

Twetnographie :

Utilisation de Twitter pour la recherche en marketing

JEAN-PHILIPPE GALAN

Maître de Conférences

CRM-Marketing (EAC-CNRS 5032)

IAE - Université de Toulouse 1 Capitole

2, rue du Doyen Gabriel Marty

31042 Toulouse Cedex 9

e-mail : jean-philippe.galan@iae-toulouse.fr

ALEXANDRA VIGNOLLES

Enseignant-Chercheur en Marketing

Inseec Business School

26, rue Raze

33000 Bordeaux

e-mail : avignolles@groupeinseec.com

Résumé : Nous présentons dans ce travail une méthode d'enquête permettant d'utiliser les données secondaires provenant des sites de microblogging et plus particulièrement du site Twitter dont la forte croissance cette année passée en fait un site incontournable pour qui s'intéresse à l'analyse des discours des consommateurs sur internet. Dans une première partie, les éléments théoriques de la collecte de données sur internet et de la condition du consommateur des services de microblogging sont discutés. Ensuite, les particularités du microblogging et de Twitter sont exposées d'un point de vue technique mais également dans leurs conséquences sur la relation du chercheur au terrain. Enfin, les utilisations passées ou potentielles de ce nouveau type de données sont présentées.

Sommaire :

INTRODUCTION

1 – ELEMENTS THEORIQUES

1.1 Les discussions entre consommateurs comme source de données

1.2 Le microblogging, reflet d'un passage aux temps hypermodernes

2 – ELEMENTS METHODOLOGIQUES

2.1 Présentation du microblogging et du service Twitter

2.2 Le recueil de données secondaires sur Twitter

3 – ELEMENTS PRATIQUES

3.1 Etudes académiques en marketing réalisées sur les données issues de Twitter

3.2 Voies de recherches ouvertes par l'accès aux données de Twitter

CONCLUSION

“Every breath you take

Every move you make

Every bond you break, every step you take

I'll be watching you”

(The Police, 1987, “Every breath you take”, A&M Records)

Si de nombreuses entreprises ont immédiatement vu dans l'arrivée du grand public sur l'internet la nécessité ou l'opportunité d'utiliser celui-ci tout d'abord comme media, puis, ses capacités évoluant, comme un véritable lieu de vente et de service, la recherche en marketing, s'est pour sa part tenue quelques temps en retrait avant d'adopter une démarche plus active. Elle n'a toutefois pas tardé à s'approprier ce nouvel objet et c'est dès la seconde partie des années 1990 que certains ont vu dans l'internet une source considérable de données sur les consommateurs (e.g. Armstrong et Hagel, 1996 ; Kozinets, 1997, Babin et Galan, 1998).

En effet, le fonctionnement et les développements de l'internet sont tels qu'ils donnent l'opportunité aux utilisateurs de produire du contenu et d'entrer en interaction. Les espaces de discussion mis à leur disposition permettent aux consommateurs de partager des idées et de fonder des communautés dont l'intérêt marketing est important, que celles-ci se structurent autour d'une marque (e.g. Sitz, 2008) ou d'une passion commune qui directement ou non, n'échappe jamais au marché (Kozinets, 2002b ; Cova, 2003). Proposant toujours plus d'interaction et la possibilité de partager toujours plus de contenu sont apparus plus récemment les sites de réseautage social qui comme leur nom l'indique permettent de constituer des réseaux sociaux sur l'internet. Ces sites ouvrent de nouvelles opportunités pour le marketing mais également de nouveaux défis.

L'objectif de cette communication est de présenter une méthodologie basée sur l'utilisation des services de microblogging qui permette aux membres d'un site de réseautage social de diffuser des informations aux autres membres de leur réseau. En particulier, nous allons nous intéresser au service Twitter qui – par ses caractéristiques techniques mais également parce qu'il nous semble être une cristallisation des changements qui bouleversent la société de consommation – présente un intérêt tout particulier pour la recherche en marketing. Cette communication sera articulée en trois parties. Tout d'abord nous présenterons l'intérêt théorique des sites de micro-blogging dans un contexte de recherche en marketing. Nous exposerons ensuite les caractéristiques techniques du service Twitter afin que le chercheur ou

le praticien comprenne comment tirer partie de celles-ci dans le cadre d'une enquête. Enfin, nous présenterons quelques applications possibles des données accessibles sur Twitter.

1 – ELEMENTS THEORIQUES

1.1 Les discussions entre consommateurs comme source de données

Parallèlement aux travaux sur le commerce électronique, la recherche marketing a rapidement perçu l'intérêt que présentait le réseau internet en tant que source de données. Si à la fin des années 1990 certains auteurs (e.g. Galan et Vernet, 2000) ont pensé que l'internet permettait de fortement réduire le coût d'accès aux données primaires (principalement par le moindre coût de diffusion, d'administration et de saisie des questionnaires) et bénéficier d'un plus fort taux de réponse, la forte sollicitation des individus à participer à des enquêtes par ce média a fortement contribué à gommer cet avantage (Florès, 2008). L'étude des fichiers log, aujourd'hui connue sous le terme de « web analytics », est une technique non intrusive, basée principalement sur l'analyse des visites et des clics qui a connu un très fort engouement, notamment de la part des professionnels. Elle est particulièrement utile pour l'analyse de l'efficacité des publicités, du référencement ou de l'ergonomie des sites (e.g. Ferrandi et Boutin, 1999 ; Galan, 2002 ; Diesbach, Galan et Chandon, 2006). Enfin, les forums de discussion et les newsgroups, par le fait qu'ils permettent l'observation non intrusive des discours de consommateurs sur les produits, services, marques ou expériences de consommation ont vite été pris comme source de données secondaires (e.g. Sivadas, Grewal et Kellaris, 1998 ; Godes et Mayzlin, 2004), en particulier pour l'étude des communautés de consommateurs (e.g. Armstrong et Hagel, 1996 ; Kozinets, 1997).

Il y a une dizaine d'années, Kozinets (1997, 2002a), présentait la netnographie, néologisme proposé par un des relecteurs de son étude sur la communauté des fans de la série X-Files (Kozinets, 1997). Kozinets utilise ce terme pour décrire une méthode « plus rapide, plus simple et moins coûteuse que l'ethnographie traditionnelle, non intrusive et plus naturelle que les focus groupes ou les interviews » (2002a ; p. 61) afin d'étudier les communautés en ligne. Cette méthode a été largement diffusée et a fait l'objet de nombreux travaux, y compris en France (e.g. Ayadi, 2005 ; Bernard, 2004 ; Bertrandias et Carricano, 2006 ; Cova et Carrère,

2002). Ce n'est pas ici le lieu pour décrire en détails cette méthode mais nous en rappellerons certains fondements qui paraissent importants pour la suite de l'exposé¹.

La netnographie est une méthode d'étude des communautés virtuelles. Cela implique plusieurs choses. Tout d'abord cela nécessite pour le chercheur de localiser les lieux d'échange de ces communautés (sites, forums, newsgroups, etc.). Ensuite, du fait que la thématique de la communauté est la base de l'objet d'analyse, il est important que le chercheur soit familiarisé, à la fois à la thématique (série télé, marque, objet, etc.) mais également au groupe. Ceci passe donc pour le chercheur par une phase « d'entrée » (Kozinets, 2002a ; p. 63) dans la communauté qui nécessite une question de recherche, une connaissance de l'objet de consommation, et de se faire accepter en tant que membre de la communauté et participer aux échanges. Une autre particularité de la netnographie est qu'elle est fondée sur l'observation de communications textuelles (quoique le recours possible aux MMORPG² laisse la possibilité de contenus de tous types) qui seront mises en perspective par les expériences du chercheur avec le produit ou marque ainsi que par des sources documentaires (ouvrages, etc.). La netnographie implique – du fait du fonctionnement même des forums de discussion – que les comportements et communications observables sont circonscrites au domaine de l'internet : ce ne sont pas des individus qui sont étudiés mais des avatars matérialisés par des pseudonymes et dont le comportement est contraint par les règles de l'interface. Enfin, Kozinets (1997 ; 2002a) appuie sur la partie éthique de ce type de recherche et préconise au chercheur (1) de faire ouvertement son « entrée » en prévenant la communauté de ses intentions de collecte et du thème de sa recherche, (2) d'assurer l'anonymat et la confidentialité des informants, (3) de chercher le feedback de la communauté et de l'intégrer, et (4) obtenir la permission des membres avant d'utiliser leurs propos.

