur l'ensemble de considération des sites de vente en ligne selon deux directions : la taille de l'ensemble de considération et l'hétérogénéité de l'ensemble de considération (évalué en fonction du poids des attributs des différentes alternatives).

En outre, un certain nombre de ces recherches aboutit à l'élaboration de typologies de sites et à un essai de comparer les sites à d'autres univers de consommation et donc d'autres univers de concurrence. Spiller et Lohse (1997) établissent une correspondance entre l'achat en ligne et l'achat sur catalogue. Néanmoins, l'interaction entre les consommateurs et les services proposés en ligne est moins porteuse de transaction et de création de valeur qu'attendue.

2. Le *Web Analytics* appliqué à une analyse concurrentielle : Fnac.fr versus Amazon.fr, description de la méthodologie de recherche et contexte

2.1. Le choix du terrain de recherche

Le choix du terrain de recherche se veut illustratif d'une démarche d'analyse d'un marché à travers les outils du *Web Analytics* et non d'une démarche d'analyse d'un marché particulier. Le marché de la vente en ligne des produits culturels et de haute technologie a été choisi pour son appétence particulière avec l'achat en ligne. Selon la Fevad², 18% des ventes de ce marché sont faites sur Internet ce qui en fait le secteur dont la part de marché sur le Web est la plus forte.

Par ailleurs, la Fnac et Amazon sont en concurrence frontale sur le marché des livres mais également sur certains produits numériques comme les liseuses. Cette catégorisation est une catégorisation par les offreurs eux-mêmes. Elle ne l'est pas forcément pour les utilisateurs sur tous les produits, Amazon développant une place de marché étendue sur des produits de consommation courante. Enfin, le dynamisme de ce marché a rendu l'analyse de l'intensité concurrentielle possible. En effet, le marché de la vente en ligne des produits culturels et high-tech est en pleine croissance. Au quatrième trimestre 2011, Amazon représentait presque 13 000 000 de visiteurs uniques³ moyens par mois, 9 730 000 visiteurs uniques moyens par mois pour Fnac. Parallèlement, ce marché subit trois évolutions

² Source Fevad, chiffres clé, vente à distance, e-commerce, 2012

³ Source Médiamétrie, observatoire des usages Internet, 4ième trimestre 2011

majeures. La première évolution tient à l'arrivée de nouveaux entrants aussi bien dans l'univers du discount des *Pure Players* que dans celui des *Clicks and Mortars* avec des sites de ventes en ligne de produits culturels comme Carrefour. La deuxième évolution tient au rôle central de la technologie qui amène les marchands à redéfinir de façon régulière leurs segments. Même le livre, produit traditionnel est révolutionné par le numérique et la distribution de contenus dématérialisés (Amazon MP3, Kindle DX) amène les sites à refondre leur infrastructure technologique (Isaac et Volle, 2010). La dernière évolution est liée à un glissement stratégique fort qui voit les deux leaders se diversifier au risque de brouiller leur image. Amazon se lance avec sa *Marketplace* dans une stratégie à la eBay et se transforme, peu à peu, en plate-forme de ventes de produits divers. Cette stratégie est suivie par la Fnac qui développe parallèlement ses activités de billetterie. Selon une étude de l'observatoire de la e-performance de Yuseo, les grands sites de produits culturels enregistrent une note de satisfaction qui se dégrade parmi les internautes. Pour la Fnac, l'élément prix est jugé peu compétitif. Pour Amazon, la profondeur de l'offre nuit à sa lisibilité.

2.2. La démarche de la recherche

La démarche de la recherche a été de sélectionner les informations apportant un éclairage sur le type d'analyse concurrentielle qu'il est possible d'effectuer grâce au *Web Analytics*. Deux outils ont été sélectionnés prioritairement :

- L'outil Alexa fonctionne sur la base d'un panel d'internautes volontaires. Il suit leur navigation et fournit des données pour les sites classés dans les 25 premiers millions. C'est un outil dit *user centric* car fondé sur l'étude du comportement des utilisateurs. Cet outil fournit, en outre, une classification des montants investis en communication par les différents sites selon trois catégories (intensité publicitaire forte, moyenne ou faible) calculée sur le total des investissements investis sur le secteur. Il suffit de s'inscrire sur le site d'Alexa pour disposer des données du panel.
- Les différents outils mis en ligne par Google Analytics notamment Google Trends et DoubleClickAdplanner fonctionnent à partir des données Logs et représentent une des offres les plus complètes du marché. Ce sont des outils dits *site centric* car fondés sur l'étude des données issues des sites eux-mêmes. Google Trends répertorie l'ensemble des requêtes des internautes au niveau mondial et sur plus de dix ans. Il comporte un volet prédictif puisqu'il calcule des courbes de régression multiples en fonction des données collectées. DoubleClickAdplanner est un outil de média-planning mesurant la puissance des sites et apportant des informations sur le nombre de visiteurs et leur profil socio-démographique. Google Trends est gratuit et disponible *via* l'adresse googletrends.com. DoubleClickAdplanner, gratuit également, est accessible *via* l'ouverture d'un compte Gmail.

