



HAL
open science

Du bruit au signal

Baptiste Kotras

► **To cite this version:**

Baptiste Kotras. Du bruit au signal: Agencements et savoirs locaux dans les tableaux de bord d'e-réputation. *Terrains et Travaux: Revue de Sciences Sociales*, 2015, 1 (26), pp.127-145. halshs-01264349

HAL Id: halshs-01264349

<https://shs.hal.science/halshs-01264349>

Submitted on 29 Jan 2016

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.

Du bruit au signal

Agencements et savoirs locaux dans les tableaux de bord d'e-réputation

Baptiste Kotras
LATTS/Université Paris-Est
bkotras [at]gmail [point] com

Résumé

A l'heure où un ensemble de logiciels proposent aux marques la mesure de leur réputation sur le web, nous interrogeons les processus par lesquels ces tableaux de bord permettent effectivement d'agréger et de qualifier les évaluations dont les acteurs économiques font l'objet sur le web. Ces outils font de fait face à un défi méthodologique majeur consistant à convertir le « bruit » en « signal », c'est-à-dire à rendre intelligibles des données verbales préexistantes, massives et hétérogènes. A travers l'étude de prestataires français et américains, nous montrons comment la façon dont sont conçus puis vendus ces tableaux de bord les rend perméables aux épistémologies et aux cadres cognitifs des organisations clientes. En demandant à l'utilisateur un constant travail d'agencement et d'interprétation des résultats algorithmiques, les procédures techniques et commerciales qui accompagnent le déploiement des tableaux de bord permettent ainsi la coïncidence entre les problématiques des organisations clientes et le flux de la parole en ligne.

Mots-clés : tableau de bord ; réputation en ligne ; classification ; quantification ; sémantique.

Abstract

While an array of software offer to monitor the brands' reputation on the web, we question the processes by which these dashboards actually aggregate and qualify the judgments made by people on the web about companies and products. These tools address in fact a major methodological challenge: converting the "noise" into a "signal", that is to say, to make intelligible preexisting, massive and heterogeneous verbal data. Through the study of French and American providers, we show how the way they are designed and sold make these dashboards permeable to epistemologies and cognitive frameworks of the client. By demanding the user constant arrangement and interpretation of algorithmic results, the technical and commercial procedures that come with the dashboards allow the matching between the client's questions and the flow of online speech.

Keywords: dashboard; online reputation; classification; quantification; semantic.

Depuis le début des années 2000, les formes multiples de l'évaluation profane sur le web (avis, arguments, commentaires, notations, etc.) ont pris une importance croissante, redistribuant les positions et les relations dans des domaines d'activité variés, qu'il s'agisse de la production du savoir encyclopédique (Cardon et Levrel 2009), de la hiérarchisation de l'information (Cardon 2013), de l'évaluation des biens et des services (Beauvisage et al. 2013) ou encore des trajectoires artistiques (Beuscart 2008). Dans chaque cas, nous retrouvons à l'œuvre des dispositifs qui, en agrégeant et en rendant visibles les évaluations des internautes, leur confèrent une *agency* spécifique sur les espaces sociaux concernés. Ainsi se trouve posée à nouveaux frais la question sociologique de la réputation, entendue comme « *représentation sociale partagée, provisoire et localisée, associée à un nom et issue d'évaluations sociales plus ou moins puissantes et formalisées* » (Chauvin 2013 : 132). En effet, la formalisation croissante des dispositifs et du travail réputationnels, notamment dans le monde marchand, impose d'interroger les traces, les métriques et les outils, profondément réflexifs, qui réalisent ce travail de centralisation et de mise en calcul d'évaluations éparées, contribuant ainsi à reconstruire la réputation des individus et des organisations.