Si l'analyse des forums et des listes de diffusion reste une source de données très pertinente pour l'étude des communautés virtuelles, il faut remarquer que depuis l'article fondateur de Kozinets (2002a), les dispositifs mis à la disposition des utilisateurs afin de produire du contenu sur l'internet ont considérablement évolué. Tout d'abord les blogs, puis très rapidement les sites de réseautage social et avec eux le microblogging, que ce soit grâce aux

¹ Le lecteur désireux d'avoir plus de détail se référera à l'article original (Kozinets, 2002a) ou à la description qui en est faite par Bernard (2004) en langue française.

² Massively Multiplayer Online Role-Playing Games. Kozinets (2002a) n'emploie ce terme mais celui de « Donjons » qui en est un cas particulier.

statuts Facebook dans un premier temps puis par le biais de systèmes spécialement dédiés tels que Twitter, Jaiku ou identi.ca.

Ces services bouleversent les comportements des consommateurs, sur internet au moins, et obligent à proposer de nouvelles méthodes qui, sans présenter une rupture totale avec la netnographie, présentent suffisamment de particularités pour constituer un autre type d'outil et d'applications. Dans les lignes qui suivent nous allons présenter rapidement le principe du microblogging, en quoi il correspond à des changements sociaux plus profonds, et enfin quels sont les implications de ses caractéristiques sur la tâche du chercheur.

1.2 Le microblogging, reflet d'un passage aux temps hypermodernes

Les sites de microblogging sont fondés sur un double principe : (1) ils permettent à chaque utilisateur de publier un contenu textuel court (en général moins de 200 caractères par message), (2) ils permettent de construire un réseau afin que chaque utilisateur puisse lire le contenu des autres. Il s'ensuit une double activité : en tant que producteur de contenu, l'utilisateur doit trouver des lecteurs et les convaincre de lire le contenu qu'il produit ; en tant que consommateur de contenu, il doit trouver d'autres utilisateurs pour lire leur production.

La recherche de l'infini

Cette recherche d'utilisateurs à lire ou à persuader est une activité à part entière qui peut être source de plaisir, ou tout du moins de valorisation. En effet, sur ce type de site, le nombre de « lecteurs³ » dont peut se vanter un utilisateur lui confère un statut à part, censé traduire son influence et, partant, la pertinence de ses propos (Vuitton, 2009). Le nombre de lecteurs est ainsi un indicateur de reconnaissance de l'intérêt de la production de l'utilisateur par les autres membres du site. Pour qu'un autre utilisateur lise le contenu, il faut en général lire soi-même le contenu de l'autre. C'est un moyen de se porter à sa connaissance dans la masse des utilisateurs du service. Aussi, de manière très majoritaire, un grand nombre de lecteurs est associé à un grand nombre de personnes que l'utilisateur lit (Galan et Vignolles, 2010). Cette mécanique illustre bien ce que rapportent Carù et Cova (2008) sur « l'axiome ontologique de la subjectivité lacanienne : plus je suis actif, plus je dois être passif ailleurs, c'est-à-dire plus il

³ Le terme sur Twitter est « followers », ceux qui sont lus par l'utilisateur sont appelés « following ». Nous traduirons ces termes respectivement par « suiveurs » et « suivis » dans la suite de ce travail.

faut qu'il y ait un autre qui soit passif à ma place ». Les deux auteurs discutent d'une clé de lecture du comportement du consommateur postmoderne : l'interpassivité, pendant de l'interactivité. Ils décrivent la jouissance née du décalage qui existe entre la puissance quasi-infinie de certains objets de consommation et l'utilisation limitée qui en est faite par le consommateur. Ceci serait d'autant plus vrai que l'objet est simple, c'est-à-dire présentant un nombre très restreint de fonctions. Les sites de microblogging semblent pour la plupart clairement appartenir à ce type d'objet. Ils ne proposent que peu de fonctions : écrire, lire, suivre quelqu'un, se faire suivre par quelqu'un ; et leur limite d'usage ne sera jamais atteinte : aucun individu ne peut lire tout ce qui circule sur un site comme Twitter, il ne peut pas suivre les millions d'utilisateurs inscrits et ne pourra jamais convaincre tous les utilisateurs inscrits de le suivre. Quels que soient ses efforts⁴.

Pour Carù et Cova (2008), la jouissance serait ainsi, non dans l'usage, mais dans la conscience de l'incomplétude certaine de l'usage et dans l'idée que l'on peut renvoyer cet usage dans le futur, un futur qui n'arrivera jamais. Cette dernière idée paraît très intéressante et permet d'aller au-delà du cadre postmoderne que se fixent les auteurs. Le consommateur postmoderne a perdu toute confiance dans les grandes structures socialisantes et les grandes idéologies (métarécits) et particulièrement dans l'idée du progrès moderne. Cette perte de confiance dans l'avenir l'amène à se recentrer sur le présent afin d'en jouir (*carpe diem*). Il a également pris conscience que les entreprises ont ôté la part créatrice des activités humaines en les faisant passer dans le domaine public (plats cuisinés, prêt-à-porter, etc.). Le rejet de la société moderne se traduit par une tendance, à travers la consommation, à se réapproprier cette part de création. Le consommateur n'est plus un individu passif soumis à des mécanismes commerciaux rationnels qui conduisent à un monde désenchanté, il choisit de s'immerger dans des expériences dont il est un véritable acteur. Dans ces expériences, il cherche l'immersion afin de rompre avec un quotidien stressant, d'atteindre le plaisir (flow), et donner un sens à sa vie. Cependant, pour Lipovetsky (2006 ; Lipovetsky et Charles, 2004), cette période euphorique et libératoire (Firat et Venkatesh, 1995) n'est que transitoire et doit laisser la place à un rapport au présent qui n'est plus vécu sous le régime de l'émancipation, mais sous celui de la crispation. La peur de l'avenir l'emporte sur la jouissance du présent.

⁴ L'acteur Ashton Kutcher et la compagnie CNN ont récemment sollicités beaucoup de moyens médiatiques pour être le premier à gagner la course au million de suiveurs sur Twitter. Quoiqu'A. Kutcher ait gagné cette course, le million d'utilisateurs n'est qu'une partie des abonnés à ce service et le nombre d'inscrits augmente tous les jours rendant cette quête impossible.

Le microblogging qui se traduit par une incomplétude de l'usage qui renvoie à un futur qui ne sera jamais atteint (Carù et Cova, 2008) paraît être le moyen idéal pour que l'individu ait de lui-même une vision infinie, de se bercer d'une illusion d'immortalité dans laquelle il y aura toujours un but à poursuivre. Le fait de suivre le contenu d'un grand nombre d'utilisateurs participe de la même logique. En tant que lecteur du contenu des autres, le consommateur a accès à un nombre infini d'informations en tous genres. Aussi, et nous le verrons plus loin, plutôt que de s'intéresser au contenu produit par chacun, il va traiter l'information de tous les utilisateurs en une seule fois. Ce rapport à la production d'autrui change le rapport à l'information. En s'intéressant à un seul utilisateur, la lecture prend une dimension temporelle. Le lecteur lit la vie de l'utilisateur : ce qu'il a fait la semaine dernière, ce qu'il a fait hier, ce qu'il a fait aujourd'hui. En s'abonnant au contenu de centaines voire de milliers d'utilisateurs, il ne peut traiter que l'instant présent. Chaque seconde des milliers d'informations défilent sur les sites de microblogging et jamais l'utilisateur ne pourra les lire. Comme pour le nombre de lecteur, cela renvoie à l'infini et, partant, à l'immortalité. Mais cette focalisation sur l'immédiat permet également à l'individu de vivre continuellement l'instant présent. Un instant présent qui défile et qui ne finit jamais : « un sentiment d'éternité dans un monde livré à la fugacité des choses » (Lipovetsky et Charles, 2004 ; p.33). Dans ce flux d'informations il n'y a pas de futur. L'individu est avec chaque utilisateur, partout, en même temps, tel une divinité il échappe au commun des mortels et peut apaiser sa peur de l'avenir et de la vie quotidienne.

L'individu au centre de la communauté

L'intérêt commun est au centre de la notion de communauté. Cet intérêt commun est souvent, même de manière indirecte, lié à un objet de consommation : marque, produit, service, pratiques diverses. Le fonctionnement de l'internet a immédiatement permis – grâce en particulier à l'ensemble de protocoles Usenet – le regroupement d'individus autour de passions communes, de la cuisine aux discussions sur les mathématiques en passant par la tonte des caniches. Le consommateur postmoderne, sujet fragmenté et multiphrénique (Firat et Venkatesh, 1995) peut appartenir à plusieurs de ces communautés, et assume leurs contradictions ou leurs paradoxes. En marketing, la communauté est souvent employée comme synonyme du terme postmoderne de tribu. Pour ce courant (e.g. Maffesoli, 1988), le consommateur est à la recherche d'appartenance à des tribus pour deux raisons : (1) il rejette le progrès et la tribu correspond à une organisation sociale plus traditionnelle et (2) la

désagrégation des grandes institutions modernes lui fait chercher un réconfort dans la participation à une aventure collective.