L'ensemble des données issues de ces deux types d'outils a été retraité et reconfiguré afin de fournir une analyse concurrentielle reposant sur les principales conceptualisations de la concurrence. Le retraitement des données s'est fait selon deux étapes. Une première étape a consisté à collecter l'ensemble des informations fournies par les outils de *Web Analytics* retenus. Une seconde étape a

été de les mettre en forme soit sous forme graphique, soit sous forme de tableau. L'essentiel du travail de reconfiguration porte sur la mise en lien des concepts issus de la littérature et des ratios et chiffres fournis par le *Web Analytics*. La correspondance des données et des concepts représente le cœur de l'analyse effectuée. Par ailleurs, nous avons fait le choix d'utiliser des outils en libre accès (d'où notre sélection d'Alexa et de Google), l'objectif étant de pouvoir tenir à disposition des praticiens et des chercheurs une grille d'analyse pouvant être aisément répliquée pour tous les secteurs et toutes les périodes.

2.3. Les résultats de la recherche

2.3.1. L'environnement marché

La visualisation des frontières du marché

Les frontières d'un marché sont, souvent, difficiles à déterminer. Dans le cas d'une approche par le *Web Analytics*, il est possible d'estimer un marché en fonction des requêtes des internautes. Si nous prenons comme référence le premier site étudié à savoir Fnac, des informations nous sont fournies sur les autres sites que les utilisateurs de Fnac ont consultés et les requêtes afférentes. A partir d'un chaînage que l'on pourrait qualifier de cognitif, le marché de la Fnac se dessine (produits et concurrents). Chaque requête faite par un internaute permet d'identifier le ou les sites prioritaires qu'il va consulter mais également les mots-clés liés⁴. Ainsi, le marché de la Fnac comprend trois segments principaux :

- Le segment des livres
- Le segment des produits technologiques de loisir
- Le segment du spectacle

Sur le segment des livres, Fnac est concurrencée par des généralistes (Amazon, France Loisirs) mais également des acteurs dont le positionnement plus spécialisé ressort.

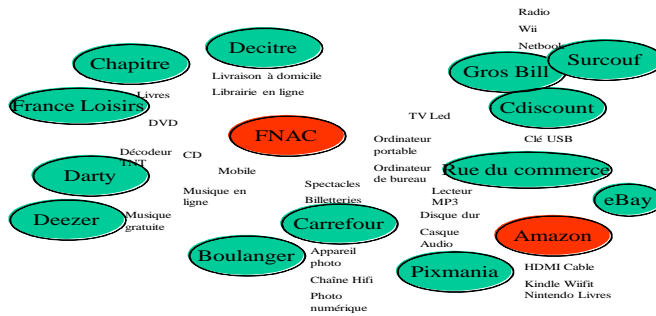
Sur le segment des produits technologiques, elle subit la concurrence de spécialistes pour certains produits (Darty et les décodeurs TNT), de discounters (Cdiscount) et d'innovateurs ayant une position de pointe sur certains segments (Amazon et son *Kindle*)

Sur le segment des spectacles, il apparaît que Carrefour est un acteur reconnu de ce segment.

Schéma 1

Visualisation du marché de la Fnac et d'Amazon par analyse des navigations et des requêtes

⁴ Pour la requête télévision, les mots-clés liés c'est à dire ayant une fréquence d'occurrence supérieure à la moyenne sont Led, écran plat, Samsung.



Le nombre d'alternatives

Le nombre d'alternatives peut être mesuré par le nombre de sites vendeurs du produit. Un premier niveau d'intensité concurrentiel peut être également estimé. Alexa, outil de *Web Analytics* permet de mesurer le nombre d'alternatives pour les différents segments de Fnac et Amazon et le niveau de popularité des mots clés représentatifs de chaque catégorie. Il donne également un indicateur de pression publicitaire (*query competition index*) par requête.

