Vers une approche instrumentale de l'identité numérique : les attributs identitaires comme structuration de l'environnement informationnel ?

Camille Alloing

▶ **To cite this version:**

Camille Alloing. Vers une approche instrumentale de l'identité numérique : les attributs identitaires comme structuration de l'environnement informationnel ?. Hermès Sciences. Enseignement, préservation et diffusion des identités numériques, Lavoisier, 2014, Traité des sciences et techniques de l'information. hal-01135911

HAL Id: hal-01135911
https://hal.science/hal-01135911

Submitted on 26 Mar 2015

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.

Vers une approche instrumentale de l'identité numérique : les attributs identitaires comme structuration de l'environnement informationnel ?

Camille Alloing – CEREGE (EA 1722)

Ceci est le « pré-print » du chapitre d'ouvrage. Pour citer le texte, merci de prendre la référence :

ALLOING, C., « Vers une approche instrumentale de l'identité numérique : les attributs identitaires comme structuration de l'environnement informationnel ? », in PINTE, J-P. (Dir.), *Enseignement, préservation et diffusion des identités numériques*, Hermes Lavoisier, Avril 2014, pp. 39-68.

2.1. Introduction

Le développement d'Internet d'un point de vue technique (haut-débit, fibre optique), technologique (interfaces, protocoles dédiés), mais aussi de l'accès (une majorité des foyers français étant aujourd'hui connectés[1]) a permis à ce « réseau des réseaux » d'intégrer peu à peu le quotidien des individus, mais aussi des organisations. Source d'informations pour diverses tâches de la vie quotidienne (par exemple chercher l'adresse d'un restaurant), Internet s'est progressivement mué en outil communicationnel à part entière sous l'impulsion du développement des interfaces Web (au-delà des premiers protocoles mails), communément qualifiées de « 2.0 » [MUS 06], tels les blogs, les wikis ou encore les « réseaux socionumériques » [STE 10] (Facebook, Twitter, Myspace, etc.), et leur appropriation grandissante par les utilisateurs d'Internet (les « internautes »).

Evoluant vers une forme d'économie de la contribution [STI 09], le Web dit 2.0 offre la possibilité à l'internaute de s'affranchir des règles communément établies de production et de diffusion de l'information, développant alors des pratiques qualifiées par E. Broudoux [BRO 03] « d'autoritatives », soit : « l'attitude consistant à produire et à rendre public des textes […] sur le www, sans passer par l'assentiment d'institutions de référence référées à l'ordre imprimé ». Comme nous l'aborderons par la suite, cette capacité à diffuser du contenu en s'affranchissant des traditionnels *gate keepers* [CAR 10] pose question à différents niveaux :

– celui des consommateurs d'informations sur ces plates-formes se trouvant à la recherche de points de repère pour évaluer la pertinence ou la qualité des informations diffusées ;

– celui des organisations et plus précisément des entreprises souhaitant collecter de l'information pour l'intégrer à leurs stratégies de communication.

Car « là où la culture de masse fabriquait des dénominateurs communs, la culture numérique tend par essence à personnaliser l'information » [MER 09a], personnalisation redéfinissant notamment les possibles interprétations d'un même document ou d'une information diffusée sur le Web.

Dans ce contexte où « *tout contenu publié devient lui-même prescriptif, sur le mode des préférences partagées* » [MER 09, b] (préférences issues notamment des traces laissées par les internautes et constitutives de leur identité numérique [ERT 09]), les organisations en quête d'informations sur certaines thématiques (environnement concurrentiel, évolution de certains marchés ou encore réputation)

1. « Un peu plus de 64 % des ménages déclarent avoir un accès à Internet à leur domicile en 2010 » selon l'INSEE (www.insee.fr/fr/themes/document.asp?ref_id=ip1340#inter1, consulté le 9 juin 2012).

doivent alors se questionner sur la manière d'intégrer cette constante réarticulation sémiotique des documents qu'elles collectent. Intégration aux stratégies de « veille stratégique », définie comme « le processus par lequel un individu […] traque, de façon volontariste, et utilise des informations à caractère anticipatif » [LES 02], pouvant alors passer par l'accaparation d'une partie de l'économie de la recommandation [PIE 11] en cours sur ces réseaux ou plates-formes : identifier les usages prescriptifs afin de définir un contexte stable d'observation, mais aussi identifier des signaux pouvant appuyer une prise de décision. Comme le souligne J. Pierre : « la valeur de la recommandation est cependant associée à son origine, et plusieurs niveaux se distinguent : l'algorithme de computation, façon Google ou Amazon, puis l'adjonction sociale, mais anonyme d'un commentaire et, dernier niveau récemment atteint, la recommandation par une personne identifiée au sein de notre réseau social » [PIE 11].

Dans ce chapitre, nous proposons de nous concentrer sur le « dernier niveau » de la recommandation, en proposant une réflexion sur la manière dont une entreprise peut « identifier » une source de recommandation/prescription d'informations et l'intégrer à son propre « réseau social »[2]. Plus précisément, nous nous interrogerons sur le fait de savoir en quoi les données/attributs identitaires en ligne permettent de mieux saisir le contexte d'émission des informations afin de définir une forme d'autorité non algorithmique. Et ce, par une recherche-action dont voici les contours.

2.2. Détection d'agents-facilitateurs et intégration au processus de veille : le cas de La Poste Courrier

Les questionnements et constats présentés en introduction sont en partie issus d'une recherche doctorale menée dans le cadre d'un contrat CIFRE au sein de la société La Poste métier Courrier, et plus précisément à la direction du Système d'Information[3]. L'un des objectifs premiers de cette recherche fut de définir des indicateurs de mesure de « l'e-réputation » des diverses innovations technologiques numériques produites par la direction de rattachement. Concept d'e-réputation que nous pouvons définir en première approche comme procédant « d'une redocumentarisation itérative automatisée ou endogène des interprétations et connaissances produites en ligne par les individus ou les organisations, forgeant un attribut identitaire servant de marqueur prescriptif destiné à cadrer dans un contexte asynchrone et atopique les formes d'engagement ultérieures, et facilitée par les

2. Il serait plus juste de parler de « réseau informationnel » ou de « réseau de sources », tel que l'entend l'intelligence économique (voir [MAR 07]).

3 Thèse en sciences de l'information et de la communication intitulée *Processus de veille par infomédiation sociale pour construire l'e-réputation d'une organisation. Approche par agents-facilitateurs appliquée à la DSIC de La Poste*, et soutenue le 2 juillet 2013.

fonctionnalités techniques des sites Web où elle opère une informatisation du microsocial au profit et avec les moyens de la logique marchande » [ALL 12a]. L'objectif global étant alors d'identifier en quoi la réarticulation constante des connaissances diffusées sur un sujet (en l'occurrence les innovations postales) peut d'une part, impacter la diffusion de celui-ci et d'autre part, définir une stratégie de veille pouvant s'appuyer sur les différentes formes de diffusions prescriptives d'informations et de documents, afin de fournir une alternative à la hiérarchisation et à la pondération offertes par les différents outils de veille stratégique présents sur le marché (qu'ils soient gratuits, comme Google, ou payants).

Dans ce contexte, une première étape a consisté en l'identification d'internautes-prescripteurs, que nous nommons *agents-facilitateurs* [ALL 11a, ALL 11b, ALL 12b], afin non pas de définir en quoi ils pouvaient potentiellement influer sur l'e-réputation d'un produit ou service, mais de définir la manière dont leur intégration à un processus de veille permettait d'identifier des tendances sur des thématiques proches de celles voulues par La Poste, ainsi que les sources informationnelles traitant du sujet et pouvant inspirer les discours de l'entreprise (ou faisant autorité pour le domaine observé). Bref, de définir un contexte d'observation propre aux internautes cibles de la communication de l'entreprise afin d'insérer certains indicateurs réputationnels pouvant appuyer les objectifs de l'entreprise.

Cette recherche visant à identifier des agents-facilitateurs pertinents a été menée durant six mois (en 2011), le terrain d'observation était le site de *micro-blogging* twitter.com. Il nous a été nécessaire de définir les attributs d'identité numérique des internautes-relais de l'information postale, et d'observer ensuite de quelle(s) manière(s) ces attributs pouvaient influer sur la collecte et la sélection de l'information. Nous proposons donc ici de détailler tout d'abord les réflexions ayant amené à la détermination d'attributs identitaires propres à Twitter. Puis, de définir en quoi la détection de ces internautes-prescripteurs insérés dans une économie de la recommandation paraît aujourd'hui nécessaire pour une entreprise souhaitant développer son activité sur le Web, particulièrement dans l'optique de détecter des sources et des contenus jugés comme faisant autorité pour les internautes qui les transmettent. Enfin, nous présenterons la méthodologie retenue ainsi que certains constats issus de l'observation de ces internautes dans le cadre d'une activité de veille stratégique.

2.3. Identité numérique et attributs identitaires : quel(s) modèle(s) de détection ?

Avant de s'interroger plus avant sur les différents attributs identitaires en ligne sur lesquels nous pouvons nous appuyer pour qualifier les sources de

recommandations/prescriptions d'informations, il semble nécessaire de définir plus en profondeur le concept même d'identité numérique.

Définir l'identité numérique passe généralement par la mise en avant de la notion de traces. Ces traces (vues comme l'inscription informatisée des activités, actes de communication et d'interaction d'un internaute par les plates-formes web), permettent donc d'observer, si ce n'est de définir la présence d'un individu en ligne. D'un point de vue instrumental[4], ces traces offrent la possibilité aux systèmes algorithmiques qui les régissent, les agrègent voire les produisent, de « prédire le comportement [d'un profil] et [d'en] user comme d'une ressource » [MER 10]. Pour les autres membres d'un même réseau ou d'une même plate-forme, ces traces permettent d'interpréter qualitativement le comportement d'un sujet [GEO 10] et sont donc inhérentes au développement des interactions entre les membres d'une plate-forme ou d'un réseau social numérique. Cette approche par la trace computée par des systèmes, et avant toute forme d'interprétation, amène à considérer « l'homme comme un document » qui dispose « d'une identité dont il n'est plus 'propriétaire', dont il ne contrôle que peu la visibilité (ouverture des profils à l'indexation par les moteurs de recherche), et dont il sous-estime la finalité marchande » [ERT 09].