Parmi la grande variété de ces dispositifs, nous nous intéresserons aux logiciels de veille qui, conformément à l'injonction décrite par Power et al. (2009), proposent depuis le milieu des années 2000 aux grandes marques la gestion de leur réputation en ligne, et notamment du « risque réputationnel » né de l'expression massive des consommateurs sur le web. Ces éditeurs logiciels offrent pour cela de tirer parti des traces d'opinion contenues dans les discussions quotidiennes des internautes (sur les forums, blogs, sites de réseaux sociaux, etc.) pour mesurer en temps réel la réputation des organisations marchandes, et informer leur action stratégique sur celle-ci¹. Dans le cadre d'une « économie de l'opinion » (Orléan 1999), où la réputation constitue un actif à valoriser et à défendre, la promesse de ces prestataires logiciels consiste donc à restituer au client la possibilité de gouverner son image de marque, par la scrutation exhaustive et en temps réel des milliers de traces discursives laissées chaque jour par les internautes. L'utopie technique qui imprègne le discours commercial est celle d'instruments qui, « branchés » sur les flux de publications du web social, donneraient un aperçu immédiat et transversal de la réputation des marques, jusqu'ici éparpillée en une infinité de lieux et de moments². De fait, les outils vendus sur ce marché adoptent tous ou presque la forme synoptique du « tableau de bord », ou *dashboard* : en fonction des mots-clés surveillés, l'utilisateur accède à un fil de retombées en temps réel, ainsi qu'à une batterie d'indicateurs, de graphes et de scores qui doivent décrire les données conversationnelles brutes et en déduire la réputation.

Ces logiciels de veille peuvent être analysés comme des « *scoping systems* » (Knorr-Cetina et Grimpe 2008), dans la mesure où ils produisent une description globale et autoréférencée du phénomène observé – ici la réputation en ligne. D'autre part, ce sont aussi et surtout des dispositifs intrinsèquement réflexifs, dont les résultats orientent la production de la connaissance et la coordination de l'action. Suivant le programme général de la sociologie de la quantification (Desrosières 2008), il apparaît alors indispensable d'interroger les conditions dans lesquelles se réalise un projet visant à centraliser et agréger les traces éparées de l'opinion en un nouvel artefact – la « réputation » – à son tour susceptible de calculs et de déplacements ultérieurs (Callon et Muniesa 2003).

¹ Pour un état de l'art, voir Boullier et Lohard (2012).

² On perçoit bien cet idéal d'exhaustivité et d'instantanéité dans les étymologies des noms commerciaux qui peuplent ce marché, qu'il s'agisse de ceux des éditeurs (Synthesio, Brandwatch) ou de leurs logiciels (Radarly, Buzz, HeartBeat).

L'intérêt spécifique des données conversationnelles en ligne réside d'abord dans leur caractère spontané : parce qu'elles ont été émises indépendamment de tout dispositif d'enquête, celles-ci évitent l'écueil de l'artificialité souvent reprochée aux résultats des sondages (Bourdieu 1973). Cette spontanéité présente cependant l'inconvénient de fournir des données massives et hétérogènes, qui ne « répondent » en soi à aucune question. L'utilisateur du tableau de bord travaille sur des données qui n'ont pas été produites pour lui ; en inférer une réputation globale suppose donc de les domestiquer, de convertir le « bruit » en « signal » intelligible et interprétable par les marques clientes.

Comment les tableaux de bord et les acteurs qui les déploient parviennent-ils à articuler ces deux univers *a priori* étrangers, celui de données verbales massives et hétérogènes, et celui des problématiques des marques qui sont leurs clients ? En décrivant les situations concrètes de travail avec les données (Dagiral et Peerbaye 2012), nous montrerons que cette articulation repose sur la négociation et l'inscription de savoirs locaux et orientés vers l'action dans la forme générique du tableau de bord. Loin de donner un accès immédiat à la réputation, ses résultats sont constamment réagencés et interprétés en fonction des buts opérationnels ponctuels de l'utilisateur. Grâce à sa perméabilité aux épistémologies locales (Dagiral et Parasie 2015), le tableau de bord thématise et problématise des traces d'opinion jusqu'ici présentes à l'état latent, permettant *in fine* la mesure de la réputation, c'est-à-dire la reconstruction du public en ligne de la marque (Gillespie 2014).