Tableau 4
Popularité des requêtes, nombre d'alternatives et pression publicitaire

Requête représentative des types de produits vendus par Amazon et Fnac (recherche France)	Nombre d'alternatives ou sites vendeurs	Indicateur de pression publicitaire sur 100/ nombre de publicités portant sur le mot clé, pression classée en trois catégories (pression publicitaire faible, moyenne, forte)	Popularité de la recherche sur 100/ indicateur de fréquence de recherche du mot clé classé en trois catégories (popularité faible, moyenne, forte)
Livres	1319	73, pression forte	36, popularité faible
DVD	1043	59, pression modérée	61, popularité modérée
Musique	5114	39, pression modérée	49, popularité modérée
Jeux vidéos	923	12, pression faible	45, popularité modérée
TV	2202	36, pression moyenne	69, popularité forte

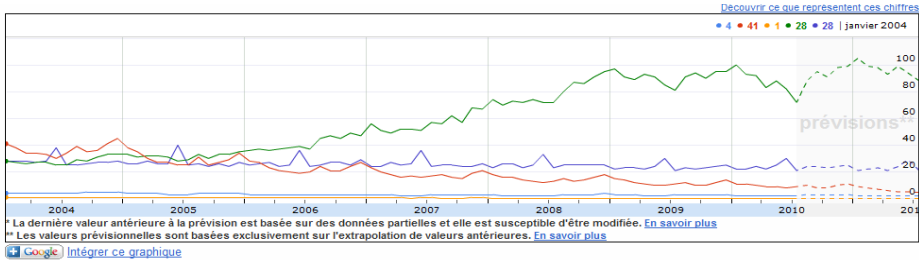
2.3.2 La recherche d'informations par le nombre de requêtes

Le nombre de requêtes peut être mesuré de façon dynamique par Google Insight. L'évolution des requêtes liées aux différents segments produits d'Amazon et de Fnac montrent :

- Une augmentation des requêtes liées aux produits bruns (télévision)
- Une stagnation des requêtes liées aux livres, DVD et jeux vidéos

Cette utilisation prédictive de Google Trends permet d'estimer le succès d'un positionnement qui, dans le cas de la Fnac est de plus en plus axé sur la vente de produits multimédias.

Schéma 2
Evolution du nombre de requêtes Télévisions, Livres, DVD, Musique Jeux vidéos



Légende : **T**élévision, **L**ivres, **D**VD, **M**usique, **J**eux vidéo

2.3.3 Le processus de catégorisation des sites en fonction des avantages produits recherchés

Catégorisation en fonction des avantages recherchés

Le processus de catégorisation des sites de ventes en ligne a pour particularité d'être élaboré par les marchands eux-mêmes et par les consommateurs. La distinction faite par Aurier (1993) entre la réalité marché et la réalité consommateur se retrouve.

La réalité consommateur c'est à dire la catégorisation des sites en fonction des avantages produits recherchés est visible à travers la structure interne des sites proposée par les moteurs de recherche. Cette structure interne ou *sitelink* est fondée sur l'adéquation entre la requête demandée et le site, cette adéquation étant mesurée par les statistiques de trafic. Elle représente les liens les plus souvent cliqués sur la page d'accueil du site.

La réalité marché c'est à dire la catégorisation faite par les marques elles mêmes est fondée sur le recours systématique aux annuaires. Tout site de ventes en ligne effectue une démarche d'inscription préalable dans un certain nombre de catégories au sein des annuaires. *Dmoz*, un annuaire mondialement contenu repose sur une arborescence de 16 catégories au sein duquel la catégorie « commerce électronique » est composée de 46 sous-catégories dont une catégorie « boutique en ligne » dans laquelle est référencée la Fnac.

Tableau 5
Catégorisation des sites via les avantages recherchés par les internautes, principaux *sitelinks*

FNAC	AMAZON
dvd	livres

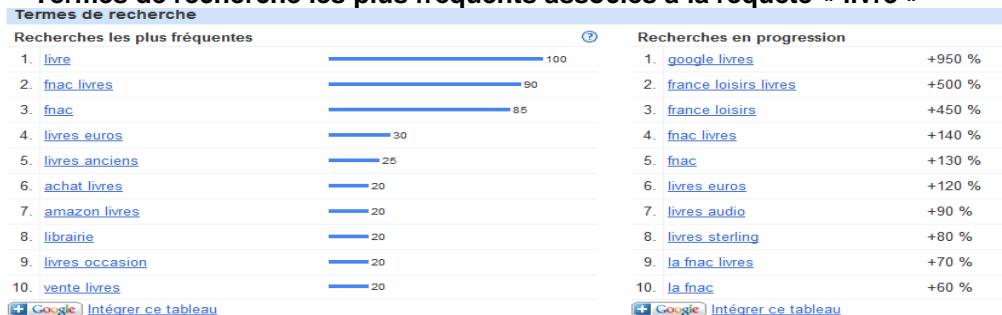
informatique	musique
Jeux vidéos	DVD
Logiciels	informatique
CD	TV et vidéo
GPS	Jeux vidéo
livres	Blu ray
Billetterie Spectacles	Affiliation

Sites de référence prototypiques

Par ailleurs, les sites de référence prototypiques d'une catégorie sont facilement identifiables. Google Insight « détermine la relativité en examinant les recherches ayant été effectuées par un groupe important d'internautes, avant et après le terme de recherche que vous avez spécifié » (Source, Google Analytics). Ainsi, Fnac et dans une moindre mesure Amazon apparaissent prototypiques du secteur du livre mais pas de celui des DVD ou des jeux vidéo.