Par le prisme de cette définition documentaire de l'identité numérique, un premier modèle d'identification des attributs identitaires en ligne émerge dans la littérature s'orientant vers la valeur sémiotique de ces traces. Modèle développé par F. Georges [GEO 07, GEO 08, GEO 09, GEO10] que nous pourrions qualifier, à la suite de A. Coutant et T. Stenger [COU 11], « d'égocentré », car se focalisant sur la *représentation de soi* permise par ces traces, représentation qui n'est « qu'une partie de l'identité numérique : cette dernière se tisse entre l'ensemble des signes saisis par le Sujet et l'ensemble des signes valorisés par le dispositif » [GEO 09]. L'identité numérique devient donc « une transposition graphique, sonore et visuelle d'une représentation en pensée façonnée par le Sujet dans le matériau de l'interface, transposition amenant à percevoir l'identité en ligne sous trois formes :

– l'identité déclarative : ensemble des informations saisies par l'individu, à savoir (selon les plates-formes) son nom ou pseudonyme, sa date de naissance, ses goûts (pour Facebook), sa biographie (pour Twitter), etc. Attributs identitaires que nous pouvons mettre en parallèle du modèle régalien et de la notion « d'identifiants socles » [ITE 08] si ceux-ci offrent la possibilité d'accéder à l'identité civile de l'internaute ;

– l'identité agissante : « constituée des messages répertoriant les activités de l'utilisateur par une initiative du Système » [GEO 09], comme la mise à jour d'un

4. Instrumentation que nous entendons ici comme la capacité d'un système informatique ou d'un individu à s'appuyer sur ces traces pour répondre à divers objectifs : de personnalisation de l'information ou encore d'analyse de l'information par exemple.

profil (enrichissement de l'identité déclarative d'un contact), une notification soulignant une nouvelle prise de contact ou encore un nouveau commentaire posté sur un profil ;

– l'identité calculée, qui représente l'agrégation des variables quantitatives (volume de contacts, d'interactions, classements, etc.) calculées par un système.

Soulignons que ce modèle, sur lequel nous nous appuierons plus loin pour définir des attributs identitaires propres à la plate-forme Twitter, propose une forte mise en avant de l'activité du système dans l'élaboration de l'identité numérique d'un individu (identité agissante et identité déclarative), laissant une faible place aux constructions identitaires basées sur les saisies endogènes (propres à l'utilisateur – identité déclarative) et exogènes (effectuées par d'autres utilisateurs et constitutives d'une identité sociale collective) [GAL 11].

Cette approche offrant la mise en avant de certains attributs identitaires, s'inscrit selon nous dans ce que J. Denouël [DEN 11] qualifie de « mouvement exprissiviste » dans l'analyse de l'identité en ligne, considérant que « *les éléments identitaires mis en ligne n'en demeurent pas moins orientés vers autrui* » et répondent à une mise en scène de soi au sens goffmanien du terme. L'auteur souligne par la suite que l'analyse de l'identité numérique peut aussi être envisagée sous l'angle de la mise en relation : s'appuyant notamment sur les travaux de D. Cardon [CAR 08], l'identité numérique répondrait à une logique relationnelle par la mise en visibilité des traces produites. Visibilité à la fois dépendante des possibilités offertes par les plates-formes, des usages et rites sociaux en rigueur, ou de l'exposition souhaitée. D. Cardon [CAR 08] parle alors d'un « design de la visibilité » reposant sur diverses formes :

– visibilité en « paravent » (identité accessible seulement par des moteurs dédiés, par exemple sites de rencontre en ligne) ;

– visibilité en « clair-obscur » (choix des contacts à qui l'on révèle ses attributs identitaires, Facebook) ;

– visibilité en « phare » (par exemple productions de photos afin de les valoriser, Flickr) ;

– visibilité en « post-it » (mise en avant de ses disponibilités, de ses activités en temps réel, Twitter) ;

– visibilité *lanterna magica* (construction d'un avatar et d'un moi virtuel, Second life).

Cette mise en relation répondrait aussi selon certains auteurs à une quête de reconnaissance, passant « par la sélection de facettes personnelles, la mise en lumière de certains traits estimés distinctifs et la valorisation de soi par contraste »

[GRA 10] dans la constitution de son identité en ligne et le choix d'attributs identitaires.

Comme le soulignent A. Coutant et T. Stenger [COU 11], ces différentes approches amènent à mettre au pluriel le concept d'identité numérique, et de s'appuyer alors sur la notion *d'identités numériques* : « elles concernent un ensemble hétéroclite d'éléments provenant de différentes sources, humaines ou logicielles (et pas uniquement de l'individu), de différents contextes (qui vont influer sur leur sens), poursuivant différentes logiques (tentative de stabilisation d'une image de soi ou au contraire invention de soi) et produits de manière plus ou moins consciente (de l'activation spontanée de dispositions incorporées aux positions les plus réflexives) ». Dans ce sens, il nous semble intéressant, si ce n'est utile pour la suite de notre réflexion, de souligner deux autres approches constitutives des attributs identitaires en ligne d'un individu :

– l'approche structurale qui repose globalement sur le positionnement dans un réseau d'un individu, permettant notamment de déceler le « capital social » [BOU 85] de l'individu dans ce contexte. Que cela soit par la présence des « trous structuraux » [BUR 92] (nombre de contacts interconnectés, densité des liens), la force des « liens » (forts ou faibles) qu'il développe avec ses contacts [GRA 73] ou encore sa centralité dans ce même réseau (de degré, de proximité ou d'intermédiarité) [FRE 79]. Le tout offrant alors une vision *macro* de l'identité dans une approche relationnelle de celle-ci, et surtout permettant de mieux définir le contexte pouvant influer sur l'interprétation du sens ;

– l'approche réputationnelle, ou de redocumentarisation [ZAC 06] de l'identité calculée. Si les plates-formes et les systèmes qui les régissent proposent une forme de *commensuration* [ESP 98] des activités des internautes, de nouveaux outils viennent ensuite recalculer ces données afin de produire des classements. Par exemple, l'outil Klout.com vise à agréger l'ensemble des activités d'un internaute sur diverses plates-formes afin de produire ensuite un « score de son influence ». Calcul propre à la notion de réputation (d'un point de vue étymologique) venant réarticuler sémiotiquement l'interprétation du comportement de l'internaute. Plus prosaïquement, un moteur de recherche comme Google se base sur certains critères d'activité pour évaluer la popularité d'une source (source pouvant être entendue ici comme le profil d'un internaute indexé par ledit moteur).

Ces divers attributs identitaires esquissés, nous proposons maintenant de les appliquer à la plate-forme de *micro-blogging* Twitter, afin de montrer un exemple concret et *in situ* de ces attributs, mais aussi dans l'optique d'intégrer par la suite ces divers attributs à la méthodologie d'identification d'agents-facilitateurs que nous présenterons.

2.3.1. *Twitter et attributs identitaires*

Créé en 2006, Twitter est un réseau social numérique offrant la possibilité à ses utilisateurs de diffuser des messages textuels et hypertextuels de 140 caractères maximum (le *micro-blogging*). La plate-forme permet de se connecter de manière asynchrone entre utilisateurs (pas de réciprocité automatique dans la prise de contact), d'échanger directement avec d'autres utilisateurs en les « interpellant » par la syntaxe *@NomDuCompte*, ou encore (comme de nombreuses plates-formes) de personnaliser son compte (couleur, avatar, nom, pseudo, biographie, etc.).

De par son modèle suivi/suiveur, Twitter permet de privilégier « une modalité de relation fondée sur la curiosité et le désir d'être informé, plutôt que par la tentation du paraître » [GEO 10], si tant est que ce désir de paraître ne se formalise pas par le choix des informations diffusées, où la mise en avant de certaines activités jugées comme valorisantes par un utilisateur ou son réseau (comme signifier que l'on se trouve à tel endroit à la mode ou avec telle personne). Ce « désir d'être informé » repose néanmoins sur un choix : celui du ou des comptes auxquels l'on doit s'abonner. Si de nombreux facteurs entrent en jeux (sur lesquels nous reviendrons partiellement dans l'explicitation de notre méthodologie), l'interface de Twitter permet de développer des attributs identitaires pouvant amener ou non à s'abonner à un compte.

Nous présentons (figure 2.1) ces critères, en les situant par zones sur l'image d'un compte[5] ; ensuite en catégorisant ces attributs en fonction des modèles cités *supra*.

Le tableau 2.1 met en corrélation les attributs présentés dans la littérature avec ceux observables sur un compte Twitter.

Nous pouvons ajouter à ces attributs ceux liés à un message en particulier (un tweet) et identifiables en cliquant sur le bouton « dérouler » présent sous ces tweets : heure et date du tweet (identité calculée), outil utilisé pour envoyer le tweet (comme un téléphone portable – identité agissante), la localisation (si l'utilisateur l'accepte – identité déclarative), les personnes ayant interagi avec ce tweet – identité agissante.

5. En l'occurrence celui de l'auteur de ce chapitre, accessible à l'adresse Web suivante : https://twitter.com/#!/CaddeReputation.