Nous mobiliserons pour cela une série d'entretiens conduits entre mai 2013 et février 2014 auprès de différents éditeurs, à Paris et à San Francisco³, en étudiant notamment le rôle du « *customer success manager* » (CSM), fonction centrale dans le modèle économique du *software as a service* dont Campbell-Kelly et Garcia-Swartz (2007) ont montré la résurgence récente. Ni développeur, ni agent de vente, ce professionnel est en charge d'un portefeuille de clients dont il est l'interlocuteur unique une fois le contrat signé. Il agit ainsi comme un traducteur entre les problèmes de chaque client et les fonctionnalités du logiciel ; très bon connaisseur de son outil, pédagogue et prescripteur des « bonnes pratiques », réparateur occasionnel et réceptacle des enthousiasmes et des frustrations du client, le CSM constitue donc un excellent moyen d'accéder à la matérialité du travail réputationnel et de son articulation avec les dispositifs automatisés. Nous montrerons dans un premier temps comment le CSM et son client, lors du paramétrage initial du logiciel, construisent une architecture sémantique de façon à circonscrire le périmètre de la réputation de marque, et à le diviser en unités de sens intelligibles. Dans un second temps, nous étudierons la façon dont les diverses métriques proposées par les plateformes doivent être combinées et réinterprétées par l'utilisateur pour faire émerger une description des publics de la réputation en ligne.

Du mot-clé à la catégorie, construction d'une architecture sémantique réflexive

Les mots-clés renseignés lors du paramétrage du tableau de bord constituent de fait le seul point de contact entre l'utilisateur et les flux de conversations qu'il veut mesurer, et dont la captation dépend intégralement. L'activité consistant à nommer, circonscrire et décrire les mots qui composent la réputation en ligne est en cela essentielle à la construction de l'espace de sa mesure. Nous montrons ici comment le CSM travaille à la réflexivité de son client, dans

³ Nous mobiliserons sept entretiens, menés auprès de cinq éditeurs distincts (deux français, trois américains). Si les produits et les prestations sont relativement identiques dans les deux pays, les acteurs principaux en termes de chiffre d'affaires se trouvent aux Etats-Unis, et jouent un rôle prescripteur pour les acteurs européens ; le métier de CSM notamment, est issu du marché nord-américain avant de devenir un standard sur le marché français également.

la construction des requêtes, puis des catégories qui définissent l'architecture sémantique d'un projet de veille.

La construction des requêtes : déterminer l'univers sémantique

Si les évaluations des consommateurs abondent sur le web, dans des formes techniques et linguistiques hétérogènes, les coordonner en une mesure de la réputation présuppose d'abord de nommer les entités qui la composent, acte qui donne *de facto* une existence et des contours précis à cette réputation. La portée performative de cette opération n'est pas sous-estimée par les éditeurs, qui consacrent une part essentielle de leur activité à la phase initiale de formation et de paramétrage de l'outil, dans laquelle est défini une fois pour toutes le plan de requêtes du projet de veille. Ici, le CSM intervient pour obliger son client à se positionner de façon réflexive quant à ce qu'il souhaite mesurer : faut-il inclure les noms des produits ? Des dirigeants de l'entreprise ? De ses concurrents ? Cette fonction proprement maïeutique constitue selon eux le cœur de leur métier : « *Vous savez, j'apprends à mes clients à tout faire. Je leur montre comment faire, je leur donne les moyens, j'agis comme leur guide. Ou un coach, quelqu'un qui fait office de conseiller* » (Joel, CSM, Engagor).

Dans ce travail de co-élaboration des requêtes, tous les enquêtés insistent sur la nécessaire problématisation, non acquise au départ, du rapport entre l'utilisateur et les flux qu'il collecte. En particulier, il est impératif de spécifier ce qui, dans la profusion des contenus impliquant potentiellement la marque, relève de sa réputation telle qu'elle est envisagée par le client. Ce travail de spécification de l'univers réputationnel de la marque apparaît le plus clairement dans les cas où le client est décrit comme peu familier de ce type d'enjeu, ou peu clair sur les objectifs qu'il souhaite poursuivre. Chercher à « savoir ce qui se dit » sur son entreprise est selon les CSM insuffisant pour démarrer un projet ; il leur faut définir des sujets d'intérêt qui thématisent et problématisent la veille par l'introduction de mots-clés spécifiques.

« La plupart des gens ne soupçonnent pas la généralité et l'ampleur des sujets. Si vous entrez 'Coca-Cola', le système va exploser, dans la mesure où ça va ramener beaucoup trop d'informations. Du coup, c'est notre travail de leur expliquer : 'ok, qu'est-ce que vous voulez savoir sur Coca-Cola ? Est-ce que vous voulez savoir ce que les consommateurs en disent ? Voulez-vous savoir s'ils sont en colère, ou s'ils l'aiment bien ? Voulez-vous savoir ce qu'ils disent sur les produits ? Quel aspect du marché vous intéresse, finalement ?' Nous avons de longues conversations avec le client sur le secteur, sur leurs stratégies, afin de comprendre la façon dont nous pouvons filtrer au mieux les requêtes » (Belén, CSM, Meltwater).