Schéma 3

Termes de recherche les plus fréquents associés à la requête « livre »

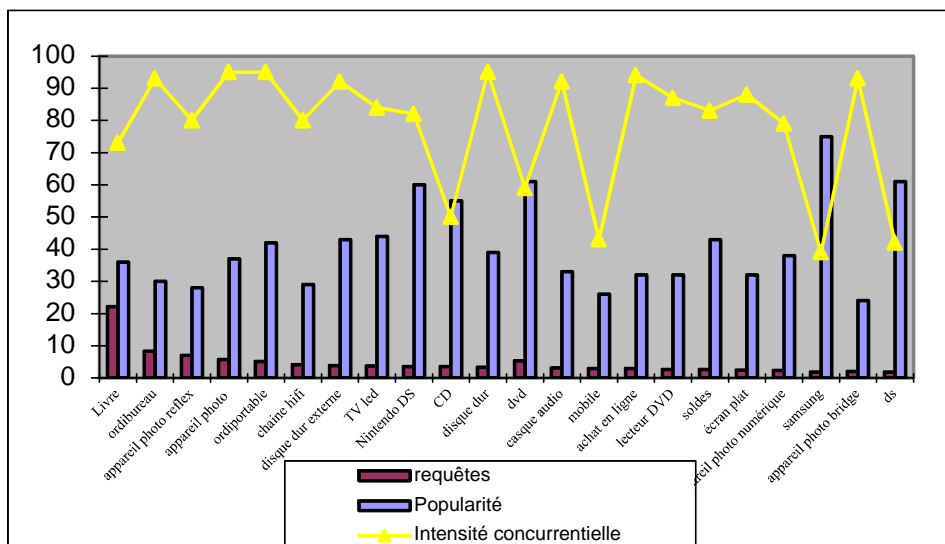


La requête Fnac livres arrive en seconde position, Amazon livres en septième. Il est, par ailleurs, intéressant de noter les recherches en progression comme celle de Google livres (+950 % de croissance).

La substituabilité d'un site par un autre grâce à l'analyse des requêtes

Enfin, les requêtes amenant le plus de trafic sur le site sont identifiées et permettent de déterminer sur quelles catégories de produit, chaque site est le plus pertinent. Il est, alors, possible de retraiter ces informations en enlevant les requêtes Fnac ou Fnac.com, ces dernières correspondant logiquement à plus de 70% des visites.

Schéma 4
Panorama des requêtes apportant du trafic sur le site de la Fnac



2.3.4. L'ensemble de considération

L'ensemble de considération des deux sites peut être déterminé par :

1. La navigation
2. La similarité des sites via l'affinité

La navigation

La navigation permet de déterminer les sites de provenance et de destination. Il apparaît que les internautes utilisateurs du site Fnac proviennent de Google à 35% et d'Amazon à 2,38%. On retrouve dans les sites de provenance Cdiscount et dans les sites de destination (sites immédiatement consultés après avoir quitté le site Fnac) Amazon à 3%. L'ensemble de considération par la navigation est brouillée par la place des moteurs de recherche et des réseaux sociaux (Facebook à près de 4%). Ces résultats fournis par Alexa peuvent être complétés par Google Ad Planner qui fonctionne sur l'affinité.

L'affinité

L'affinité est calculée selon la même méthode que celle pratiquée en médiaplanning à savoir le rapport exprimé en pourcentage ou en indice entre l'audience d'une cible et l'audience totale d'un support, d'une émission, d'une tranche horaire. Dans le cas d'Internet, la base est la population des internautes. Il apparaît que les sites les plus en affinité avec la Fnac sont Surcouf et

Chapitre.com, ceux les plus en affinité avec Amazon sont Web distrib et Price Runner. On retrouve un certain nombre de sites qui avaient été préalablement identifiés lors de la délimitation du marché par les requêtes.