Figure 2.1. *Attributs identitaires sur Twitter, exemple d'un compte*

Zones Attributs	Identité déclarative	Identité agissante	Identité calculée	Visibilité
Zone 1	Nom de l'auteur (*Camille A*), nom du compte (*caddereputation*), biographie, pays (*France*), avatar, lien hypertexte externe			Accessible à tous, attributs référencés par les moteurs de recherche
Zone 2		Volume de tweets envoyés sur le réseau	Volumes d'abonnements et d'abonnés	Accessible à tous si profil « ouvert »
Zone 3		Tweets envoyés	Date de l'envoi	Accessible à tous si profil « ouvert » et référencés par les moteurs de recherches (de manière aléatoire)
Zone 4		Abonnements (accès direct aux comptes procédant de l'usage	Abonnés	Accessible à tous si profil « ouvert »

		de la plate-forme), favoris, listes[6]		

Tableau 2.1. *Attributs identitaires et représentation de soi sur Twitter*

De plus, il convient de souligner que Twitter propose au détenteur d'un compte de recevoir des notifications quant à l'activité de ses abonnements (mise en favoris et abonnement à un nouveau compte). Notifications en provenance des contacts qui, comme le souligne F. Georges [GEO 09] « incitent à interpréter leur comportement selon des règles qui dépendent entièrement du Sujet ». De même, les attributs que nous pourrions qualifier de réputationnels et de structuraux nécessitent de s'appuyer sur d'autres outils de computation et d'évaluation offrant la possibilité de mesurer : le volume de reprises des contenus diffusés (les « re-tweets ») ; le volume d'interaction moyen pour un compte ; l'évolution du nombre d'abonnés/abonnements ; les sources des informations diffusées (vers où pointent les liens hypertextes) ; le volume de clics sur ces liens ; les comptes avec lesquels il y a les plus forts ou faibles volumes d'interactions, etc. Ensemble de *commensurations* permettant de traduire l'activité de l'internaute en chiffres (puis éventuellement de les classer en les comparant avec d'autres comptes), mais dont la collecte de données reste fortement dépendante des capacités de calculs de la plate-forme Twitter (et de l'accessibilité que Twitter souhaite offrir à ses données).

Twitter apparaît donc comme une plate-forme où la construction de l'identité numérique d'un individu s'avère possible et repose sur divers attributs identitaires, principalement liés à la notion d'identité agissante, et dont la visibilité que D. Cardon [CAR 08] qualifie en « post-it » devient de plus en plus perméable à la computation par d'autres plates-formes. Ceci étant, et si Twitter peut être perçu avant tout comme un réseau dont l'activité principale s'oriente sur le partage d'informations plus que sur la construction de la *représentation de soi*, il convient de s'interroger sur la manière dont ces attributs identitaires peuvent influer sur l'information qui est diffusée sur ces profils : y a-t-il une forme de personnalisation du à l'usager ou aux computations de la plate-forme ? Ces attributs nous permettent-ils d'évaluer l'autorité de la source ? La crédibilité ou la popularité de l'information ?

2.4. Les attributs identitaires pour la structuration de l'information sur le Web : vers une nouvelle forme d'autorité ?

6. Un utilisateur peut réunir sous forme de listes plusieurs comptes afin notamment d'opérer un regroupement thématique. Ici, nous pouvons avoir accès aux listes créés par l'utilisateur, mais aussi celles où il est inscrit par d'autres.

« Après avoir été pensée comme une cible, qui venait après une information déjà constituée, la personne est devenue une ressource, un agent de pertinence et un opérateur de liens entre les informations. » [MER 09, b]. Cette idée que l'individu, et sa « transposition » numérique par le biais d'attributs identitaires en ligne, apparaisse comme un agent de pertinence ne se retrouve pas seulement d'un point de vue du sujet souhaitant collecter de l'information en ligne, mais aussi dans la structuration même (et surtout à venir) du fonctionnement des différents acteurs « dominants » de la recherche d'informations sur le Web. Avec le développement de son réseau social numérique « Google+ », les premières analyses[7] soulignaient que Google visait principalement par sa mise en place à concurrencer ceux déjà présents (Facebook en tête). Si cette stratégie semble effectivement se confirmer, il s'avère que la création de ce réseau apporte au moteur de recherche une tout autre valeur ajoutée : la personnalisation de la recherche ; et une nouvelle manière de pondérer et hiérarchiser les résultats de son moteur de recherche. En effet, la création d'un compte pour accéder aux nombreux services Google entraine systématiquement la création d'un compte Google+. Une fois l'utilisateur connecté à son compte, le moteur de recherche lui propose alors, pour une requête donnée, des résultats classés non pas en fonction seulement de ses critères « classiques » d'évaluation (le *PageRank* notamment), mais aussi en fonction des préférences de ses contacts (figure 2.2) : les attributs identitaires servent donc à personnaliser l'information et à orienter le choix par la mise en avant de préférences.

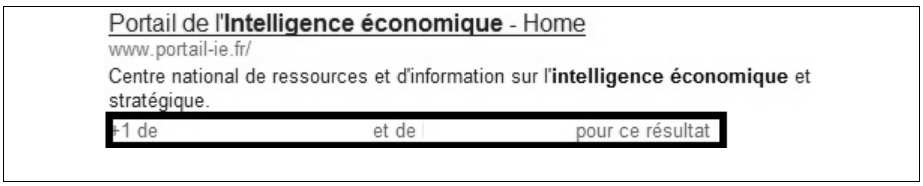

Figure 2.2. *Exemple d'un résultat, après connexion à un compte Google, pour la requête « intelligence économique ». Le moteur de recherche met en avant les résultats déjà partagés par d'autres*

Mais au-delà de la personnalisation des résultats en fonction des recommandations effectuées par d'autres membres d'un même réseau (ici le bouton « +1 » équivalent du *Like* de Facebook), le moteur de recherche effectue d'autres tests afin de mettre en avant, dans les résultats, certains attributs identitaires de l'auteur d'un texte ou d'une source (figure 2.3).

7. Analyses essentiellement en provenance de praticiens, par exemple www.mediassociaux.fr/2011/10/14/pourquoi-le-succes-de-google-est-assure (consulté le 9 juin 2012).

Figure 2.3. *Exemple de résultat mettant en avant l'identité de l'auteur d'une page Web, pour la requête « site:mediassociaux.fr facebook » sur le moteur Google.fr le 9 juin 2012, sans connexion à un compte Google*

Ces deux exemples illustrent la personnalisation constante des résultats (qui n'est pas à proprement parler nouvelle, puisque de nombreux sites, tel Amazon, en ont fait un avantage concurrentiel) s'appuyant en particulier sur l'identité des utilisateurs (dont les contacts font partie intégrante), mais aussi la redéfinition constante de ce qu'est l'autorité et la popularité d'une source.

La recherche et la sélection d'informations sur Internet questionnent depuis quelques années les modèles classiquement établis d'autorité de la source et de crédibilité de l'information [SER 04]. L'intégration des attributs identitaires en ligne comme facteurs de pondération et de hiérarchisation des résultats ne fait que complexifier cette problématique. E. Broudoux [BRO 07] propose un concept de l'autorité informationnelle basé sur l'imbrication d'autres formes d'autorités, et dont l'arrivée du Web nécessite de prendre en compte certaines spécificités :

– l'autorité énonciative (celle de l'auteur, sa notoriété/réputation, volume et type des publications) ;

– l'autorité de contenu (genre, architexte, paratexte associés au contenu publié sur la source) ;

– l'autorité du support-logiciel ou en l'occurrence de la plate-forme (type et volume de lectorat/visiteurs) ;

– l'autorité du groupe (comme une communauté d'intérêts sur le Web).

Là où les moteurs de recherche comme Google (principale porte d'entrée des internautes sur le Web en France) s'axaient principalement sur l'évaluation de l'autorité de contenu (en mesurant le nombre de liens hypertextes pointant vers une source notamment), il semblerait donc que l'autorité énonciative soit aussi un facteur de pondération forte. Ainsi que, dans une moindre mesure pour l'instant, celle du groupe.

L'autorité énonciative se base sur la notoriété et la réputation de l'auteur d'une source. Notoriété et réputation dont le calcul repose sur une forme de redocumentarisation de l'identité calculée. Cette nouvelle forme de pondération des

résultats, mais aussi de sélection pour un internaute souhaitant prescrire du contenu, s'appuie donc en partie sur les attributs identitaires propres à un profil s'exprimant sur une source Web. Ce passage de l'autorité (telle qu'elle était déléguée par des institutions de référence) à la popularité (telle qu'elle est calculée aujourd'hui par certaines plates-formes), et de la pertinence (de l'information ou de la source) à l'influence (reposant sur un calcul de la présence et de l'agir d'un profil) amène les professionnels tout comme les individus en recherche d'information à redéfinir leurs pratiques de sélection [LED 06]. Qui plus est si ces professionnels (au sein des organisations particulièrement) ou internautes cherchent du contenu « expert ». Si l'expertise d'un auteur pour son sujet ne répond pas à la question de la crédibilité de l'information, elle permet tout de même d'identifier les contenus qui, pour un sujet donné, seront favorisés et mis en avant par les divers systèmes traitant ces contenus. Sur le Web, l'expertise peut être vue comme la capacité à produire du contenu répondant à une requête d'un internaute, contenu dont la visibilité reposera là encore sur diverses computations et divers calculs, qu'ils portent sur la médiatisation de ce contenu (positionnement sur un moteur de recherche, volume de liens entrants dénotant la popularité, etc.) ou « l'influence » du profil (volume de contacts, d'échanges sur un réseau, de sollicitations, de citations) [ALL 10]. Les points de repère semblent donc brouillés par la constante redocumentarisation de l'identité calculée des internautes et son intégration à la sélection de l'information, et par le glissement de certaines valeurs quant à la pertinence d'un contenu lorsque celui-ci n'est pas seulement accessible par une forme de médiation algorithmique, mais lorsque cette médiation même vise à déterminer ce qui est pertinent ou ce qui ne l'est pas pour un sujet précis. De plus, il semble intéressant de noter que certains internautes souhaitant positionner leur expertise sur un sujet (comme les blogueurs) instrumentalisent eux aussi les divers critères d'évaluation de leur expertise/notoriété/réputation/popularité, transformant en une forme de jeu le choix des attributs identitaires ou des formes de leurs contenus afin de renvoyer l'image (généralement celle d'un expert sur le sujet qu'ils traitent) qui leur semble la plus appropriée pour atteindre leurs objectifs [MAR 12].