Si la réputation que l'on cherche à mesurer est celle, globale, de la marque, celle-ci n'existe matériellement que comme le cumul de mots-clés spécifiques qui définissent les thématiques de la veille. Le CSM joue un rôle intermédiaire entre les objectifs du client et ce qu'il anticipe, par expérience, le degré de spécificité et de pertinence des requêtes envisagées ; le cas échéant, il rectifie, précise, ajoute. La maîtrise de l'algèbre booléenne⁴ est ici essentielle, puisque c'est l'ajout de mots-clés et de conditions qui permet la composition de requêtes complexes, donc précises (cf. Fig.1).

⁴ Opérateurs logiques nécessaires au paramétrage des requêtes en recherche d'information, qui spécifient le lien entre les différents mots-clés de la requête (dans notre cas, essentiellement : « ou », « et », « sauf », et les parenthèses). Les requêtes les plus complexes remplissent l'équivalent de plusieurs pages au format A4.

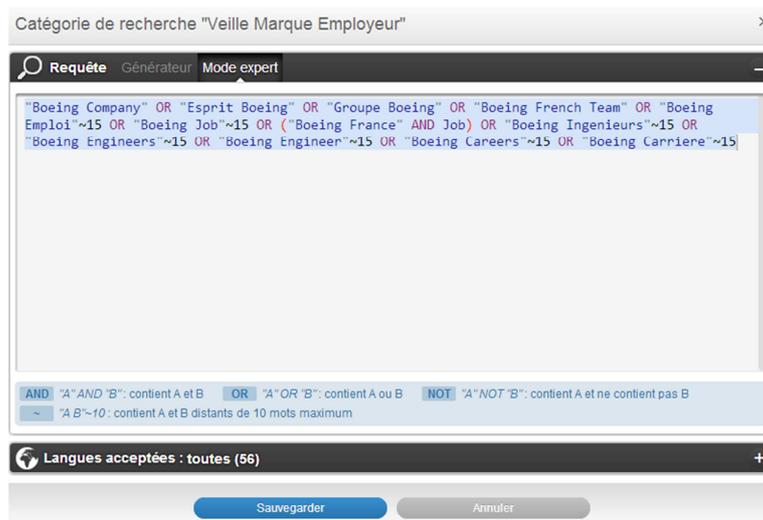


Fig.1. Exemple de requête sur le logiciel Radarly⁵

La nécessité de sélectionner un matériau linguistique surabondant est par ailleurs inscrite dans les logiciels eux-mêmes, qui proposent des fonctionnalités destinées à faciliter l’itération dans la définition des mots-clés. Le logiciel de l’éditeur Meltwater permet ainsi de tester une requête avant sa fixation définitive, en indiquant la volumétrie estimée, ainsi qu’un échantillon aléatoire de *verbatim* correspondant aux termes de recherche. De même, de nombreuses plateformes comme Synthesio affichent, pour une requête donnée, un « nuage de mots » correspondant aux termes les plus cités dans les résultats d’une requête. Un tel outil vise à mettre au jour l’univers sémantique du client, dans lequel le veilleur est alors invité à affiner sa requête par sérendipité :

« Vous vous dites : tiens, il y a quand même pas mal de gens qui parlent du design [d’un produit], donc vous allez vous dire finalement, je pourrais peut-être faire quelque chose de spécifique sur le design. [...] Quand vous rentrez un seul mot-clé, vous pouvez avoir un nuage de mots-clés associés, donc ça vous aide » (Thomas, dirigeant, Synthesio – Paris).

On voit ici comment le CSM aussi bien que le logiciel prennent en charge l’indétermination qui entoure le périmètre de la réputation de marque. Par le dialogue et les suggestions automatisées, l’usager est invité à faire des choix qui l’inscrivent dans un univers sémantique fixe. Ses inclusions et exclusions, comme le choix des entités nommées, définissent seuls les contours de l’univers que l’instrument doit mesurer.