Schéma 5⁵
Principaux sites en affinité avec Fnac

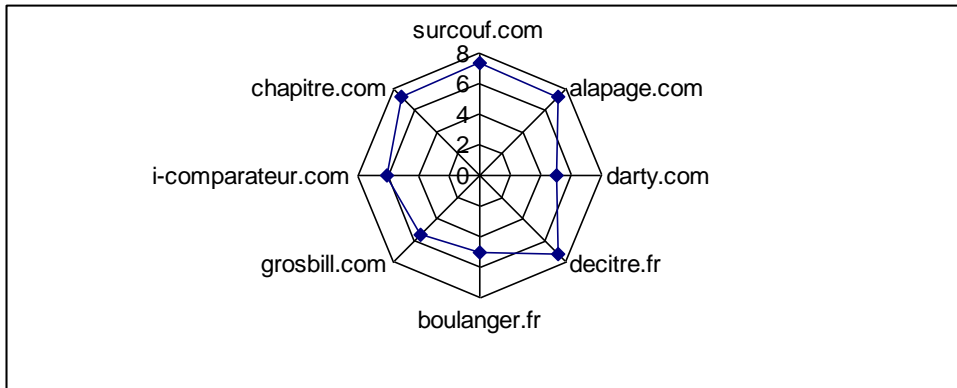
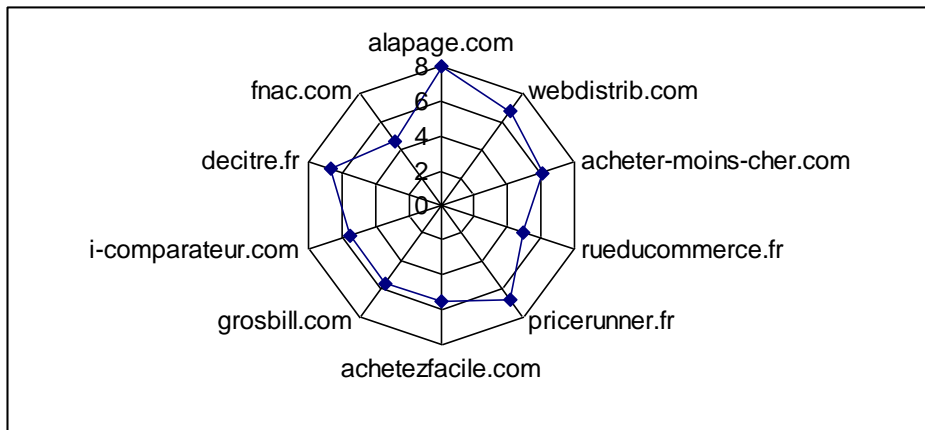


Schéma 6
Principaux sites en affinité avec Amazon



2.3.5 L'attitude à l'égard du site

L'attitude à l'égard des sites est également enregistrée par les outils de *Web Analytics*. Cette attitude est mesurée de façon qualitative et quantitative.

Sur le plan qualitatif, l'outil Alexa donne la possibilité aux internautes de poster des commentaires sur la qualité du site (livraison, produits, ergonomie...) et de le noter (nombre d'étoiles accordées à chaque site). Amazon dispose de 70 notations contre à peine 4 pour la Fnac. Ce résultat est à mettre en corrélation avec les sites

⁵ Il est à noter que le site alapage.com a disparu en mars 2012 soit quelques mois après notre terrain de recherche.

pointant sur Fnac et sur Amazon. Cet indicateur est un indice de la e-réputation des sites de ventes en ligne. 19723 sites pointent sur Amazon contre à peine 6999 sites sur Fnac. La politique d'affiliation d'Amazon est, certainement déterminante, dans la construction de sa réputation en ligne. Le lancement d'un dispositif similaire par la Fnac (dit de *marketplace*) n'est pas encore tracé par les outils du *Web Analytics*, ce qui souligne sa difficile émergence.

Sur le plan quantitatif, la qualité de navigation sur le site peut être un indicateur de la confiance accordée aux sites. La profondeur de navigation, le nombre de pages vues et le temps passé sur le site sont mentionnés dans la littérature, non pas uniquement comme des indicateurs de performances commerciales, mais également de confiance et d'implication à l'égard du site. Les résultats des deux sites sont assez comparables. 40% des internautes se connectant sur le site de la Fnac ne téléchargent qu'une seule page contre 35% pour Amazon. 7 pages sont vues en moyenne par internaute pour Amazon contre 4,5 pour la Fnac. La faiblesse moyenne du temps passé sur chaque site (quatre minutes environ) est plus problématique si l'on considère ces sites comme des canaux d'information sur les produits culturels et multimédia.

La littérature montre qu'il existe un lien positif entre le classement dans les moteurs de recherche (référencement naturel) et l'image des sites (Enquiro, 2007). L'outil l'espionnage d'Alexa donne les référencements des sites en fonction des requêtes. Par rapport à la requête « livres », Fnac arrive en septième position sur Google derrière notamment Evane, Google Books, Amazon ou Chapitre.com. Il en est de même pour la requête CD (rang 7 derrière Amazon, Priceminister, Wikipedia ou Cdiscount).