Les attributs identitaires en ligne sont donc aujourd'hui un facteur de pondération de l'information, de sélection et de mise en avant de celle-ci, que cela soit effectué par des systèmes algorithmiques ou que (plus naturellement pourrait-on dire) cette sélection repose pour un individu sur la *représentation de soi* que fournissent les auteurs de contenus. Si nous prenons alors le prisme des organisations souhaitant identifier des contenus experts ou des internautes faisant autorité dans leur domaine dans le cadre d'une collecte d'informations, ce constat questionne la manière dont une source sera identifiée et intégrée à un processus de veille stratégique. Il nous semble que cette intégration ne doit pas viser à identifier des indicateurs de pertinence « dans l'absolu » d'une information ou de sa source. Mais plutôt à intégrer le fait que le choix (à but prescriptif) d'une information ou d'une source, effectué par certains internautes sur les réseaux socionumériques,

permet d'appréhender ce qui – pour l'internaute « fait autorité », considérant que cette autorité repose plus sur la capacité d'une source à se médiatiser, à développer sa notoriété et son expertise par l'appui d'indicateurs réputationnels, qu'à fournir une information crédible (l'un n'empêchant pas l'autre). A l'inverse, et de manière récursive, certains attributs identitaires semblent permettre d'identifier ces profils que les indicateurs réputationnels amènent à considérer comme experts, influents ou encore potentiellement producteurs ou relayeurs d'informations pertinentes sur un sujet. Pertinence qui s'entend ici au niveau de la propre « réalité sociale » de l'individu sélectionnant et diffusant l'information, et de son public [SCH 87], et non pas, encore une fois, de façon absolue.

Si les attributs identitaires permettent de définir des indicateurs de pertinence pour la sélection de l'information, si « l'identité numérique [est] le seul dénominateur commun d'une masse de données hétérogènes, qu'aucune classification a priori ne peut plus ordonner » [MER 09b], il paraît maintenant intéressant de se questionner sur l'intérêt qu'une organisation pourrait avoir à se reposer sur ces attributs, avant d'aborder quelques notions méthodologiques de détection et d'emploi de ces attributs.

2.5. Rechercher l'information sous l'emprise des dispositifs : de la nécessité de s'appuyer sur la prescription des internautes ?

De nombreux chercheurs comme D. Boyd et K. Crawford [BOY 11] s'interrogent sur la manière, épistémologique comme méthodologique, d'aborder la profusion de données informatisées présentes en ligne, lorsque ces données (leurs modes de diffusion, leur traitement, etc.) sont régies par les plates-formes sur lesquelles elles sont produites, et dont on ne connait que peu de choses sur leur mode de traitement. Ce que soulignent ces chercheurs ne s'applique pas seulement aux activités de recherches scientifiques s'appuyant ou observant le Web, mais aussi selon nous à chaque individu souhaitant collecter et analyser des informations/données sur des supports Web. Les organisations voulant mettre en place des dispositifs de veille stratégique se trouvent alors face à un dilemme similaire : comment faire la part entre ce que ma méthodologie et mes outils me permettent de collecter, et ce que les plates-formes par/sur lesquelles je collecte l'information veulent bien me fournir ? Cette problématique des « techniques en boite noire » [RIE 10] nécessite non seulement de repenser la manière d'aborder la recherche, la collecte et le traitement des informations présentes en ligne, mais aussi d'intégrer aux prises de décisions issues de cette collecte (particulièrement en entreprise) ce que nous pourrions qualifier de « subjectivité algorithmique ». Car aujourd'hui, pour collecter des données issues d'une plate-forme comme Twitter par exemple, il est nécessaire de

passer par les API[8] fournies par le système. Et, comme toute entreprise commerciale, Twitter vise à atteindre un équilibre économique par la monétisation des données et informations circulant sur son réseau. Il est donc admis que la plate-forme (au gré de ses accords commerciaux ou de ses objectifs en termes de *business model*) modifie constamment ce à quoi elle souhaite donner accès (gratuitement ou de manière payante[9]).

Parallèlement à cela, et comme nous l'avons signalé précédemment, les moteurs de recherche personnalisent constamment les résultats qu'ils fournissent (géolocalisation, historique de navigation, etc.) redéfinissant ainsi la manière dont sont hiérarchisées les informations, voire le volume d'information (spécifiquement pour Google) auquel ils donnent accès [VIS 12]. Et, pour les professionnels, internautes ou chercheurs, souhaitant s'appuyer sur les outils mis à disposition par les moteurs pour rechercher plus précisément des informations, l'on s'aperçoit que les options de recherches permettant de filtrer plus spécifiquement les résultats (et donc de diminuer certains biais liés à leur personnalisation) diminuent sensiblement, laissant l'utilisateur du moteur face à un nombre réduit d'options [ERT 11] (et donc « dépendant » des choix du moteur). Face à cette double contrainte, le professionnel de la veille stratégique se retrouve non seulement à questionner la capacité des outils de recherche d'informations à lui fournir l'ensemble des données/informations qu'ils stockent ou qui circule sur leurs plates-formes, mais aussi la manière dont cette donnée/information est présélectionnée (et donc hiérarchisée) avant de lui être transmise. Si certains de ces critères de hiérarchisation sont identifiables (comme le fait de s'appuyer sur les attributs identitaires de l'auteur d'une source), il nous semble qu'une manière complémentaire de collecter de l'information tout en s'appuyant sur une certaine subjectivation dans sa collecte, et de s'appuyer sur la prescription informationnelle opérée par certains internautes, en s'insérant ainsi dans « l'économie de la recommandation ». Economie qui, nous rappelle [PIE 11] « se situe dans le chemin cognitif et la pertinence de l'information nécessaire au processus décisionnel. ».

Cette insertion dans l'économie de la recommandation comme complément aux fonctionnements des algorithmes des plates-formes, passe par l'identification des pratiques informationnelles pouvant s'avérer utiles aux attentes des professionnels

8. API : *Application Programming Interface*, est une interface fournie par un logiciel (ou une plateforme web en l'occurrence) permettant les échanges avec d'autres logiciels/programmes/ plateformes. Twitter permet notamment de développer des outils statistiques sur les comptes de ses contributeurs, les échanges générés, le traçage des liens hypertextes, etc.

9. Nous pouvons citer, à titre d'illustration, cette modification des règles de collecte de l'information sur le réseau Twitter, survenue en avril 2012, et diminuant sensiblement le nombre de requêtes possibles dans une même heure (et donc par extension le volume et le type d'informations à collecter) : https://dev.twitter.com/docs/rate-limiting (consulté le 9 juin 2012).

ou des organisations. Pratiques définies comme l'ensemble des actions et des choix opérés par un utilisateur lors d'une phase de recherche d'information sur un autre utilisateur ou sur lui-même » [GEO 10]. Si les motivations, les intentions, ayant poussé au choix de diffuser telle ou telle information sur une plate-forme restent sensiblement inconnues à l'observateur de ces actions, il est raisonnable de supposer qu'en « en laissant un contenu informationnel sur le Net (ou en le relayant), l'internaute présuppose un lecteur et cette présupposition induit le type de contenu qu'il envoie sur le Net » [MON 11]. Identifier le lecteur potentiel d'une information diffusée ouvre donc la possibilité de circonscrire le choix ayant motivé à la diffusion de cette information. A l'inverse, l'information diffusée par un profil sur le Web présuppose une certaine appétence du public récepteur de cette information. Ce constat permettra par la suite de définir de quelle manière la sélection d'internautes prescripteurs peut intervenir dans la mise en place d'une veille stratégique spécifiquement axée sur la collecte d'informations visant à construire le discours communicationnel d'une entreprise : en identifiant des internautes ayant déjà relayé des informations ou contenus de/sur l'entreprise, l'on présuppose qu'ils s'adressent à un public portant un intérêt à l'entreprise (ou aux thématiques attenantes) et par extension les informations diffusées par ces internautes sont potentiellement utiles pour appréhender le contexte global dans lequel doit s'insérer le discours de l'entreprise (et, pourquoi pas, construire des indicateurs liés à ce contexte informationnel). De même, l'approche structurale citée *supra* trouve ici écho : l'identité numérique d'un internaute prend sens dans le contexte dans laquelle elle s'insère, par les liens qu'elle noue avec les autres membres du réseau, formant des communautés d'intérêts. Qui plus est sur des plates-formes comme Twitter où l'information diffusée est le principal facteur d'adhésion entre les membres. De plus, observer le lectorat de profils diffusant des contenus informationnels permet de définir un contexte propre aux profils observés. En effet, si l'identité est dépendante du contexte dans lequel elle se construit [GEO 09], l'appréhension des attributs identitaires qui la constitue ne pourra se faire sans prise en compte de ce contexte, que nous pouvons définir à trois niveaux :

– macro : le Web de manière globale, avec les indicateurs constitutifs de cette forme d'autorité basée sur la notoriété, l'influence, l'expertise et certains indicateurs réputationnels. Mais aussi le contexte informationnel dans lequel s'insèrent les contenus diffusés par l'internaute (nombre de résultats déjà présents sur le sujet, positionnement par rapport aux autres sources de références, etc.) ;

– méso : la plate-forme sur laquelle se construisent ces attributs identitaires (comme dans l'exemple de Twitter présenté plus haut), comprenant notamment le public présent. Mais aussi certains usages ou « rites » propres à la plate-forme observée ;

– micro : au niveau des informations, documents, contenus, produits et diffusés par le profil.