Pour Maffesoli (1988 ; 2004), l'appartenance communautaire renvoie à une subjectivité commune, à une passion partagée. Cette approche est fondée sur la subjectivité qui s'oppose à la rationalité moderne qui a conduit au désenchantement du monde. Mais cette subjectivité n'est pas individuelle. Pour l'auteur (1998/2000), l'individu est le dernier avatar des substances qui servent de fondement aux analyses modernes et à ce titre doit disparaître. Dans la communauté, le soi disparaît au profit du soi collectif. L'émotif individualiste devenant l'émotionnel collectif et l'imaginaire du progrès se fond dans un présentéisme collectif. Ce présentéisme se retrouve dans l'aspect éphémère des groupes (Merle, 2003) qui se caractérisent par des rassemblements ponctuels suivis d'un éparpillement. Cette disparition de l'individu dans la masse collective est également un point sur lequel les théories de l'hypermodernité reviennent (e.g. Jauréguiberry, 2004 ; Lipovetsky 2006). Dans ce cadre au contraire c'est l'individu qui fait sens. Il ne disparaît pas, il se construit. Et au travers de son appartenance à de multiples communautés il « relève le défi de la gestion de son identité » (Jauréguiberry, 2004 ; p.165).

Les sites de microblogging et de réseautage social viennent questionner la notion de communauté et semblent également cristalliser le passage d'un consommateur postmoderne à un consommateur hypermoderne. Ils correspondent pourtant d'un premier abord à une illustration parfaite des tribus de Maffesoli (1998) et en particulier de leur plasticité (Sitz, 2008). Ils permettent aux consommateurs de former des communautés aux contours flous, dont la composition change perpétuellement. Plus encore que les communautés constituées autour de forums de discussion thématique ou de sites de marque, les groupes formés sur les sites de microblogging sont la plupart du temps éphémères, liés à l'émergence d'un centre d'intérêt commun qui précède l'éparpillement une fois que le thème a perdu son intérêt. A l'instar des flash mobs, des rencontres sportives, des concerts musicaux qui rassemblent ponctuellement des individus autour d'une activité ou passion commune, le microblogging peut rassembler des individus autour d'un sujet d'actualité ou d'une information pendant une durée limitée. Enfin, du fait que tous les regroupements entre individus sont possibles, toutes les communautés sont virtuellement présentes sur un site de réseautage social, allant dans le sens d'un sujet fragmenté et multiphrénique.

Pourtant, ce caractère éphémère des groupes, qui se matérialise par des individus se rassemblant autour d'un thème l'espace de quelques secondes, minutes, journées apparaît en contradiction avec la ritualisation postmoderne de ces regroupements (Cova, 2003). D'ailleurs, pour Sitz (2008), l'utilisation des réseaux sociaux pour l'étude des communautés conduit à négliger ce qui circule entre les individus et ainsi à manquer l'importance des rituels. Dans l'approche postmoderne, le sujet est décentré et se confond avec l'objet (Firat et Venkatesh, 1995). Ce qui fait sens dans la communauté, c'est l'éthos commun (Merle, 2003) : l'objet. Le rituel dans son acception postmoderne permet au consommateur de retrouver la notion du sacré qui a disparu dans le monde moderne désenchanté et participe d'un retour désiré à une société moins marchande. L'idée même d'une communauté est fondée sur une certaine notion de partage et d'échappatoire à la société de consommation (Kozinets, 2002b). Le rituel est utilisé comme moyen d'authentification des membres de la communauté (Badot et Cova, 2003). Le rituel est donc conforme à la perspective postmoderne : le sujet doit se fondre à la fois dans l'objet et dans le collectif. Le rituel n'est rien d'autre que le symbole de la transformation d'un individu en membre de la communauté. Ce qui prime c'est le groupe.

Sur le site de microblogging, un sujet peut attirer l'adhésion spontanée et ponctuelle de milliers d'individus. Il n'y a besoin d'aucun rituel, au mieux la marque de l'adhésion (le hashtag, cf. infra). L'entrée et la sortie sont totalement libres. Ceci semble cohérent avec la perspective hypermoderne pour laquelle la marchandisation de la société est totalement assumée et exacerbée. Il n'y a besoin d'aucun prétexte et d'aucun simulacre pour se donner l'impression d'échapper au marché. Ce qui prime ce n'est pas la thématique de la communauté. Ce qui donne sens aux communautés c'est l'individu. L'individu se produit lui-même, il construit son identité à travers son appartenance à divers groupes. Pour Jauréguiberry (2004), ce comportement renvoie à l'expérimentation des limites du *moi* par le sujet et provoque chez lui un nouveau questionnement sur ses limites identitaires. Dans cette perspective, l'appartenance à la communauté serait une ressource mobilisée par l'individu pour se mettre en récit et produire son identité.

Ceci ne remet nullement en cause le bien fondé du recours aux théories postmodernes dans l'étude du comportement du consommateur dans bien des domaines (notamment celui de la culture et des loisirs). Ceci étant dit, le microblogging semble bien être un lieu privilégié pour le consommateur hypermoderne. Dans les lignes qui suivent les aspects méthodologiques de

la collecte de données de microblogging sont décrits à travers la présentation du service le plus populaire de cette pratique : Twitter.

2 – ELEMENTS METHODOLOGIQUES

2.1 Présentation du microblogging et du service Twitter

Apparu en 2006 dans le paysage du microblogging, Twitter a récemment connu une forte progression de notoriété et d'utilisateurs. L'IFOP rapporte qu'en France la notoriété de ce service a été multipliée par 7 durant le premier semestre 2009 pour arriver à plus de 9 millions de français ayant entendu parler de Twitter et estime le nombre d'utilisateurs abonnés à 640 000 (Cann, 2009). Ce service gratuit de microblogging et de réseautage social propose à ses utilisateurs de diffuser du contenu textuel sous la contrainte que celui-ci ne dépasse pas 140 caractères (d'où le terme de microblogging). Ces courts messages sont dénommés « tweets » et peuvent être diffusés via un ordinateur ou un téléphone mobile. A l'instar des blogs, ce service permet donc à ses utilisateurs de s'exprimer auprès de leur « réseau ». Le positionnement initial de Twitter était de proposer un service permettant à ses utilisateurs de tenir en permanence leur entourage de leurs faits et gestes. L'utilisation a cependant considérablement évolué et permet à tout individu de diffuser des informations sur n'importe quel sujet, donner son avis sur l'actualité ou encore de confesser des idées ou réflexions.

Twitter permet également la construction d'un réseau social : pour lire les « tweets » d'un individu il faut le suivre, c'est-à-dire s'inscrire sur le réseau Twitter et s'abonner au flux que propose l'utilisateur. Contrairement au réseau Facebook qui nécessite un accord de l'auteur des contenus (statuts Facebook), sur Twitter, l'utilisateur n'a pas besoin de l'accord de l'émetteur des « tweets » pour les lire⁵. Enfin, l'utilisateur suivi reçoit un courrier électronique le notifiant que quelqu'un le suit. Il peut alors choisir de le suivre en retour, se retrouvant alors dans une relation mutuelle de suiveur-suivi et symétrique comparable à celles qui se nouent sur le réseau Facebook.

⁵ Par défaut, chaque utilisateur accepte que quiconque le suive, c'est-à-dire ait accès au contenu qu'il produit. Le service donne toutefois l'opportunité de paramétrer le compte de manière à ce qu'une demande soit acceptée pour le suivi, mais très peu d'utilisateurs ont recours à cette option. A posteriori, un utilisateur peut bloquer un autre utilisateur, l'empêchant ainsi de le suivre.

De ce fait, il est aisé de suivre plusieurs utilisateurs que l'on ne connaît pas. De la même manière, la procédure pour arrêter de les suivre se fait en un « clic » et ne nécessite aucun accord. Ceci a pour conséquence que, même si pour chaque utilisateur le contenu est archivé tout comme dans un blog, l'accès aux informations se fait en temps réel auprès de tous les utilisateurs suivis en même temps plutôt que par la consultation de la page « personnelle » de chacun. Le service se présente ainsi comme une page actualisée en permanence sur laquelle défilent tous les « tweets » produits par les individus que l'utilisateur a décidé de suivre. Cette particularité de focaliser l'utilisateur sur l'instant présent a rapidement fait de Twitter un concurrent redoutable de tous les services d'actualités. Twitter possède en effet un fort avantage vis-à-vis des sites traditionnels d'information : l'information est livrée en temps réel par les utilisateurs, de manière brute. Alors que d'autres doivent faire des choix éditoriaux pour choisir quelles informations méritent plus que les autres d'être communiquées aux lecteurs, Twitter les fournit toutes. Cet apparent fatras a poussé les utilisateurs à s'adapter afin de faciliter l'accès à cette masse d'information. Sont alors apparus les hashtags et les retweets.