Enfin, l'image des sites passe par leur activité sur les réseaux sociaux et les contenus générés par les utilisateurs (*consumer generated media*). Ce buzz des sites en ligne est désormais identifié grâce à de nombreux outils. *Social Seek* analyse les blogs, les Tweets, les vidéos. La Fnac capte un buzz positif important grâce aux spectacles et événements qu'elle sponsorise. Amazon voit son trafic social conditionné par ses offres commerciales et l'activité de ses partenaires sur sa *Marketplace*. Le pic de fin juillet est dû aux innovations annoncées sur son produit phare le « *Kindle* » et mesure l'intérêt suscité par le marché.

Schéma 7
Présence sociale de la Fnac sur le Net (vidéos, images, posts)

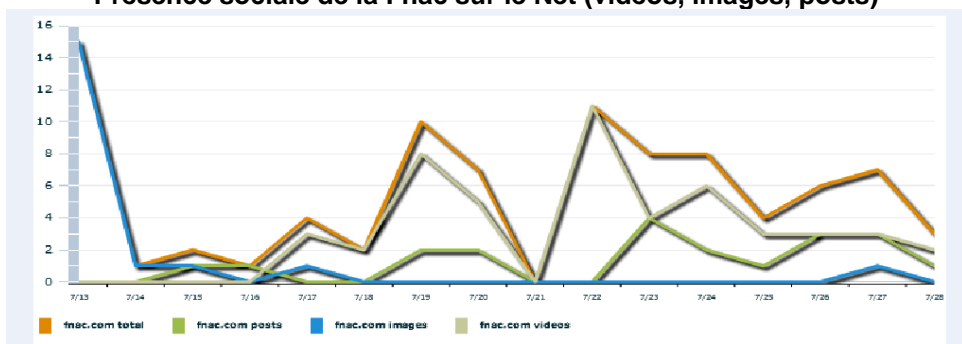
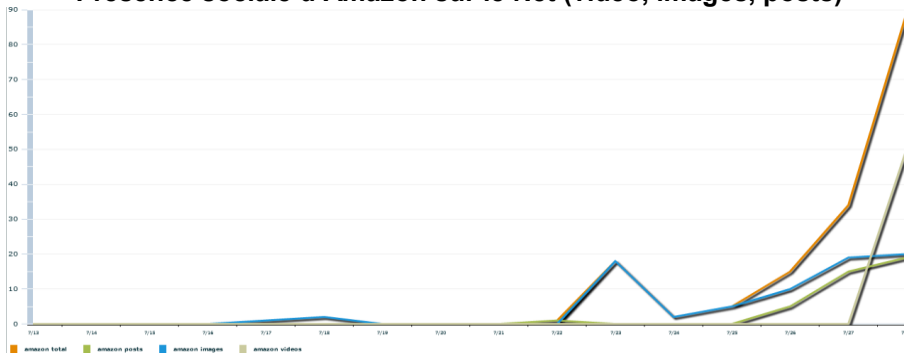


Schéma 8
Présence sociale d'Amazon sur le Net (vidéo, images, posts)



3. Recommandations managériales

L'analyse concurrentielle faite à travers les outils du *Web Analytics* montre que le marché des produits culturels et technologiques de loisir présente des champs concurrentiels différents. Par ailleurs, Amazon et Fnac ne semblent pas réellement concurrents. En effet, tant sur le plan de l'affinité, que de l'analyse des requêtes ou du positionnement social, Amazon et Fnac sont en compétition sur un faible nombre de segments.

Le *Web Analytics* représente donc un outil intéressant d'analyse de la concurrence puisqu'il permet de définir un marché en fonction des comportements des internautes. Le principal intérêt du *Web Analytics* est de se situer sur une analyse comportementale de la navigation. Les requêtes, le temps passé sur un site, la transformation ou non de la visite en achat permettent une matérialisation du concept de concurrence non plus sur les intentions mais sur les comportements. De ce fait, il peut être un outil majeur de réflexions managériales.

Néanmoins, du fait de la multiplicité des solutions proposées et des indicateurs, une démarche d'analyse concurrentielle pertinente doit suivre des objectifs préalablement définis. Elle peut permettre d'identifier ses concurrents les plus dangereux mais aussi de valider sa stratégie produit ou de suivre l'évolution de son marché. Deux apports du *Web Analytics* ouvrent des voies de réflexions intéressantes pour les managers: l'analyse concurrentielle fondée sur les requêtes et la notion de positionnement social.

L'analyse concurrentielle fondée sur les requêtes donne des résultats en accord avec la littérature sur la concurrence.

Elle permet de mieux comprendre la notion de substituabilité des offres. En effet, reposant sur les recherches d'information des consommateurs, elle est une illustration des différentes situations d'usage auxquelles une marque peut être confrontée.