Au niveau micro, nous pouvons considérer que les informations diffusées par les internautes de manière prescriptive (puisqu'elles s'adressent à un public – réel ou souhaité) font partie intégrante des traces constitutives de son identité numérique. Si ces informations sont clairement rattachées à son identité (de par les moteurs de recherches, ou si elles se trouvent sur le compte de l'internaute en question) et se situent sur des plates-formes dont la production (comme les blogs) ou le partage d'informations (comme Twitter) sont le principal moteur. Le contexte créé par les attributs identitaires de l'internaute, du macro au micro, vient alors participer à la réarticulation sémiotique et documentaire (la redocumentarisation) des contenus qu'il diffuse, que ce soit implicitement (par « l'autorité » prêtée à l'internaute qui diffuse) ou explicitement (par l'ajout de métadonnées ou par la reformulation pour répondre aux attentes du lectorat).

Cette capacité d'un internaute à collecter du contenu, le hiérarchiser, le mettre en forme, le redocumentariser pour l'adapter aux attentes de son lectorat ou au contexte d'émission, puis à le diffuser – en résumé à opérer comme un filtre informationnel sur un réseau ou pour une communauté d'intérêts – nous permet de définir ce que nous nommons un *agent facilitateur*.

2.5.1. *Les agents-facilitateurs : filtres de l'information sur les réseaux socio numériques*

La notion d'agent-facilitateur vise à circonscrire et qualifier certaines pratiques informationnelles d'internautes sur le Web [ALL 11a, ALL 11b, ALL 12b]. Agents pour lesquels nous présenterons une méthodologie d'identification dans la section 2.6.

Cette notion vise à souligner plusieurs aspects des pratiques des internautes :

– en tant « qu'agents » ils prennent part à un système, en l'occurrence une plate-forme Web 2.0, dans lequel ils jouent un rôle, ici de sélection et de prescription de l'information :

– par « facilitateur », nous soulignons ici leur capacité à faciliter l'accès aux informations sur des thématiques données pour leurs contacts ou leurs lecteurs sur ces plates-formes.

A proprement parler, ces pratiques ne sont pas nouvelles sur le Web : de la diffusion d'informations sur les premières listes de diffusion propres à l'Internet (comme UseNet) à l'utilisation des plates-formes de *bookmarking* (comme Diigo ou Delicious), de nombreux internautes ont pour principale démarche de sélectionner et diffuser de l'information qu'ils jugent pertinente (ou qui peut d'ailleurs le devenir après sa diffusion si leurs attributs identitaires les dotent d'une certaine autorité sur le sujet). De plus, et comme nous le soulignons à la suite de T. Stenger [STE 09], les

activités informationnelles sur les réseaux socionumériques en général se font sur un mode prescriptif. Cependant, ces activités de filtrage et de redistribution de l'information sur le Web trouvent aujourd'hui un nouvel écho et concentrent les attentions avec le développement de ce que de nombreux praticiens (et à leur suite des chercheurs) nomment : la curation[10]. Du latin *curare*, la curation suppose que certains internautes vont « prendre soin » des informations qu'ils souhaitent conserver et/ou diffuser à leur lectorat. Cette pratique a engendré le développement d'une véritable industrie : en à peine deux ans, de nombreuses plates-formes dédiées à la curation sont apparues sur le Web. Si le modèle économique de ces plates-formes reste difficile à cerner, leur développement a le mérite d'attirer l'attention sur la capacité que certains internautes ont à sélectionner, rééditer, mettre en contexte et diffuser de l'information. Du point de vue des entreprises, les activités de curation sur des plates-formes dédiées apparaissent comme utiles afin de détecter des « signes d'alerte précoce » [LES 01], en proposant l'agrégation de ressources thématiques parfois mal indexées par les outils de veille, et surtout en proposant une mise en contexte pouvant favoriser l'interprétation de ces signes [ALL 11a].

Ainsi, ces plates-formes proposent des attributs identitaires propres aux « curateurs » pouvant ici encore favoriser la sélection des profils, mais aussi la pertinence des informations collectées : thématiques abordées (identité agissante ou calculée), volume de contenus présents (identité calculée), nombre d'abonnés au compte sur la plate-forme (identité agissante, permettant de définir une certaine forme de notoriété voire d'expertise reconnue sur le sujet), description de l'auteur du compte (identité déclarative), etc.

Nous préférons le terme d'agent-facilitateur à celui de « curateur » pour plusieurs raisons : il englobe des pratiques plus larges que celles observables sur les seules plates-formes dites de curation, il nous semble plus explicite et surtout moins connoté que « curateur » (notamment pour sa signification dans la langue française, et par l'aspect commercial et marketing qu'il peut revêtir). Les agents-facilitateurs sont pourvoyeurs de signes d'alertes précoces, mais ils participent aussi à la construction de la notoriété d'un sujet ou du produit/service d'une entreprise [ALL 11, ALL 11b]. De par l'affordance aux outils ou plates-formes qu'ils utilisent, et pour personnaliser l'information qu'ils diffusent afin d'attirer ou de répondre aux attentes d'un lectorat, les agents-facilitateurs opèrent une forme de redocumentarisation par prescription : par la reformulation, l'ajout de commentaires ou de métadonnées du document qu'ils diffusent ; mais aussi par l'expertise que l'on peut interpréter des différents attributs identitaires qu'ils rendent visibles. Pour un professionnel de la veille ou pour tout autre observateur, et au-delà de leurs capacités à identifier des sources ou des informations non prises en compte par certains outils de veille, il apparaît que la redocumentarisation constante effectuée

10. Voir à ce sujet : *Documentaliste-Sciences de l'Information*, vol. 49, n°1, 2012.

par ces agents-facilitateurs peut permettre de structurer l'environnement informationnel dans lequel la stratégie de collecte d'informations est opérée. A la suite de L. Quéré [QUE 97], nous entendons que « la structuration de l'environnement réduit la gamme des faits à sélectionner et des choix à opérer, en même temps qu'elle permet de remplacer différentes tâches cognitives (impliquant des représentations et des manipulations de symboles) par une perception directe ». Ici, la gamme des faits à sélectionner et des choix à opérer (en termes de recherche et de sélection de l'information) est réduite par la sélection opérée par l'agent-facilitateur. De plus, cette structuration par l'intégration d'agents-facilitateurs à un processus de veille vient remplacer certaines tâches cognitives liées à la compréhension des fonctionnements des API et autres algorithmes (puisque l'on se repose ici sur les capacités de l'agent-facilitateur à opérer ces tâches). La perception est donc directe : collecte de l'information telle qu'elle est diffusée par l'agent-facilitateur dans son contexte.

Enfin, précision que le rôle d'agent-facilitateur doit être appréhendé comme un construit stratégique de l'entreprise souhaitant s'appuyer sur certaines pratiques pour collecter de l'information : en fonction des thématiques, des objectifs et des attentes de l'entreprise, les pratiques de sélection et de diffusion de l'information n'auront pas la même importance. Même si certains utilisateurs de plates-formes telles Twitter semblent cependant ne se focaliser que sur cette activité de filtrage de l'information, d'autres usages des outils (conversation, par exemple) sont généralement observables.

Dans un contexte où l'information sur le Web est de plus en plus traitée par des algorithmes dont les usagers connaissent peu les tenants et aboutissants, il apparaît que l'appui sur les pratiques informationnelles prescriptives de sélection et de diffusion de l'information, dont les pratiquants les plus aguerris peuvent être qualifiés d'agents-facilitateurs, peut devenir une solution complémentaire dans la mise en place d'une stratégie de veille. Par l'autorité qu'ils acquièrent aux yeux d'un public ou pour certaines plates-formes, par la redocumentarisation qu'ils effectuent, et par la structuration de l'environnement informationnel qu'ils offrent à une organisation les intégrants dans ses stratégies de collecte et de traitement de l'information, ces internautes-prescripteurs peuvent participer à l'atteinte de certains objectifs de l'entreprise. Seulement : de quelle(s) manière(s) identifier ces profils ? Et en quoi leurs attributs identitaires jouent un rôle dans la sélection de l'information qu'ils diffusent ?

2.6. A la recherche d'une congruence informationnelle : le cas de La Poste et de l'identification d'agent-facilitateurs

Avant d'aller plus avant dans la présentation de la recherche effectuée à La Poste Courrier, il nous semble intéressant de tracer certaines limites à la méthodologie mise en place, et de globalement en définir les contours.

Comme toute recherche-action en entreprise, celle effectuée ici part en partie d'une demande de l'entreprise : comment collecter sur le Web de l'information permettant d'une part, d'établir des tendances sur des sujets donnés ; et d'autre part, d'identifier des ressources pertinentes pour construire des discours et stratégies communicationnelles en adéquation avec les discours déjà tenus par certains internautes sur ces mêmes sujets. Globalement, de quelle(s) manière(s) les discours concernant les innovations technologiques produites par La Poste peuvent-ils s'insérer dans le contexte informationnel numérique dans lequel les cibles de la communication naviguent au quotidien (permettant ainsi de développer peu à peu une *e*-réputation en adéquation avec les attentes de l'entreprise et les représentations déjà formées des internautes) ?