Les hashtags sont des termes qui, à l'intérieur des tweets, sont précédés d'un symbole dièse (#) et associent le tweet à un concept plus général, à une catégorie, ce qui permet de faciliter le traitement de l'information. Ainsi la présence des hashtags #politique #cuisine etc permet immédiatement de contextualiser le message et d'en faciliter le traitement. Tout comme les tags constitutifs des folksonomies (Roxin et Bernard, 2007), ces hashtags peuvent être créés sur simple décision de l'utilisateur sans faire référence à une taxonomie « officielle » des sujets abordés sur Twitter. Ce faisant, ils proposent une thématique de discussion et deviennent le symbole d'une communauté qui se réunit ponctuellement autour d'un thème. Parfois certains thèmes apparaissent et disparaissent le temps d'un « événement » : #ns (nouvelle star), #vieillescharrues (festival des vieilles charrues) ou persistent. Dans ce dernier cas, on trouve des hashtags autour de thèmes généraux (#environnement, #emprunt, etc) ou de marques au sens large : produits et services (#iphone, #flickr, etc), partis politiques (#ps, #ump, etc.), artistes (#madonna, #mj : Michael Jackson, etc.).

Les retweets (RT) sont une information qui permet à l'utilisateur de citer sa source lorsqu'il reprend à son compte l'information d'un autre utilisateur de Twitter. Il précise alors qu'il n'est pas la première source de l'information en indiquant « RT @ » suivi du pseudonyme de l'utilisateur source : « RT @toto : le texte hadopi2 a finalement été voté ». Ce faisant, le RT

marque l'intérêt porté par un utilisateur au contenu d'un autre utilisateur et s'impose au lecteur comme le signe de la valeur informationnelle particulière du message.

Enfin, lorsqu'un individu s'adresse à un autre individu, s'il veut que son message soit lu par la personne à qui il est destiné, il faudra qu'il précise le pseudonyme de l'interlocuteur à l'intérieur même du tweet, précédé d'un arobase (e.g. « @alexandra : on se retrouve où pour aller à Dijon mercredi ? #jrmb14 »). La présence d'un arobase suivi d'un pseudonyme est donc le signe d'une interaction sociale

Ainsi, dans le flux d'informations qui défile sur son écran, l'utilisateur peut rapidement et visuellement repérer les thématiques (#), les messages importants (RT) et les discussions (@). Cette immédiateté de l'information brute, partiellement catégorisée et hiérarchisée par les retweets et les hashtags combinée à la fonction de réseau social du service Twitter paraît particulièrement intéressante pour étudier les processus sociaux à l'œuvre dans les groupes de consommateurs, les phénomènes de bouche à oreille et d'influence interpersonnelle sur l'internet ainsi que pour toute recherche s'intéressant aux discours des consommateurs.

2.2 Le recueil de données secondaires sur Twitter

Conformément à ses particularités techniques tant qu'aux éléments théoriques qu'il convoque, le microblogging – en particulier le service Twitter – offre de nouvelles possibilités quant au recueil des données qui y transitent.

Caractéristiques techniques : les API Twitter

La principale caractéristique du service Twitter, et probablement la principale raison de sa forte démocratisation, est qu'il propose deux API (Application Programming Interface) ou interfaces de programmation qui mettent à la disposition des utilisateurs l'ensemble des données présentes sur Twitter.

*** API REST**

Cette API fournit les données à un format XML (Extensible Markup Language) qui permet d'obtenir des données balisées, c'est-à-dire préformatées dans des champs arborescents dans afin de faciliter l'échange des données entre Twitter et les utilisateurs de l'API. Quoique cette

API soit principalement destinée à l'attention des développeurs qui désirent proposer aux consommateurs de nouvelles manières d'utiliser Twitter (logiciels spécifiques, téléphones portables, téléviseurs, panneaux d'affichage électroniques, widgets, etc.) cette API n'en demeure pas moins une formidable opportunité pour les chercheurs comme pour les praticiens désireux d'analyser les discours des consommateurs.

Pour chaque « tweet » posté sur Twitter, l'API fournit à ses utilisateurs tout un ensemble d'informations (cf. annexe 1) dont, entre autres, la date et l'heure du tweet, son contenu, le fait qu'il soit en réponse d'un autre tweet, l'identification de l'expéditeur du tweet, la date de création de son compte, le nombre de suiveurs, de suivis et de tweets qu'il comptabilise depuis la création de son compte. L'API permet également d'obtenir des renseignements sur les relations (suiveur, suivi) qu'entretiennent les utilisateurs de Twitter entre eux (cf. annexe 2). L'API fournit d'autres champs qui permettent de créer des programmes pour suivre automatiquement plusieurs personnes en même temps, arrêter de les suivre, les bloquer, etc. Ces fonctions, si elles font la force de ce service, ne présentent que peu d'intérêt pour le chercheur.

Enfin, Twitter impose des limites à l'utilisation de son API REST. Ces limites définissent le nombre de tweets envoyés par jour (1000), le nombre de courriers privés envoyés aux autres utilisateurs (250 « direct messages » / jour), et enfin – le plus important en ce qui nous concerne – le nombre maximum de requêtes dans l'API (150 / heure) avec une limite de 800 tweets collectés par requête (soit un maximum de $150 \times 800 = 120\,000$ / heure).

* API Search

Cette API fournit des données au format JSON (Java Script Object Notation). L'utilité de l'API Search est de permettre une sélection des contenus présents sur twitter sur n'importe quel mot clé. Aussi, dans l'exemple où le chercheur voudrait analyser uniquement les échanges contenant un (ou plusieurs) mot(s) précis (nom de marque par exemple), il ne serait pas obligé de collecter tous les tweets publiés pour ensuite sélectionner uniquement ceux contenant le(s) mot(s). Il collecte d'emblée uniquement les tweets qui correspondent à la requête. Cette API permet en outre de sélectionner les tweets en fonction de leur langage (si l'utilisateur a déclaré son langage) ou de la localisation de l'émetteur (cf. annexe 3).

Implications pour le chercheur

L'existence d'une API bouleverse le rapport que le chercheur peut avoir au terrain. Arnould et Wallendorf (1994) après avoir souligné l'importance des situations réelles préconisent la participation de longue durée dans l'univers de consommation, le croisement des sources de données et le fait que la procédure de collecte de données soit cohérente avec la dynamique du phénomène étudié et non pas fixée *a priori*. L'idée générale de cette approche est que le chercheur est lui-même l'instrument d'une « *grounded theory* » (Glaser et Strauss, 1967) qui s'élabore à mesure qu'elle émerge des données. Sur ce point, ce que nous appelons la « *twetnographie* » présente une particularité notable. Le site de microblogging représente le croisement de toutes les individualités et de toutes les communautés. Le flux d'information y est tel qu'il peut atteindre plusieurs milliers d'énoncés publiés par seconde. Le chercheur ne peut pas matériellement recopier cette vague informationnelle. Même si, comme le rappellent Galan et Vernet (2000) l'avantage des données sur internet est qu'elles sont déjà retranscrites et que le chercheur n'a plus qu'à les copier, cela n'est possible que dans le cadre de l'analyse d'une communauté délimitée comme c'est le cas dans les études netnographiques (Kozinets, 2002a). Dans le cadre du microblogging, c'est l'API qui va permettre l'accès aux données. Le chercheur fait des requêtes dans l'API pour tout récupérer. Cette collecte est considérable et en quelques semaines peut représenter des millions d'énoncés.

Il s'agit ensuite de faire ce que Kozinets (2002a, p.64) appelle le « grand tour » d'interprétation ce qui signifie coder les énoncés en quatre catégories selon que les messages sont sociaux ou informationnels, dans le sujet ou hors-sujet, le chercheur devant alors se concentrer sur les messages informationnels et dans le sujet. Pour faire ce « grand tour » comme pour analyser les énoncés les plus importants, l'approche ethnographique propose de réduire l'écart d'interprétation entre le chercheur et les données par la lecture répétée (Arnould et Wallendorf, 1994). Dans le cas des données de micro-blogging, le chercheur va devoir encore s'en remettre à l'API et au traitement automatisé. Nous avons vu plus haut que les tweets contenaient des informations sur le caractère social du message (présence d'un « @ » suivi d'un pseudonyme) qui peuvent servir de filtre dans les données collectées. Cependant cette présence ne signifie pas uniquement des interactions de courtoisie, il peut y avoir une véritable production d'informations lors de ces échanges. Il paraît donc important de croiser cette donnée avec une recherche sur mots clés. De la même manière, les hashtags sont pour leur part le signe d'une thématique. La présence d'un « # » suivi d'un terme est le

signe d'une thématique qui doit s'étendre au-delà de l'énoncé considéré pour en regrouper plusieurs. Cependant, étant donné le fait que tous les utilisateurs n'utilisent pas systématiquement, ni de la même manière les hashtags (Galan et Vignolles, 2010) il paraît nécessaire de croiser cette information avec une recherche sur des mots clés spécifiques à la thématique de recherche abordée. Si le chercheur peut faire la lecture répétée d'un certain nombre de tweets pour paramétrer son codage, la collecte du matériel à analyser ne peut être qu'automatisée.