La similarité est une similarité affinitaire. Ainsi, elle n'est pas fondée sur une similarité produit mais une similarité d'utilisateurs. Alors que l'intégration de la similarité dans les modèles de concurrence est largement critiquée (car la concurrence provient de la différenciation et non de la similitude), la similarité affinitaire est un signe de concurrence peu contestable. Son étude est rendue possible par la double fonction du Web : une fonction commerciale mais également un média qui, à ce titre, peut profiter des indicateurs de concurrence utilisés traditionnellement en média-planning.

L'intensité concurrentielle est particulièrement bien décrite par les outils du *Web Analytics*. Elle s'exerce à deux niveaux :

- Au niveau des requêtes et donc de l'achat de mots clés, certains étant plus apporteurs de trafic que d'autres et donc plus convoités par les annonceurs.
- Au niveau des sites eux-mêmes *via* le suivi des visiteurs uniques et des taux de conversion.

Par ailleurs, la notion de positionnement social peut être intégrée à une analyse de la concurrence car elle est révélatrice, non seulement de l'image d'un site et de son positionnement perçu, mais également de sa capacité à accroître le potentiel viral de cette image. L'image de Fnac doit son potentiel viral au sponsoring de spectacles, celle d'Amazon au réseau de ses affiliés.

De façon plus générale, la diversité des requêtes montre une forte capacité de différenciation des sites, notamment sur certains produits technologiques voire une tendance à l'extrême customisation de l'offre. Les requêtes permettent d'estimer ou de découvrir les préférences les plus pointues des consommateurs. Enfin, il est à noter que les coûts de changement mentionnés par la littérature semblent s'exercer sur le Net. La relative faiblesse des sites consultés pour chaque requête tend à prouver qu'une certaine fidélité à un petit panier de sites est d'usage. Quant à la question du prix, elle n'apparaît pas, paradoxalement, comme centrale dans l'activité concurrentielle des sites. En effet, l'analyse des navigations et des requêtes (schéma 1) montre des recherches essentiellement centrées sur les produits. La requête « prix » n'apparaît pas ce qui indique un volume de recherches trop faibles pour être identifiées. Tout se passe comme si les consommateurs avaient intégré un certain nombre de schémas cognitifs à savoir une relative faible dispersion des prix pour la majorité des sites et une spécialisation (qui est un positionnement en soi) sur des bas prix pour quelques sites discount.

Conclusion

Ainsi, le concept d'analyse concurrentielle est revisité à la lumière d'un nouvel outil de collecte de données (le Web Analytics) plus proche des comportements que des attitudes. En ce sens, il permet d'enrichir la littérature sur l'analyse concurrentielle qui s'est en grande partie construite à travers le déclaratif des consommateurs. Il complète utilement les études liant avantages concurrentiels et image du site. Si des auteurs comme Lim et Dubinsky (2004) mettent en exergue quatre facteurs explicatifs de l'achat en ligne (produit, interactivité, confiance et navigation), les analyses de régression effectuées se fondent sur des attitudes et des perceptions. De même des auteurs comme Van der Heijden and Verhagen (2004) montrent que des sites développant des avantages concurrentiels forts

bénéficient d'images fondées sur la facilité d'utilisation, le style du site, les performances logistiques ou l'image de l'entreprise. Ces variables sont positivement corrélées à l'intention d'achat et non à l'achat qui est une variable mesurée par le Web Analytics. Des recherches futures pourraient être entreprises en vue de croiser variables attitudeinales et comportementales ce qui permettrait d'enrichir considérablement les résultats issus du Web Analytics plus descriptifs qu'explicatifs.

Par ailleurs, les limites de cette recherche sont à souligner. La première limite tient à l'hétérogénéité des technologies utilisées (fondées soit sur les logs, soit sur les tags) qui fragilise la robustesse des résultats obtenus. La deuxième limite tient au terrain proposé (le marché des produits culturels). Notre terrain pourrait s'enrichir de l'analyse d'autres marchés permettant de renforcer la validité externe de nos résultats fondés sur le postulat qu'Amazon et Fnac évoluaient dans un univers concurrentiel relativement identique ce qui n'est le cas que pour certains segments de produits. D'une façon générale, l'analyse des données Web gagnerait en pertinence par l'intégration d'autres variables comme le type d'expériences créées, les ergonomies utilisées, les différents modes de navigation... Les voies futures de recherche sont nombreuses, cet article étant un premier essai exploratoire d'utilisation des données du *Web Analytics* à des fins de meilleures conceptualisations de la concurrence sur le Web. Elles peuvent être de deux types. Une voie plus conceptuelle permettrait d'affiner le concept de concurrence sur le Web. En quoi est-il différent de la concurrence dite *off line* ? Une deuxième voie serait de développer un des indicateurs de la concurrence comme la substituabilité en se focalisant uniquement sur les données utilisateur ce qui permettrait de construire des cartes de positionnement ou des arbres concurrentiels. Cette voie nécessite d'affiner les données des outils de *Web Analytics user centric* qui fournissent généralement uniquement les navigations des sites entrant vers les sites sortant afin de disposer de chaînages plus complets. Néanmoins, comme nous l'avons démontré, la relative simplicité d'accès aux données proposées par Alexa et Google par exemple, offrent une véritable opportunité pour les chercheurs curieux d'optimiser plus encore les stratégies de marques sur Internet. Nous espérons simplement qu'ils seront de plus en plus nombreux à s'y intéresser⁶.