Comme nous l'avons détaillé dans ce chapitre, il nous a semblé que pour répondre en partie à ces interrogations, il était nécessaire d'introduire aux processus de veille stratégique déjà existants les pratiques informationnelles d'internautes-prescripteurs, mais aussi la personnalisation de l'information inhérente à ces pratiques, et sources de structuration de l'environnement d'observation. Notons que même si la recherche de ces relais n'était pas explicite au lancement du projet, les résultats d'une première observation ne pouvaient qu'orienter vers ces dits relais (comme nous le verrons plus loin, 50 % et plus des interactions concernant le produit observé ont été des relais sur Twitter). Cette détection d'agents-facilitateurs est donc passée par :

– la catégorisation de ceux identifiés comme ayant traité du lancement d'un produit postal (Digiposte) ;

– la qualification à l'aide de leurs attributs identitaires en ligne ;

– le tri dans le temps en fonction des informations qu'ils diffusent et de l'intérêt réel en termes de prise de décision et de construction du discours pour La Poste.

Il paraît alors important de préciser que la caractérisation des données ou attributs identitaires sélectionnés dépend fortement des besoins exprimés par l'entreprise, du contexte dans lequel s'insère la collecte d'informations, et que la catégorisation répond à une approche inductive visant justement à mettre en corrélation certains attributs identitaires à l'interprétation possible de l'information collectée. Car, comme le mettent en avant A. Coutant et T. Stenger [COU 11] : « il est impossible de caractériser *a priori* une donnée identitaire, car elle évoluera selon le contexte, l'audience et même dans le temps ». Le processus de recherche a donc visé à définir une typologie d'acteurs en fonction des observations effectuées (approche inductive), plutôt que d'essayer de définir au préalable des attributs

identitaires permettant d'identifier par la suite certains acteurs. De même, l'approche que nous proposons par la suite ne vise pas à déterminer des catégories immuables d'acteurs afin de prédire leur comportement ou de le reproduire, mais bien encore une fois (nous insistons sur ce point) d'observer de quelle manière les pratiques informationnelles de ces acteurs peuvent être levier de la structuration de l'environnement informationnel de l'entreprise, et en quoi leurs attributs identitaires peuvent être vecteurs de sélection de l'information. Nous nous inscrivons donc dans certaines réflexions d'Erving Goffman, reprises par A. Monseigne [MON 11] : « *Des situations et des Hommes et non des Hommes et des situations*, pour indiquer qu'il faudrait pouvoir lister la complexité des paramètres en jeu dans chaque situation pour pouvoir prédire les comportements futurs d'un individu à partir de son passé ». Liste difficilement réalisable dans un environnement aussi mouvant qu'est le Web. Par congruence, nous entendons donc la capacité à identifier des agents-facilitateurs dont le filtre opéré de l'information vient répondre aux attentes exprimées par l'entreprise.

La première étape de notre action fut l'identification (à l'aide d'outil gratuits, comme le moteur Google ou encore d'outils plus spécifiques[11] en fonction des types de plates-formes) des sources et informations concernant le lancement sur le marché français du produit Digiposte, et ce du 1er janvier au 30 juin 2011. Parmi les 390 résultats collectés[12] contenant l'expression Digiposte ou pointant un lien vers des contenus traitant de Digiposte (dont le site du produit lui-même), 186 se trouvaient sur Twitter (déduction faite des tweets redondants en provenance d'un même profil). Nous nous sommes donc focalisés sur cette plate-forme en tant que principale source d'information sur le sujet. Le choix de se concentrer, puis d'intégrer à une stratégie de veille ces profils, reposait sur le fait qu'ils s'adressaient à un public réceptif aux discours concernant certains produits postaux (pour des thématiques biens spécifiques – en l'occurrence ici la dématérialisation du courrier, le stockage en ligne, ou plus largement les innovations numériques), c'est-à-dire le public qui sera visé par la suite par l'entreprise, et dont il paraît intéressant de capter les informations auxquelles il est potentiellement exposé chaque jour.

Le premier fait marquant fut que l'ensemble des tweets identifiés contenait un lien pointant soit vers le site de Digiposte (une minorité), soit vers des articles traitant de Digiposte. Deuxième constat, seuls 30 % des tweets identifiés étaient « originaux », les autres n'étant que des reprises (des « re-tweets). Nous nous trouvions alors à deux niveaux de redocumentarisation : le premier provenait de la

11. Comme l'outil Topsy.com dédié à Twitter, ou Boardreader.com pour les forums de discussion.

12. Tout du moins, que les outils et les plates-formes nous ont rendus accessibles. Ce volume ne prenant pas en compte, par exemple, les commentaires attenants à des articles de blogs ou presse en ligne.

diffusion d'un lien et du contexte qui l'accompagne (attributs identitaires de celui qui diffuse, et reformulation potentielle de la dénomination de la ressource vers lequel le lien pointe) ; le deuxième provenait de la reprise de ce tweet par un autre profil (associant ainsi ses propres attributs à ceux implicitement induits par la source reprise). Ce constat interroge d'ailleurs sur l'unité documentaire à construire pour l'analyse : doit-on considérer seulement le tweet original comme influant sur la compréhension du contexte et l'appropriation possible par le public de ce tweet ? Ou y intégrer aussi la reprise et les attributs identitaires de reprenant ? Lequel indexer ?

Suite à cette première identification, nous avons choisi volontairement de réduire le champ d'observation en opérant une présélection des profils à observer et à intégrer par la suite à un processus de veille. Cette présélection s'est effectuée sur certains attributs identitaires relevant selon nous d'une marque forte pour l'appréhension du contexte. Nous avons donc sélectionné des profils dont au moins l'un des critères identifiables était le suivant :

– présence d'un site dans la partie biographique de Twitter (zone 1 de la figure 2.1), considérant que le fait de pouvoir relier un profil à d'autres comptes ou pages Web propres à la présence de l'auteur du profil serait pertinent pour élargir la définition de son autorité (et éventuellement intégrer de nouvelles sources à surveiller) ;

– volume d'abonnés supérieur de 20 % à la moyenne des autres profils identifiés : ainsi, nous sélectionnions des profils dont le public s'avèrait plus large, et donc plus intéressant pour l'entreprise ;

– production d'un tweet original, soulignant ainsi les capacités à aller collecter et diffuser du contenu sur le Web (tout du moins à s'appuyer sur ses propres pratiques *autoritatives*) ;

– reprise d'un *tweet (re-tweet)* concernant Digiposte, permettant d'identifier les profils s'appuyant sur d'autres comptes estimés pertinents. Que ces reprises concernent des utilisateurs *preservers* (qui rediffusent à l'identique) ou *adapters* (qui l'adaptent au contexte et contraintes de la plate-forme) [BOY 10] ;

– profil ayant vu son tweet relatif à Digiposte repris par d'autres membres de la plate-forme, soulignant ainsi que (pour ce sujet) le profil dispose d'un public prêt à interagir avec le contenu diffusé (et dans un certain sens à légitimer ce contenu en le diffusant à d'autres) ;

– publication d'un tweet le jour de la sortie de Digiposte, ou dès qu'un contenu concernant le produit est publié. Cette réactivité laissant présager par la suite la possibilité d'obtenir rapidement (en « temps réel ») des informations sur les sujets traités par les profils sélectionnés ;

– ajout de métadonnées au tweet (qu'il soit original ou repris), tels les *hashtags* (#) permettant de relier des sujets similaires sur Twitter, ou plus simplement l'ajout

d'un commentaire. Cela nous permettant d'appréhender une forme de redocumentarisation directe.

De ce premier tri sont ressortis 61 profils possédant au moins un des critères. Par la suite, nous avons voulu produire une typologie de ces internautes à partir de leurs attributs identitaires. L'objectif étant de définir (sur le plus long terme) s'il était possible d'observer une certaine consistance comportementale « impliquant une cohérence des décisions à venir avec les décisions prises par le passé » [BOU 10] dans la manière de traiter l'information des acteurs en fonction de cette typologie. Au-delà de cette observation dans la durée, l'intérêt premier de cette typologie était de pouvoir identifier et classer des profils en fonction de l'apport de leurs attributs identitaires à la qualification de l'information qu'ils diffusent. Les attributs choisis étaient au nombre de 44 au total. Nous ne souhaitons pas donner ici la totalité de ces attributs, et ce pour plusieurs raisons : d'une part, le choix de ces critères reposait sur des objectifs propres à l'entreprise et au contexte de la recherche, ce qui serait trop long à détailler (et perdrait du sens) ; d'autre part, notre propos n'est pas ici de fournir une méthodologie reproductible hors contexte, mais bien de définir le mécanisme par lequel certaines observations ont pu être effectuées.

Le tableau 2.2 donne une liste de 10 attributs principaux (en plus de ceux choisis dans la phase de sélection). Ils sont présentés en fonction de la forme d'identité à laquelle ils font appel, ainsi que de l'apport en termes de qualification de l'information qu'ils peuvent offrir à l'entreprise et au professionnel.

Attributs	**Identité**	**Apports**
Personne physique ou morale	Déclarative	Permet de différencier ce qui motive la diffusion d'informations : valorisation de l'entreprise d'un côté, reconnaissance de ses compétences propres de l'autre
Pseudonyme ou identité civile	Déclarative	L'identité civile laisse supposer une meilleure identification de l'autorité informationnelle de l'auteur, et engage ses actes au-delà de son compte Twitter
Biographie	Déclarative	La présence d'une biographie offre la possibilité d'identifier le statut que l'internaute se donne (de nombreux profils observés se décrivant comme « expert »), ainsi que les thématiques qu'ils souhaitent aborder sur leurs comptes
Nombre d'abonnements	Calculée	Ce volume souligne la capacité du profil à recouper son information et à se tenir informé par le réseau
Nombre d'abonnés	Calculée	Ce volume souligne l'attrait et la reconnaissance/ notoriété du compte

Types d'abonnements	Agissante	Permet d'élargir l'intégration de sources, mais aussi de définir si celles-ci correspondent aux attentes de l'entreprise
Types d'abonnés	Calculée	Identification du public du profil, afin de définir notamment si ce public correspond à celui recherché par l'entreprise. Et si dans ce public se trouvent d'autres profils jugés comme « experts »
Volume de tweets moyen par jours	Calculée et redocumentarisée	Afin de définir si le volume d'informations diffusées peut s'intégrer aux capacités de traitement du système de veille
Part de conversations versus part d*e tweets* informatifs	Usage et calculée	Cette différenciation permet d'identifier les profils s'axant principalement sur la diffusion d'informations (et qui ne semblent pas réceptifs aux interactions de leur réseau), et ceux plus conversationnels (ouvrant ainsi la possibilité de débattre ou préciser certaines informations reçues)
Volume de liens dans les *tweets*	Calculée et d'usage	Afin de définir les profils s'axant principalement sur la mise à disposition de ressources, de ceux utilisant la plate-forme comme outil « post-it » ou de conversation

Tableau 2.2. *Attributs identitaires, type d'identité et apports en termes d'analyse de l'information*

Ces attributs ne sont donc qu'indicatifs (et s'associent à ceux présentés dans la première phase de sélection), nous pourrions, par exemple, y ajouter les types de sources (presse en ligne, blogs, etc.) privilégiées en fonction des attentes de l'entreprise.