Vient ensuite l'analyse des données sélectionnées, informationnelles et dans le sujet. Wallendorf et Belk (1989) suggèrent que la recherche ethnographique requiert un certain temps pour développer et comprendre un phénomène, un groupe ou une culture. Le temps nécessaire est variable et les auteurs préfèrent parler d'« engagement prolongé et d'observation persistante » (p.70). Il est nécessaire que le chercheur soit véritablement impliqué dans cette observation tout en gardant une distance critique afin de voir les choses avec des « yeux neufs » (p.70). On comprend que dans un contexte ethnographique en milieu « naturel », l'implication du chercheur peut rapidement connaître une érosion, l'amenant à progressivement se détacher du sujet et conduire des observations de moins en moins utiles. Dans le cadre de communautés virtuelles où l'observation se fait tout de même de manière active (aller sur le site ou le forum, lire les échanges, les recopier dans un fichier, etc.), le problème est le même. En revanche, en présence d'une API, le rapport à la durée de l'observation et au caractère soutenu de celle-ci est changé. Le chercheur conçoit un programme qui, une fois connecté à l'API, va aller chercher des données à sa place, pendant une durée infinie si besoin et de manière totalement exhaustive (car la diversité des comportements sur un site internet est tellement restreinte que la totalité de l'information peut être capturée) dans les limites fixées par Twitter (150 requêtes par heure). De ce fait, le chercheur peut laisser son programme collecter les données pendant des heures, des semaines, des années et analyser l'ensemble des observations *a posteriori*.

Enfin, la dernière particularité de la « twetnographie » concerne les considérations éthiques. En effet, les particularités du site de microblogging rendent quasi-impossible l'application des préconisations de Kozinets (2002a) que nous avons présentées dans la première partie de ce travail (p. 3). Trois de ces préconisations consistent dans la communication du chercheur vers la communauté afin de faire son entrée dans la communauté, chercher le feedback et obtenir une permission. Sur les sites de microblogging en général, et sur Twitter en particulier, il n'est

pas possible d'écrire à l'ensemble de la communauté. Cette fonctionnalité, par son absence, vise sûrement à protéger les utilisateurs des publicités intempestives (SPAM). Pour qu'un utilisateur puisse écrire à d'autres utilisateurs, il faut que ces utilisateurs soient des lecteurs de ses informations (suiveurs). Aussi, étant donné qu'il est impossible que le chercheur se fasse « suivre » par l'ensemble d'une communauté, il ne peut pas envoyer des informations à celle-ci. Cette impossibilité d'avertir les utilisateurs de ce qui va être fait des contenus qu'ils publient a encouragé Twitter à prendre les devants et à établir des conditions générales d'utilisation⁶ du service et de l'API qui encadrent l'utilisation des contenus. Les utilisateurs sont avertis qu'ils concèdent à Twitter et à tous ses partenaires le droit d'utiliser leur contenu sous toutes les formes. Twitter précise toutefois que les auteurs restent propriétaires de leur contenu. Ceux-ci ont le choix entre trois solutions : (1) ne rien faire, (2) s'inscrire dans une liste de comptes sous licence « Creative Commons », (3) préciser qu'ils ne veulent pas que leur contenu soit publié en dehors de leur réseau. La quasi-totalité des utilisateurs se trouve dans la première situation. L'utilisateur de l'API (ici le chercheur) doit s'en remettre alors aux instructions suivantes : (1) identifier l'émetteur du tweet sauf dans le cas où l'on publie des résultats sous forme agrégée ; (2) maintenir l'intégrité des tweets (ils peuvent être abrégés pour des raisons d'espace d'affichage) ; (3) demander l'autorisation à l'utilisateur avant de lui envoyer un message ; (4) obtenir la permission de l'utilisateur avant d'intégrer ses tweets dans un bien ou produit commercial (t-shirt, poster, etc.). L'instruction (1) vient donc à l'encontre de la préconisation de Kozinets (2002a) de maintenir l'anonymat des participants. Dans le cadre de Twitter, les conditions d'utilisations obligent à identifier l'utilisateur dès lors que l'on cite leur contenu. L'instruction (3) est inhérente au fonctionnement de Twitter : on ne peut envoyer un message sans être suivi par l'utilisateur ce qui est l'expression du consentement. L'instruction (4) ne concerne pas *a priori* la recherche.

Dans les lignes qui suivent nous allons exposer comment certaines recherches s'organisent en tenant compte de ces spécificités. Nous présenterons les principaux résultats, et sur la base de ce qui a déjà été fait, ainsi que sur ce qui a été présenté tout au long de cette communication, nous proposerons quelques voies de recherche possibles utilisant la twetnographie.

⁶ La première particularité de ces « terms of services » est qu'ils sont susceptibles de changer à n'importe quel moment. Aussi nous construisons notre propos sur la dernière version en date du 12/09/2009 (twitter.com/tos).

3 – ELEMENTS PRATIQUES

3.1 Etudes académiques en marketing réalisées sur les données issues de Twitter

Twitter étant un service récent (trois ans), il n'est pas surprenant qu'il existe peu d'études conduites sur les données qu'il fournit. Aucune revue marketing n'a encore, à notre connaissance, publiée un article sur le sujet. Cependant, dans les disciplines plus étroitement liées à l'informatique, il est possible de trouver des travaux qui ont de véritables implications marketing. Il s'agit des études réalisées par Java *et al.* (2007) et Jansen *et al.* (2009). Nous ajouterons à ces deux recherches un travail réalisé par nous-mêmes et qui demeure encore à publier (Galan et Vignolles, 2010). Le tableau 1 résume les caractéristiques de ces recherches.

Tableau 1 – Recherches utilisant Twitter comme source de données secondaires

	Java <i>et al.</i> (2007)	Jansen <i>et al.</i> (2009)	Galan et Vignolles (2010)
Question de recherche	Proposer des algorithmes pour détecter les communautés d'individus	Etudier le bouche à oreille concernant les marques	Proposer un indicateur d'identification des leaders d'opinion
Méthode de collecte	API : public_timeline (voir annexe 1 et 2)	Interface spécialisée « summarize4 »**	API : friends_timeline (annexe 1)
Etendue de la collecte	Tous les contenus postés par tous les utilisateurs du site	Tous les contenus postés par tous les utilisateurs, filtrés par noms de marque	Tous les contenus postés par 2 000 utilisateurs français
Durée de la collecte	8 semaines	13 semaines	9 semaines
Nombre d'énoncés (tweets) collectés	1 348 543	149 472	150 295
Nombre d'individus à l'origine des énoncés	76 177	Non calculé par les auteurs	2 000

** équivalent aujourd'hui à l'utilisation de l'API Search (cf. annexe 3)

Accès aux données

Avant d'exposer plus en détails les résultats de ces travaux, il est intéressant de noter qu'ils emploient des méthodes différentes quant à la collecte des données et à son étendue par rapport à l'ensemble des énoncés produits sur Twitter. L'article de Java *et al.* (2009) est celui qui collecte de manière la plus exhaustive. Durant deux mois, tous les tweets postés sur Twitter par l'ensemble des utilisateurs n'ayant pas bloqué l'accès public (la quasi-totalité) à leur contenu sont collectés (méthode statuses public_timeline ; cf. annexe 1). La recherche a

été conduite en avril 2007 ce qui explique sa faisabilité car les auteurs rapportent que seuls 76 177 utilisateurs différents ont été repérés lors de cette collecte. Sur environ 8 semaines ils ont collecté 1,3 millions de tweets à raison d'une requête dans l'API toutes les 30 secondes (soit 120 requêtes par heure). Actuellement cette méthode serait plus complexe à mettre en œuvre. En effet, il y a aujourd'hui des millions d'utilisateurs de Twitter publiant plusieurs milliers de tweets par seconde. Etant donné que le maximum de données que peut collecter un programme via l'API Twitter (120 000/ heure soit 33 tweets / seconde), il faudrait accéder à l'API de plusieurs comptes en même temps pour capturer l'ensemble des publications.