Classifications et catégories : définir le sens et la granularité des agrégats

Outre le fait de nommer les entités qui la composent, la mesure de réputation en ligne implique également d’inscrire les données conversationnelles dans des catégories qui les décrivent et les rendent intelligibles. Lors du paramétrage, les tableaux de bord demandent en effet de regrouper *a priori* les requêtes sémantiques dans des catégories de veille où sont ainsi acheminés les *verbatim* qui correspondent à chaque requête (cf. Fig.2). Le nom, le nombre et l’étendue sémantique de chaque catégorie sont laissés à la libre appréciation de l’usager ; une fois constituée, cette architecture cognitive exerce un puissant effet descriptif en inscrivant systématiquement les données dans les catégories cognitives du client.

⁵ Source : <http://linkfluence.com/2014/07/18/innover-en-rh-avec-radarly/>

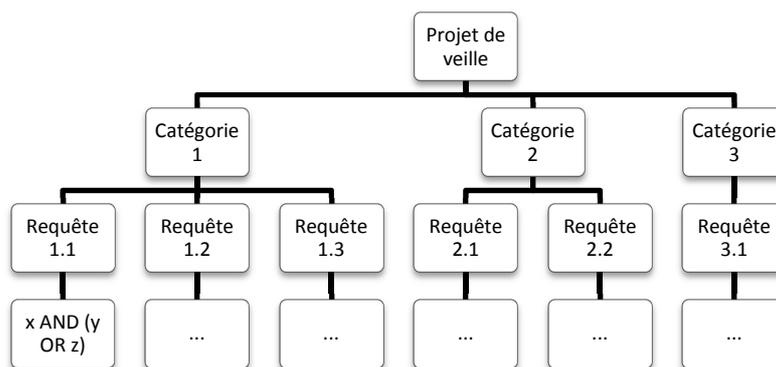


Fig. 2. Architecture sémantique d'un projet de veille

Au moment de leur captation, les données sont donc également catégorisées : les *verbatim* qui « *matchent* » l'une des requêtes du plan sont triés dans les catégories qui leur donnent sens. Un prestataire réalisant la veille pour un célèbre fast-food dispose par exemple d'une catégorie « santé », laquelle comprend une requête comportant les mots-clés « nausée », « indigestion », « malade », etc. De ce fait, le veilleur peut à tout moment isoler les *verbatim* de cette catégorie particulièrement sensible pour décider ou non d'alerter la direction ; les projets et préoccupations des entreprises s'inscrivent ainsi dans des catégories qui rendent possible l'action sur le réel (Bowker et Star 1999). Le codage systématique des données leur assigne un sens qui pour le client les rend significatives, donc gouvernables (Desrosières 2008).

« C'est vous qui définissez vous-même votre classification. Vous allez vous dire : je suis dans l'automobile, mon premier niveau, ça va être les marques : Renault, Peugeot, Citroën. Ensuite, mon 2^e niveau de classification, soit je peux me dire que je fais les modèles, mais je peux me dire aussi : tiens je vais me concentrer sur tel niveau – je dis n'importe quoi : les modèles électriques, le design, le prix, la sécurité » (Thomas, dirigeant, Synthesio – Paris).

En remplissant les catégories proposées par le tableau de bord, l'utilisateur réalise donc l'agrégation des *verbatim* en des totalités importées de son environnement commercial, technique, concurrentiel, etc. Dans tous les tableaux de bord étudiés, ces catégories sont les unités de base indispensables au calcul de la réputation, par les humains et par le logiciel. Elles permettent à l'utilisateur de se concentrer sur une dimension spécifique de sa réputation en ligne, ou au contraire de les comparer entre elles, selon qu'il souhaite par exemple lire tous les *verbatim* se rapportant à une question particulière, ou plutôt réaliser des classements, des oppositions entre les divers agrégats ainsi offerts à sa manipulation. De fait, dans la forme tableau de bord, la réputation de marque n'existe jamais en soi, pour soi : elle est toujours le fruit de ces manipulations intermédiaires permises par le balisage préalable de l'espace sémantique et cognitif de la veille.