Par la suite, des « panels » ont été constitués en associant certains de ces critères en fonction des attentes de l'entreprise notamment (établir un contexte), mais aussi des formes d'autorité pouvant se dégager de ces combinaisons d'attributs (établir des critères de sélection de l'information diffusée), et enfin dans l'optique d'observer en quoi ces attributs peuvent modifier l'analyse de l'information collectée. Voici trois exemples issus de ces typologies.

Les « experts »[13] (15 profils) regroupent les attributs suivants : personne physique ; identité civile affichée ; un plus grand volume d'abonnés que d'abonnements, des abonnements généralement similaires en termes d'attributs identitaires (d'autres « experts ») ; une part égale de conversations et de tweets informatifs ; un fort volume de liens dans les tweets informatifs ; un recours quasi

13. Les dénominations ne sont ici qu'indicatives.

systématique aux *hashtags* afin de relier leurs messages à des thématiques discutées sur le réseau. Les experts diffusent généralement une information dès sa parution. Le fait que ces profils soient identifiables au-delà de leur présence sur Twitter permet d'établir plus largement leur notoriété, leurs autorités énonciatives et de contenus (la majorité des profils identifiés proposant un lien vers leurs blogs ou d'autres plates-formes). Le fait qu'ils soient suivis par un grand nombre de personnes et que parallèlement ils en suivent que peu, démontre une certaine reconnaissance de la part d'un public pouvant s'intéresser à leurs thématiques de prédilections. Contrairement à d'autres profils, ils abordent des sujets assez larges (les technologies Web ou le marketing mobile, par exemple), et conversent avec certains de leurs abonnés ou de leurs abonnements, répondant à des questions posées directement. Le fait qu'ils aient recours à certaines métadonnées, ou qu'ils enrichissent leurs messages de commentaires, permet de faire le lien au niveau de l'analyse avec d'autres informations diffusées dans le réseau. Voire, dans certains cas, de définir la crédibilité portée à une information par « l'expert » (par exemple : « ces chiffres sont faux, une autre étude ici »). Les informations qu'ils diffusent sont généralement facilement accessibles depuis les outils de veille (peu de sources « de niche »). La teneur de leurs informations est généralement « grand public » (vulgarisation scientifique, par exemple), et le fait qu'ils diffusent des contenus dès leur parution (pour développer la reconnaissance de leur expertise sur un sujet ?) permet au professionnel de la veille d'identifier rapidement certaines grandes tendances sur un sujet (voire les sujets ou contenus qui peuvent « faire événement »).

Les « collectionneurs » (11 profils) regroupent les attributs suivants : personne physique ; pseudonyme ; un faible volume d'abonnements et un volume d'abonnés variables ; quasi exclusivement des *tweets* informatifs dont la totalité comporte un lien ; ~~ils~~ produisent la plupart des messages qu'ils envoient (peu de reprises) ; un très fort volume de *tweets* envoyés chaque jour (comparativement à l'ensemble des profils observés) ;un recours quasi systématique aux métadonnées de type #. Les collectionneurs sont au cœur de la notion d'agents-facilitateurs puisqu'ils conversent très peu et se focalisent sur la diffusion intensive d'informations. Ils abordent généralement des thématiques assez précises, voire de niche (les systèmes d'exploitation pour téléphones mobiles, par exemple). Le fait qu'ils diffusent un grand nombre d'informations permet d'obtenir une vision précise des sujets qu'ils traitent, mais nécessite un travail accru de qualification des sources et des informations. Les abonnés sont généralement des spécialistes des sujets que ces profils traitent, et leur nombre restreint (ainsi que le fait que ces « collectionneurs » soient sous pseudonymes) laisse supposer que leur notoriété ou autorité est propre à la plate-forme Twitter (la majorité des profils ne faisant pas de liens vers d'autres comptes, ou alors vers des comptes d'agrégation de contenus, type « curation »). De même, leur faible volume d'abonnements amène à constater qu'ils vont généralement chercher leurs informations et leurs sources hors de la plate-forme, permettant ainsi de favoriser la détection d'informations mal référencées par les

outils traditionnels (ce qui, après observation, est le cas : par exemple, près de 40 % des informations diffusées n'ont pas pu être collectées par des outils classiques d'alertes mails). La redocumentarisation constante qu'ils effectuent dans leurs messages permet d'établir une cohérence dans les informations qu'ils diffusent, et ainsi de mettre en contexte certaines informations pouvant paraître disparates de prime abord pour un observateur extérieur.

Un dernier exemple est celui des « veilleurs mosaïques » (10 profils). L'aspect le plus proéminent est leur propension à effectuer un grand nombre de re-tweets, et d'avoir de ce fait un plus grand nombre d'abonnements que d'abonnés. Généralement présents sous leurs identités civiles, ils traitent des thématiques larges, et produisent peu de messages par jour. Si le fait qu'ils aient un faible volume d'abonnés peut questionner quant à leur légitimité ou la reconnaissance que leur portent les autres membres de la plate-forme traitant de la thématique (et que par conséquent ils dialoguent peu avec d'autres membres), l'intérêt de leur intégration à un processus de veille repose sur leur capacité à identifier de nouvelles sources. Pour les thématiques voulues, ils permettent alors de mettre en avant les profils leur paraissant pertinents sur ces thématiques. Il est d'ailleurs intéressant de souligner que ces profils ont été pour la plupart identifiés après avoir rediffusé le message d'un profil « expert ».

Cette étude étant exploratoire, elle nécessite certains développements en cours ou à venir. Notamment l'observation de la manière dont ces informations sont utilisées par l'entreprise ; l'identification de profils pour d'autres thématiques et dans un autre contexte (par exemple nouveau produit) ; la prise en compte de l'évolution des pratiques informationnelles des profils observés (pratiques qui, si elles présentent de prime abord certaines consistances, peuvent changer rapidement) ; développement d'un processus inverse consistant à construire un « panel » en fonction des attributs identifiés afin d'observer si dans un contexte plus large les attributs dénotent systématiquement les mêmes modes de traitement de l'information ; ou encore, une comparaison plus poussée avec certains outils présents sur le marché et visant eux aussi à « catégoriser » les internautes en fonction de leurs attributs identitaires en ligne. Au final, il nous semble que la notion d'agent-facilitateur, et la mise en corrélation de certains attributs identitaires saillants pouvant faciliter l'appréhension des informations diffusées sur un Web toujours plus déstructuré, ne doit pas amener à la création de catégories fermées et reproductibles hors contexte, mais bien à la construction d'indicateurs de pertinences au service des stratégies informationnelles des individus et des organisations.

2.7. Conclusion

Les identités numériques s'avèrent être un champ d'étude et de réflexion complexe, tant elles englobent une multitude de modélisations et d'approches possibles : représentation de soi, reconnaissance, visibilité, traces, réputation, etc. Dans ce chapitre, nous avons voulu nous concentrer sur une approche instrumentale de cette notion d'identité numérique, nous questionnant sur la manière dont les attributs identitaires d'internautes-prescripteurs d'informations (et que nous nommons agents-facilitateurs) sur des plates-formes comme Twitter pouvaient offrir à une organisation une forme de congruence informationnelle lui permettant par la suite de structurer son environnement. Les premiers résultats d'une étude exploratoire menée sur le sujet montrent que des attributs identitaires en ligne, similaires à certains groupes d'acteurs, permettent d'identifier des consistances comportementales quant au traitement effectué de l'information. Comportements informationnels dont l'incorporation à un processus plus large de veille stratégique permet par la suite de collecter des informations dont la mise en contexte issue de ces attributs ou effectuée par les acteurs eux-mêmes favorise l'intégration de ces informations aux discours et stratégies communicationnelles développés par l'organisation.

Mais aborder les attributs identitaires en ligne comme source de structuration de l'information diffusée sur le Web suppose alors de s'interroger sur la notion d'autorité informationnelle, sur sa mutation en notoriété ou encore en influence. Car le Web produit ses propres indicateurs, les algorithmes qui dirigent les plates-formes constituent leurs propres critères de hiérarchisation, s'appuyant sur l'identité numérique des internautes pour construire ces règles de visibilité, de notoriété ou d'autorité, autant que pour s'appuyer sur les dites règles afin de fournir à leurs utilisateurs une information toujours plus personnalisée.

Dans un environnement numérique où se développent progressivement de véritables « bulles informationnelles » [PAR 11], où l'identité et les données qui l'accompagnent deviennent la principale valeur financière de nombreux acteurs économiques, la question de l'interprétation et de l'utilisation des identités numériques à des fins aussi bien stratégiques (pour les organisations ou les internautes-consommateurs de contenus) que de vie privée reste entière. Et, le Web étant en constante évolution depuis ses débuts, il nous semble que cette question sera constamment en reformulation…

2.8. Bibliographie

[ALL 10] ALLOING C., MOINET N., « Des réseaux d'experts à l'expertise 2.0. Le Web 2.0 modifie-t-il la création et la mise en place de réseaux d'experts ? », *Les Cahiers du numérique*, vol. 6, n°1, 2010.