Les recherches plus récentes se caractérisent par un filtrage du contenu à collecter afin d'éviter les problèmes techniques liés à la surcharge d'information. Jansen *et al.* (2009) utilisent un site spécialisé. Ce service dénommé Summize a été racheté par Twitter courant 2008 et se trouve intégré directement dans l'API (Summize est devenu l'API Search, ce qui explique probablement que les données soient dans un format différent – JSON – que dans l'API REST qui fournit les données en XML). Jansen *et al.* (2009) établissent tout d'abord une liste de 50 marques et ne collectent que les tweets en langue anglaise qui contiennent au moins une de ces marques. Ceci leur permet de réduire le nombre de tweets à 149 472 sur une période comparable à celle de Java *et al.* (2009) à une date pourtant plus récente. Pour notre part, nous avons choisi de passer par l'API REST et d'utiliser la méthode `friends_timeline` (Galan et Vignolles, 2010). Ceci permet de choisir non pas les sujets des tweets mais leurs sources. Nous avons tout d'abord sélectionné 2000 personnes selon des critères précis (pas d'entreprises, pas de célébrités, langue française uniquement, localisation en France). Cela nous permet d'arriver à 150 295 tweets sur une période de 9 semaines.

Réseaux sociaux et structure des communautés

Java *et al.* (2007) sur la base d'une collecte exhaustive de tout le contenu publié sur Twitter établissent l'existence de 76177 utilisateurs actifs différents (principalement localisés en Amérique du Nord). En utilisant l'algorithme HITS (Kleinberg, 1999), ils repèrent les connecteurs (hubs) et les utilisateurs ayant de l'autorité. Ils définissent les connecteurs comme étant des personnes ayant beaucoup d'amis (c'est-à-dire qu'ils suivent beaucoup de personnes les suivant réciproquement). Ils définissent les utilisateurs ayant de l'autorité comme ayant beaucoup de suiveurs mais ne suivant que peu d'individus. Les auteurs établissent les relations (suiveur, suivi, ami) de chacun des utilisateurs actifs et parviennent à un ensemble de

829 053 relations, regroupées autour de 87 897 « nœuds » distincts (c'est-à-dire sans aucune relation d'un nœud à l'autre). Pour les auteurs la communauté est définie par un regroupement d'individus et non par des thématiques (ce qui est cohérent avec l'approche que nous avons développée plus haut même si les auteurs ne font aucune référence à la postmodernité ou à l'hypermodernité ni à un quelconque courant de sciences humaines ou sociales). L'intérêt de cette recherche est que les auteurs vont au-delà de la notion de communauté indépendante qui est communément admise dans leur discipline pour reconnaître au consommateur la possibilité d'appartenir à plusieurs communautés simultanément. Sur la base d'un algorithme dénommé CPM (Clique Percolation Method ; Palla *et al.*, 2005), les auteurs matérialisent automatiquement les liens qui existent entre les membres de communautés différentes. Ils observent qu'un utilisateur peut avoir des fonctions tout à fait différentes d'une communauté à l'autre (être une source d'information dans une, être à la recherche d'information dans une autre). Sur la base des communautés d'individus, ils font la liste des mots clés utilisés dans chaque communauté. Ils remarquent en comparant deux communautés qui ont des thématiques similaires que le degré d'usage des termes est différent d'une communauté à l'autre et que, dans un cas ce qui est abordé de manière quasi professionnelle dans une communauté, est traité de manière très informelle dans l'autre. Ils terminent en notant que « l'intention » d'utiliser Twitter est partagée par les membres d'une communauté. Ils distinguent à ce titre quatre types d'intention : parler de sa vie quotidienne, discuter, partager des informations, se faire le relais de l'actualité. Ceci leur permet d'arriver à trois types de membres dans la communauté : la source d'information, l'ami, le chercheur d'information. On voit bien que les implications marketing ne sont pas au centre des préoccupations des auteurs, cependant, les algorithmes qu'ils proposent pour établir le degré d'autorité ou de connectivité des utilisateurs ainsi que pour établir les contours d'une communauté sont riches d'implications et ouvrent des voies de recherche prometteuses (cf. infra).

Bouche-à-oreille sur les marques

Jansen *et al.* (2009) étudient le microblogging comme nouvelle forme de bouche à oreille sur les marques. Les auteurs ont collecté durant 13 semaines tous les tweets de langue anglaise ayant mentionné l'une des 50 marques que les auteurs avaient présélectionnées dans 12 secteurs d'activité (prêt-à-porter, automobile, informatique, alimentaire, etc.), soit 149 472 énoncés. Les auteurs se sont intéressés spécialement aux énoncés mentionnant une des marques ainsi que l'expression d'une opinion ou d'un sentiment sur cette marque. L'outil

sumimize4 utilisé par les auteurs – disparu aujourd’hui suite à son rachat par Twitter – permettait à travers l’analyse des mots associés à la marque de donner un sentiment global du tweet sur une échelle en cinq points : lamentable, mauvais, bof, super et génial (wretched, bad, so-so, swell, great). Il ressort des résultats que plus de 60% des opinions sur les marques sont favorables et 12% neutres. Les résultats suggèrent également que les consommateurs dont l’évaluation est extrême (génial ou lamentable) sont ceux qui s’expriment le plus. Ces productions auraient pour objectif de gagner l’approbation des autres et partager son expertise (pour les évaluations positives) ou sont motivés par l’hostilité à l’égard de la marque et la vengeance. Les auteurs finissent par se focaliser sur les relations que peut entretenir une marque (Starbucks) avec ses consommateurs sur Twitter. Sur la période considérée, Starbucks a reçu 1585 tweets de la part de 1 038 utilisateurs sur les 7 779 que la marque suit. Réciproquement, la marque Starbucks a publié 322 tweets dont 245 à l’attention d’utilisateurs particuliers (les 77 tweets restants étant d’ordre plus général). Compte tenu des 7 751 suiveurs et des 7 779 suivis, les auteurs expliquent que le nombre de communications est plutôt faible et suggèrent qu’à l’instar d’autres médias électroniques (listserv et wiki) le nombre de membres actifs est faible, la majorité étant des « lurkers » (individus ne faisant que lire les échanges sans jamais participer). Les auteurs concluent leur travail sur le constat que 20% des tweets contiennent des noms de marque ce qui montre que Twitter peut servir de support pour les campagnes de marketing viral, le CRM et influence le bouche-à-oreille. En outre cela doit attirer les marques sur le fait que ces 20% d’énoncés peuvent affecter l’image et la notoriété de la marque et que les entreprises doivent tenter d’intégrer cela dans leur stratégie.

Leadership d’opinion

Durant 8 semaines, nous avons suivi 2000 personnes afin de proposer un indicateur susceptible de permettre l’identification des leaders d’opinions parmi les utilisateurs du service Twitter (Galan et Vignolles, 2010). Après avoir suggéré que le nombre de suiveurs débouche en retour sur un grand nombre de suivis et correspond mieux à l’identification de connecteurs que d’influenceurs, les influenceurs sont définis comme ayant beaucoup de suiveurs et peu de suivis, et plus particulièrement, ayant un rapport suiveurs/suivis (S/s) élevé. Une étude de régression confirme que ce ratio est assez peu influencé par les relations « mécaniques » provoquées par l’utilisation de Twitter et qu’il est en outre un facteur explicatif du nombre de retweets (tweets circulant d’un individu à l’autre). Afin de vérifier la pertinence du ratio S/s comme indicateur de l’influence des utilisateurs, trois groupes de

trente individus ont été créés : les influenceurs (grand nombre de suiveurs et ratio $S/s > 3$), les connecteurs (grand nombre d'amis, ratio proche de 1), les suiveurs (grand nombre de suivis, ratio $< 0,3$). Sur la base de 9254 tweets publiés par ces 90 utilisateurs, il est apparu que les groupes étaient significativement différents dans leur utilisation des hashtags, des liens, du caractère social des tweets et de l'interface utilisé pour accéder au service. Les principaux résultats à cette étape sont que les influenceurs se caractérisent par une plus faible utilisation des hashtags et que les connecteurs sont caractérisés par de plus nombreuses interactions sociales. Une étude sur le contenu des tweets « sociaux » produits par chaque groupe montre que les influenceurs produisent beaucoup plus de réponses que de questions et inversement pour les suiveurs. Enfin, une étude des tweets informationnels suggère que les influenceurs donnent plus facilement leur avis sur leur environnement politique et juridique et moins sur leur vie intime, à l'inverse des suiveurs. Tout ceci amène à conclure que le ratio suiveurs / suivis, particulièrement facile à obtenir et à calculer, représente un indicateur pertinent du leadership sur Twitter qui permet des économies de temps et qui est moins coûteux par rapport aux autres méthodes d'identification du leadership d'opinion par les méthodes basées sur les données primaires par exemple.