⁶ Google l'a bien compris et propose depuis 2008 avec le Groupe WPP des bourses de recherche de \$80,000 pour le développement de recherche visant à mieux comprendre et mesurer l'impact d'Internet sur le marketing (Google and WPP Marketing Research Awards : <http://research.google.com/university/marketingresearchawards/>)

Bibliographie

Aurier, P. (1993) "Analyse de la structure des marchés : réflexions et propositions théoriques sur la relation entre deux alternatives de choix", *Recherche et Applications en Marketing*, volume 8, n°1, p.76-95.

Belveaux, B. et Flores, L. (2010) "L'utilisation de proxies du Web pour la prédiction des marchés, une application au marché cinématographique", *Décisions Marketing*, n°57, p.9 -18.

Chandon, J.L. et Strazzeri, A. (1986) "Une analyse de structure de marché sur la base de la mesure de l'ensemble évoqué", *Recherche et Applications en Marketing*, volume 1.

Engel, J. F., Kollat, D et Blackwell, R.D. (1968) *Consumer Behavior*, eds Rinhart & Winston, New York.

Enquior, E. (2007) "The Brand life of Search", <http://pages.enquiro.com/whitepaper-the-brand-lift-of-search.html>

Filser, M. (1985) "Analyse de l'adéquation des enseignes aux attentes des acheteurs : méthodologie et implications stratégiques", *Actes de la première conférence de l'AFM*, Le Touquet, p.192-221.

Helme-Guizon, A. (2001) "Le comportement du consommateur sur un site marchand est-il fondamentalement différent de son comportement en magasin ? Proposition d'un cadre d'appréhension de ses spécificités", *Recherche et Application en Marketing*, volume 16, n°3, p.25-38.

Howard, J.A. (1963) *Marketing management: Analysis and planning*, Homewood, IL:Irwin.

Isaac, H. et Volle, P. (2010) *E-commerce, de la stratégie à la mise en œuvre opérationnelle*, Pearson Education, Paris.

Lim, H. et Dubinsky, A. (2004), " Consumers' perceptions of e-shopping characteristics: an expectancy-value approach", *Journal of Services Marketing*, 18(7), p. 500-513

Lynch, J. G. et Ariely, D. (2000) " Wine online : search costs and competition on price, quality, and distribution", *Marketing Science*, volume 19, n°1.

Malo, N. et Warren, J. (2010) *Web Analytics*, Eyrolles, Editions d'organisations, Paris.

Park, JS., Kim, JJ. et Koh, J. (2008) " Determinants of continuous usage intention" in *Web Analytics services*, *Electronic Commerce Research and Applications*, n°9, p.61-72.

Pras, B. et Summer, J.O. (1975) "A comparison of linear and nonlinear evaluation process models", *Journal of Marketing Research*, n°12, p.276-281.

Shocker, S. et Zahorik S. (1990) “ Market Structure Analysis : Practice, Problems and Promise”, in *the Interface of Marketing and strategy*, G.Day, B.Weitz et R. Wensley éd.s., Greenwich, Jai Press.

Spiller, P., et Lohse, G. L. (1997), “ A classification of Internet retail catalogs”, *International Journal of Electronic Commerce*, 2(2), p.29-56.

Tversky, A. (1977) “ Feature of similarity”, *Psychological Review*, n°84, p. 327-352.

Van der Heijden, H. et Verhagen, T. (2004), “Online store image: conceptual foundations and empirical measurement”, *Information and Management*, vol. 41, n° 5.

Wilson, C.E. (1981) “A procedure for the analysis of consumer decision making”, *Journal of Advertising Research*, n°21.

Webographie

<http://www.lsa-conso.fr/de-grands-sites-de-produits-culturels-et-multimedias-recales-par-les-internautes,113614>

<http://www.mediametrie.fr/internet/communiques/barometre-trimestriel-de-l-audience-du-e-commerce.php?id=127>