[ALL 11a] ALLOING C., DESCHAMPS C., « Veille stratégique et Internet participatif : les usages des agents facilitateurs remettent-ils en question le concept de signal faible ? », *Colloque spécialisé en Sciences de l'information : management de l'information, défis et tendances*, Université de Moncton, Canada, juin 2011.

[ALL 11b] ALLOING C., « Curation sociale et agents facilitateurs : quel(s) impact(s) sur les stratégies d'information et de communication des organisations sur le Web ? », *Médias 011 : Y a-t-il une richesse des réseaux*, Aix-en-Provence, France, décembre 2011.

[ALL 12a] ALLOING C., PIERRE J., « Construire un cadre d'analyse avec les SIC pour comprendre les pratiques et les enjeux de la réputation en ligne (des individus et des organisations) », *XVIII^e Congrès de la Société française des Sciences de l'Information et de la Communication*, Rennes, France, 30 mai-1er juin 2012.

[ALL 12b] ALLOING C., « De surveiller à « prendre soin »: comment repenser la veille sur les réseaux sociaux numériques en termes de management de réseaux d'acteurs ? », *Revue internationale d'Intelligence économique* (R2IE), 2012.

[BOU 85] BOURDIEU P., « The forms of capital », dans J.G. Richardson (dir.), *Handbook of theory and research for the sociology of education*, p. 241-258, Greenwood, New-York, 1985.

[BOU 10] BOUTET C-V., AMOR S.B., « Vers l'active SEO 2.0 », *Les Cahiers du numérique*, vol. 6, n°1, p. 179-198, 2010.

[BOY 10] BOYD D., SCOTT G., « Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter », *HICSS-43, IEEE*, Kauai, Hawaii, 6 janvier 2010.

[BOY 11] BOYD D., CRAWFORD K., « Six Provocations for Big Data », *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, 2011.

[BRO 03] BROUDOUX E., « Autoritativité, support informatique, mémoire », *Journées hypertextes, mémoire, fiction*, Montréal, Canada, novembre 2003.

[BRO 07] BROUDOUX E., « Construction de l'autorité informationnelle sur le Web », article en ligne : http://archivesic.ccsd.cnrs.fr/sic_ 00120710_v1, 2007.

[BUR 92] BURT R.S., *Structural Holes. The social structure of competition*, Harvard University Press, Cambridge (MA), 1992.

[CAR 08] CARDON D., « Le design de la visibilité. Un essai de cartographie du Web 2.0 », *Réseaux*, vol. 6 n°152, p. 93-137, 2008.

[CAR 10] CARDON D., *La démocratie Internet*, Le Seuil, Paris, 2010.

[COU 11] COUTANT A., STENGER T., « Production et gestion d'attributs identitaires », *Les Cahiers du numérique*, vol. 7, n°1, p. 61-74, 2011.

[DEN 11] DENOUËL J., « Identité », *Communications*, n°88, p. 75-82, 2011.

[ERT 09] ERTZSCHEID O., « L'homme est un document comme les autres : du World Wide Web au World Life Web », *Journal of Language and Communication Studies*, *Hermes* 53, p. 33-49, 2009.

[ERT 11] ERTZSCHEID O., « Vers une sanctuarisation du search », blog *Affordance.ino*, affordance.typepad.com/mon_weblog/2011/12/sanctuarisation-du-search-.html, publié le 12 décembre 2011.

[ESP 98] ESPELAND W.N., STEVENS M.L, « Commensuration as a Social Process », *Annual Review of Sociology*, vol. 24, p. 313-343, 1998.

[FRE 79] FREEMAN L.C., « Centrality in social networks : conceptual clarifications », *Social Networks*, vol. A, p. 215-239, 1979.

[GAL 11] GALINON-MELENEC B. (dir.), *L'Homme trace. Perspectives anthropologiques des traces contemporaines*, CNRS éditions, Paris, 2011.

[GEO 07] GEORGES F., Sémiotique de la représentation de soi dans les dispositifs interactifs. L'hexis numérique, Thèse de doctorat, Université Paris 1 Panthéon-La Sorbonne, 2007, http://tel.archives-ouvertes.fr/tel-00332747.

[GEO 08] GEORGES F., « Les composantes de l'identité dans le Web 2.0, une étude sémiotique et statistique. Hypostase de l'immédiateté », *Actes du76e congrès de l'ACFAS : Web participatif : mutation de la communication, Institut national de la recherche scientifique*, Québec, Canada, 6-7 mai 2008.

[GEO 09] GEORGES F., « Représentation de soi et identité numérique. Une approche sémiotique et quantitative de l'emprise culturelle du Web 2.0 », *Réseaux*, n° 154, p. 165-193, 2009.

[GEO 10] GEORGES F., « Pratiques informationnelles et identités numériques », *Etudes de communication* 35, p. 105-120, 2010.

[GRA 73] GRANOVETTER M.S., « The strength of weak ties », *American Journal of Sociology*, vol. 78, 1973.

[GRA 10] GRANJON F., DENOUËL J., « Exposition de soi et reconnaissance de singularités subjectives sur les sites de réseaux sociaux », *Sociologie*, vol. 1, p. 25-43, 2010.

[ITE 08] ITEANU O., *L'identité numérique en question*, Eyrolles, Paris, 2008.

[LED 06] LE DEUFF O., « Autorité et pertinence *versus* popularité et influence : réseaux sociaux sur Internet et mutations institutionnelles », 2006, en ligne : http://archivesic.ccsd.cnrs.fr/sic_00122603.

[LES 01] LESCA, H., « Veille stratégique : passage de la notion de signal faible à la notion de signe d'alerte précoce », *Actes du colloque, t.1, VSST 2001*, Barcelone, Espagne, 2001.

[LES 02] LESCA H., BLANCO S., « Contribution à la capacité d'anticipation des entreprises par la sensibilisation aux signaux faibles », *6e Congrès international francophone sur la PME*, HEC, Montréal, Canada, octobre 2002.

[LET 11] LE THRAN, D.K., CHEUNG-MON-CHAN, P., BOTHOREL, C., Conception et développement de fonctionnalités innovantes liées à Facebook pour un système de recommandation, Rapport, Telecom Bretagne, 2011.

[MAR 07] MARCON C., « Analyse de réseaux en intelligence économique : éléments pour une approche méthodologique », *Market Management*, vol. 7, n°4, p. 110-134, 2007.

[MAR 12] MARCON C., ALLOING C., « Quelle(s) perception(s) des classements de blogs par leurs auteurs ? Le cas Wikio », *Colloque international : la médiatisation de l'évaluation*, Paris, France, 15-16 mars 2012.

[MER 09a] MERZEAU L., « De la surveillance à la veille », *Cités* 39, PUF, Paris, 2009.

[MER 09b] MERZEAU L., « Présence numérique : les médiations de l'identité », en ligne : http://archivesic.ccsd.cnrs.fr/docs/00/48/32/93/PDF/ENJEUX-MERZEAU.pdf, 2009.

[MER 10] MERZEAU L., « La présence plutôt que l'identité », *Documentaliste – Sciences de l'Information* 47, 1, p. 32-33, 2010.

[MON 11] MONSEIGNE A., GALINON-MELENEC B., « La sémiotique des 'signes-traces' appliquée au recrutement : le cas de la recherche du 'bon candidat' *via* les traces numériques », *Communication et organisation*, 2011.

[MUS 06] MUSSER J., O'REILLY T., *Web 2.0: principles and best practices*, O'Reilly Media, 2006.

[PAR 11] PARISER E., *The Filter Bubble: What the Internet Is Hiding from You*, The Penguin Press, New-York, 2011.

[PIE 11] PIERRE J., « Génétique de l'identité numérique » Sources et enjeux des processus associés à l'identité numérique », *Les Cahiers du numérique*, vol. 7, n°1, p. 15-29, 2011.

[QUE 97] QUERE L., « La situation toujours négligée ? », *Réseaux*, vol.15, n°85, p. 163-192, 1997.

[RIE 10] RIEDER B., « Pratiques informationnelles et analyse des traces numériques : de la représentation à l'intervention », *Etudes de communication* 35, 2010.

[SER 04] SERRES A., « Evaluation de l'information sur Internet », *BBF*, n°6, p. 38-44, 2005, http://bbf.enssib.fr/consulter/bbf-2005-06-0038-006.

[SCH 87] SCHÜTZ A., « Le monde social et la théorie de l'action sociale », *Le chercheur et le quotidien*, p.95, Méridien Klincksieck, Paris, 1987.

[STE 09] STENGER T., COUTANT A., « La prescription ordinaire sur les réseaux socionumériques », *Médias 09, entre communautés et mobilité*, Aix-en-Provence, France, 2009.

[STE 10] STENGER T., COUTANT A., « Les réseaux sociaux numériques : des discours de promotion à la définition d'un objet et d'une méthodologie de recherche », *Journal of Language and Communication Studies*, *Hermes*, 44, 2010.

[STI 09] STIEGLER B., *Pour une nouvelle critique de l'économie politique*, coll. Débats, Galilée, Paris, 2009.

[VIS 12] VISEUR R., « Le chercheur peut-il se fier aux volumétries indiquées par les moteurs de recherche commerciaux ? », *Séminaire M@rsouin 2012 Usages du numérique*, Brest, France, juin 2012.

[ZAC 06] ZACKLAD M., « Une approche communicationnelle et documentaire des TIC dans la coordination et la régulation des flux transactionnels », Document de travail en ligne : http://archivesic.ccsd.cnrs.fr, 2006.