3.2 Voies de recherches ouvertes par l'accès aux données de Twitter

Influence interpersonnelle

Les considérations liées à l'influence interpersonnelle (e.g. leadership d'opinion) ont beaucoup d'implications en marketing. Il apparaît que les données de Twitter permettent de fournir des informations précieuses sur les liens qui existent entre utilisateurs et autorisent ainsi un nouveau type d'approche pour l'étude de l'influence. A ce titre, il paraît nécessaire de continuer les travaux déjà commencés et comparer par exemple l'approche basée sur l'algorithme HITS (Java *et al.*, 2009) avec celle du simple ratio suiveurs/suivis (Galan et Vignolles, 2010) et explorer la convergence de ces approches avec des approches plus traditionnelles de mesure (e.g. Fejlaoui et Vernet, 2009).

Etudes des mondes de marque.

Le chercheur a le choix dans son approche des données pour établir l'existence d'un groupe de consommateurs. Soit il le fait par le thème autour duquel s'est construite la communauté

(forum de discussion, newsgroup), soit pas l'identification des liens qui existent entre les individus d'une population. Les études présentées plus haut offrent à ce titre des méthodologies permettant de suivre les deux « chemins » : de la thématique vers les membres (Jansen *et al.*, 2009) ; des relations entre les membres vers la thématique (Java *et al.*, 2009). Ici également, une étude de convergence, ou tout au moins une comparaison des méthodes peut être fertile pour la recherche⁷. En effet, outre la comparaison des résultats, elle offre des perspectives théoriques permettant d'expliquer la relation de l'individu au collectif, ceci éventuellement en ayant recours aux courants postmoderne et hypermoderne. La notion de monde de marque que Sitz (2008) développe semble adéquate aux regroupements sur Twitter puisque d'après l'auteur, le monde de marque, qui associe les réseaux sociaux et les systèmes culturels, est un monde social qui résulte « d'un flux d'interactions entre des acteurs qui en viennent à partager des codes, des répertoires d'interprétation et des représentations collectives » (p. 28)

Retweets

Dans l'étude conduite sur les leaders d'opinion (Galan et Vignolles, 2010) nous avons volontairement éludé l'analyse de contenu des retweets qui sont les messages qui présentent le plus d'intérêt pour la communauté. Une voie de recherche importante est l'analyse de ces messages afin d'essayer de comprendre ce qui fait sens dans la communauté Twitter et d'établir la relation particulière qu'entretiennent les influenceurs avec ces messages. Est-ce que le message est digne d'intérêt parce qu'il provient de l'influenceur ou est-ce qu'au contraire c'est le statut d'influenceur qui découle de la capacité de l'individu à découvrir un contenu intéressant qu'il utilise ensuite comme ressource pour établir son influence ?

Marques

Jusqu'à présent, et à l'exception du seul cas de la marque Starbucks (Jansen *et al.*, 2009), les études académiques se sont focalisées sur les discussions sur Twitter sans chercher à s'intéresser en particulier aux énoncés émis par les marques à l'attention de leurs publics. Twitter est pourtant un support de communication qui est déjà investi par les entreprises. Il paraît dès lors important de s'intéresser à ce type de communication pour en comprendre les

⁷ A ce titre l'analyse des réseaux sociaux (ARS) fournit au marketing un ensemble d'outils particulièrement utiles pour la collecte, la mesure et l'analyse (voir Mercanti-Guérin 2009 pour un exposé détaillé).

spécificités et les implications. Ainsi, en prenant la méthode de collecte `friends_timeline` (Galan et Vignolles, 2010) on peut ne collecter que les tweets publiés par des entreprises qui auront été préselectionnées. Ce type de données peut-être particulièrement utile pour étudier les différentes stratégies (communication, CRM, etc.) selon les secteurs d'activité et le type d'entreprises et voir la capacité de la communication sur ce support à générer du trafic sur les sites d'entreprise ou encore à provoquer des achats.

Conclusion

Les services de microblogging sont en perpétuelle expansion et attirent chaque jour de nouveaux utilisateurs. Le site Twitter en particulier, dont la forte croissance en 2009 lui permet aujourd'hui d'obtenir plus de 20 millions de visiteurs uniques chaque mois (source *compete*⁸), progresse également en France où il compte déjà 640 000 utilisateurs (Cann, 2009) alors que l'interface en langue française n'existe pas encore. En tant que réseau social mais également comme site de microblogging, Twitter fournit des informations très précieuses, sur les relations interpersonnelles (communautés de marque, réseaux de consommateurs, relations entreprises-marques, etc.) ainsi que sur les contenus échangés (expériences de consommation, prescription, communication commerciale, etc.), qui sont gratuitement et aisément accessibles grâce à l'existence de deux interfaces de programmation (API).

Nous avons abordé ici les éléments théoriques, méthodologiques et pratiques afin de présenter la *twetnographie* comme une méthode d'enquête sur Twitter à l'attention des chercheurs et praticiens en marketing. La *twetnographie* vient en renfort des études qualitatives pour faire face aux défis posés par les bouleversements que connaît la société de consommation. Cette méthode, même si elle présente des particularités qui peuvent amener le chercheur à s'éloigner parfois du terrain (lors de la collecte) se veut fondée sur la démarche de l'ethnomarketing (Badot *et al.*, 2009) en s'intéressant avant tout à la dimension culturelle de la consommation et aux significations que les consommateurs projettent sur leur réalité d'une part et faisant du chercheur le principal instrument de l'analyse (la collecte étant automatisée) d'autre part.

⁸ <http://siteanalytics.compete.com/twitter.com/?metric=uv>

Références bibliographiques

Armstrong, A., Hagel, J. III (1996), The real value of on-line communities, *Harvard Business Review*, 74, 3, 134-41.

Arnould E. J., Wallendorf M. (1994), Ethnography: Interpretation Building and Marketing Strategy Formulation, *Journal of Marketing Research*, 31, 4, 484-504.

Ayadi N. (2005), Etude du comportement de prise de risque dans le cadre de l'expérience de consommation du parachutisme : Apports de la netnographie, *10^{èmes} Journées de Recherche en Marketing de Bourgogne*, Dijon, Novembre 2005.

Babin O., Galan J.-Ph. (1998), L'internet : outil et objet de la recherche en marketing, *Actes de la 1ère Journée Marketing et Communication de l'AFM/ESUG*, Toulouse, 89-125.

Badot O., Carrier Ch., Cova B., Desjeux D., Filser M. (2009), L'ethnomarketing : un élargissement de la recherche en comportement du consommateur à l'ethnologie, *Recherche et Applications en Marketing*, 24, 1, 93-111.

Badot O., Cova B. (2003), Néo-marketing, 10 ans après : pour une théorie critique de la consommation et du marketing réenchantés, *Revue Française du Marketing*, 195, 5, 79-94.

Bernard Y. (2004), La netnographie : Une nouvelle méthode d'enquête qualitative basée sur les communautés virtuelles de consommation, *Décisions Marketing*, 36, 49-62.

Bertrandias L. Carricano M. (2006), Le prix derrière l'écran : un étude netnographique d'une communauté virtuelle professionnelle de Pricing Managers, *Décisions Marketing*, 41, 53-66.

Cann Y.M. (2009), Twitter, beaucoup de bruit pour rien, *Observatoire des Réseaux Sociaux*, été 2009, <http://bit.ly/tRqHQ>

Carù A., Cova B. (2008), Le consommateur interpassif : nouvel avatar postmoderne ?, *7th International Marketing Trends Conference*, Venise, CD-ROM.

Cova B. (2003), Pourquoi parler de tribus qui consomment ?, in E. Rémy, I. Garabuau-Moussaoui, D. Desjeux, et M. Filser (eds.), *Sociétés, consommation et consommateurs*, L'Harmattan, Paris, 69-84.

Cova B., Carrère V. (2002), Les communautés de passionnés de marque : opportunité ou menace sur le net ? *Revue Française du Marketing*, 189/190, 119-130.

Diesbach B. P., Galan J.-Ph., Chandon J.-L (2006), Impact de la présence d'un agent virtuel incarné sur le pouvoir de rétention du site web : une analyse comportementale par les fichiers log, *11èmes Journées de Recherche en Marketing de Bourgogne*, Dijon, CD-ROM.

Fejlaoui Y., Verneette E. (2009), Repérer les leaders d'opinion dans les communautés virtuelles de consommation : une nouvelle échelle de mesure, *Actes du XXVème Congrès International de l'Association Française du Marketing*, Londres, CD-ROM.

Ferrandi J.-M., Boutin E. (1999), Un outil de mesure de l'audience d'un site Internet : l'analyse réseau, *Actes du XVème Congrès International de l'Association Française du Marketing*, Strasbourg, 669-696 .

Firat A. F., Venkatesh A. (1995), Liberatory postmodernism and the reenchantment of consumption, *Journal of Consumer Research*, 22, 3, 239-267.

Florès L. (2008), Web 2.0 : Des études ayant du répondant!, *Décisions Marketing*, 50, 79-82.

Galan J.-Ph. (2002), L'analyse des fichiers log pour étudier l'impact de la musique sur le comportement des visiteurs d'un site Web culturel, *Actes du XVIIIème Congrès International de l'Association Française du Marketing*, Deauville, 255-272.

Galan J.-Ph., Verneette E. (2000), Vers une 4ème génération : les études de marché "on-line", *Décisions Marketing*, 19, 39-52.

Galan J.-Ph., Vignolles V. (2010), Identification des leaders d'opinion sur internet : utilisation des données secondaires issues de Twitter, *9th International Marketing Trends Conference*, Venise, à venir.

Glaser B.G., Strauss A. (1967), *The Discovery of Grounded Theory: Strategies for Qualitative Researches*, Chicago, Aldine.

Godes D. Mayzlin D. (2004), Using Online Conversations to Study Word-of-Mouth Communication, *Marketing Science*, 23, 4, 545-560.

Hoffman D.L., Novak T.P. (1996), Marketing in Hypermedia Computer-mediated Environments: Conceptual Foundations, *Journal of Marketing*, 60, 3, 50-68.

Jauréguiberry F. (2004), Hypermodernité et manipulation de soi, in N. Auber (éd.) *L'individu hypermoderne*, Toulouse, Erès, 155-168.

Java, A., Finin, T., Song, X., Tseng B. (2007), Why we Twitter: Understanding Microblogging Usage and Communities, *Proceedings of the 9th Web Knowledge Discovery in Databases Workshop*, <http://bit.ly/3pO1Iz>

Jansen B. J., Zhang M., Sobel K., Chowdury A. (2009), Twitter Power: Tweets as Electronic Word of Mouth, *Journal of the American Society for Information Science and Technology*, 60, 9, 1-20.

Kleinberg J. M. (1999), Authoritative Sources in a Hyperlinked Environment, *Journal of the Association of Computing Machinery*, 46, 5, 604-632.

Kozinets R. V. (1997), "I want to believe": A Netnography of the X-Philes' Subculture of Consumption, in M. Brucks and D. J. MacInnis (eds.) *Advances in Consumer Research*, Provo, UT: Association for Consumer Research, 24, 470-475.

Kozinets R. V. (2002a), The Field behind the Screen: Using Netnography for Marketing Research in Online Communities, *Journal of Marketing Research*, 39, 1, 61-72.

Kozinets R. V. (2002b), Can Consumers Escape the Market ? Emancipatory Illuminations from Burning Man, *Journal of Consumer Research*, 29, 1, 20-38.

Lipovetsky G. (2006), *Le bonheur paradoxal : essai sur la société d'hyperconsommation*, Paris, Gallimard.

Lipovetsky G., Charles S. (2004), *Les temps hypermodernes*, Paris, Bernard Grasset.

Maffesoli M. (1988), *Le temps des tribus : le déclin de l'individualisme dans les sociétés de masse*, Paris, La Table Ronde, 3^{ème} édition (2000).

Maffesoli M. (2004), *Le rythme de la vie : variations sur les sensibilités postmodernes*, Paris, La Table Ronde.

Mercanti-Guérin M. (2009), Analyse des réseaux sociaux et communautés en ligne : quelles applications en marketing ? 8^{ème} Journée de Recherche sur le e-Marketing, Paris, CD-ROM.

Merle A. (2003), Apports du point de vue postmoderne à l'étude des comportements de consommation de groupe : un début de clarification, 2^{èmes} Journées Normandes de Recherche sur la Consommation, IAE Caen, CD-ROM.

Palla G., Derenyi I, Farkas I., Vicsek T. (2005) Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society, *Nature*, 435, 814-818.

Roxin V., Bernard Y. (2007), Etiquetage collaboratif et nuages de mots : quels apports pour les sites marchands, 6^{ème} Journée Nantaise de Recherche sur le e-Marketing, <http://bit.ly/4sWtev>

Sivadas E., Grewal R., Kellaris J. J. (1998), The Internet as a Micro Marketing Tool: Targeting Consumers through Preferences Revealed in Music Newsgroup Usage, *Journal of Business Research*, 41, 3, 179-186.

Sitz L. (2008), Les mondes de marques : l'exemple du monde Apple, *Décisions Marketing*, 52, 19-32.

Vuitton T. (2009), La course aux « followers », *Le Monde*, 12 septembre 2009.

Wallendorf M., Belk R. W. (1989), Assessing Trustworthiness in Naturalistic Consumer Research, in E.C. Hirschman (ed.) *Special Volumes – Interpretative Consumer Research*, Provo, UT: Association for Consumer Research, 69-84.

Annexe 1 : API REST – Méthode : statuses public_timeline

(la méthode statuses_friends_timeline conduit au même résultat mais n'affiche que les utilisateurs que l'individu à l'origine de la requête suit)

```

<status>
<created_at>Tue Nov 11 22:52:51 +0000 2009</created_at>
<id>147266936123123</id>
<text>@alexandra : on se retrouve où pour aller à Dijon mercredi ? #jrmb14 </text>
<source>&lt;a href="http://www.tweetdeck.com/">TweetDeck&lt;/a></source>
<truncated>>false</truncated>
<in_reply_to_status_id>147266936123987</in_reply_to_status_id>
<in_reply_to_user_id>9876543210</in_reply_to_user_id>
<favorited>>false</favorited>
<in_reply_to_screen_name>alexandra</in_reply_to_screen_name>
<user>
  <id>0123456789</id>
  <name>Jean-Philippe Galan</name>
  <screen_name>jphgalan</screen_name>
  <location>Toulouse</location>
  <description>EC en marketing.</description>
  <profile_image_url>http://s3.../avatar_normal.png</profile_image_url>
  <url>http://www.iae-toulouse.fr</url>
  <protected>>false</protected>
  <followers_count>1027</followers_count>
  <profile_background_color>9ae4e8</profile_background_color>
  <profile_text_color>000000</profile_text_color>
  <profile_link_color>0000ff</profile_link_color>
  <profile_sidebar_fill_color>e0ff92</profile_sidebar_fill_color>
  <profile_sidebar_border_color>87bc44</profile_sidebar_border_color>
  <friends_count>293</friends_count>
  <created_at>Sun Mar 18 06:42:26 +0000 2007</created_at>
  <favourites_count>0</favourites_count>
  <utc_offset>-18000</utc_offset>
  <time_zone>Paris (France)</time_zone>
  <profile_background_image_url>http://s3...twitter_bg_grass.jpg</profile_background_image_url>
  <profile_background_tile>>false</profile_background_tile>
  <statuses_count>3390</statuses_count>
  <notifications>>false</notifications>
  <following>>false</following>
  <verified>>true</verified>
</user>
</status>

```

Annexe 2 : API REST – Méthode : friends ids

Cette méthode renvoie, pour un utilisateur donné, tous les individus – par leur numéro d'identification (voir user_id en annexe 1) – que l'utilisateur suit. La méthode followers ids conduit au même résultat mais affiche à la place les individus qui suivent l'utilisateur.

```
<id_list>
  <ids>
    <id>30592818</id>
    <id>21249843</id>
    (...)
  </ids>
  <next_cursor>1288724293877798413</next_cursor>
  <previous_cursor>-1300794057949944903</previous_cursor>
</id_list>
```

Annexe 3 : API Search – méthode search

```
{ "results": [
  { "text": "@twitterapi http://tinyurl.com/ctrefg",
    "to_user_id": 396524,
    "to_user": "TwitterAPI",
    "from_user": "jkoum",
    "id": 1478555574,
    "from_user_id": 1833773,
    "iso_language_code": "nl",
    "source": "<a href='http://twitter.com'>twitter</a>",
    "profile_image_url": "http://s3.../b_normal.jpg",
    "created_at": "Wed, 08 Apr 2009 19:22:10 +0000" },
  (...),
  "since_id": 0,
  "max_id": 1480307926,
  "refresh_url": "?since_id=1480307926&q=%40twitterapi",
  "results_per_page": 15,
  "next_page": "?page=2&max_id=1480307926&q=%40twitterapi",
  "completed_in": 0.031704,
  "page": 1,
  "query": "%40twitterapi" }
```