Dès lors, la granularité de la classification ainsi élaborée fait l'objet un soin tout particulier dans l'interaction initiale entre le client et son conseiller. Dans la définition des catégories de la veille, les acteurs arbitrent constamment entre généralité et spécificité, dans la mesure où il leur faut construire des agrégats assez généraux pour réduire la polysémie de la parole en ligne, mais suffisamment spécifiques pour autoriser des manipulations fines, monographies et comparaisons. Dans cette délicate négociation, le savoir-faire incorporé par le CSM est là encore essentiel. Lors d'un entretien avec Jérémie, *account manager* pour le logiciel Radarly, celui-ci me montre le tableau de bord déployé pour le compte de la chaîne de télévision France 2. Je lui demande alors pourquoi il ne crée pas un tableau de bord à l'échelle du

groupe France Télévisions tout entier, par ailleurs client du logiciel, quitte à créer ensuite une catégorie de veille pour chaque chaîne du groupe. C'est selon lui une mauvaise idée :

« Ce serait trop le merdier parce que déjà, tu aurais énormément de retombées. Et puis à l'intérieur d'une catégorie, si tu as 'France 2 émissions', 'France 2 dirigeants', tu peux pas... Tu peux filtrer que par catégorie aujourd'hui. [...] Ça veut dire qu'il vaut mieux avoir des catégories très précises : c'est plus simple à analyser. [...] il vaut mieux faire des catégories 'France 2 *corporate*', 'France 2 émissions', etc. La thématique globale, ça reste France 2, à l'échelle du projet. C'est-à-dire qu'un projet, c'est un plan de requête cohérent » (Jérémy, *account manager*, Linkfluence).

Son expérience de la veille pour cet important client indique à notre enquête que ma suggestion aboutirait à des catégories insuffisamment granulaires pour l'intelligibilité du projet. Il serait complexe, voire impossible de « dire » quoi que ce soit d'une catégorie qui mêlerait des *verbatim* mentionnant aussi bien des émissions que des dirigeants de la chaîne France 2 ; le niveau de zoom serait insuffisant pour obtenir des agrégats significatifs. Au contraire, un plan de requêtes précis couplé à des catégories suffisamment granulaires permet un jeu plus fin sur les différents ordres de réalité décrit par la classification des *verbatim*. La granularité des catégories fait donc l'objet d'une pédagogie dédiée :

« Ce qu'on conseille à nos clients de faire, c'est de prendre un fichier Excel, un document Word, et de mettre tous les mots-clés qu'ils souhaitent écouter, et de les séparer par grandes thématiques. De faire une première séparation : qu'est-ce qui concerne le *corporate*, qu'est-ce qui concerne la pub, les dirigeants, les produits, etc. » (Jérémy, *account manager*, Linkfluence).

La phase de paramétrage du tableau de bord est donc un moment décisif, en ce qu'il fixe l'architecture sémantique de la veille. En lui demandant de nommer les entités pertinentes et de les regrouper par catégorie, le logiciel et le *customer success manager* poussent l'utilisateur à se déterminer, et donc à déterminer en retour l'univers sémantique, et donc cognitif, où sera calculée la réputation en ligne. Par itérations successives, le chargé de veille construit une architecture qui trie et qualifie des données conversationnelles *a priori* très nombreuses et polysémiques, rendant possible l'application ultérieure de filtres catégoriels. Paradoxalement, la forme voulue globale et synoptique du tableau de bord implique concrètement de circonscrire pour les distinguer des agrégats assez nombreux et assez fins pour permettre une manipulation heuristique.

Agencer et interpréter les indicateurs : des algorithmes agnostiques

Nous nous intéresserons ici à la deuxième phase de la veille réputationnelle sur le web : celle de l'interprétation des données. Pour cela, les logiciels offrent un ensemble de modules (« *analytics* ») qui mettent en calcul les *verbatim* collectés à partir de leurs métadonnées associées⁶. Ces boîtes noires incarnent par leur juxtaposition la forme « tableau de bord » en mettant l'utilisateur face à un ensemble de graphes, courbes et tableaux, chacun apportant une information précalculée et univoque.

⁶ Classiquement, on aura ainsi des courbes de volumétrie, des ventilations par support (blogs, réseaux sociaux, etc.), des classements d'« influenceurs », ou encore des baromètres de tonalité (positive, négative ou neutre) calculés par des algorithmes de traitement automatique du langage (TAL).

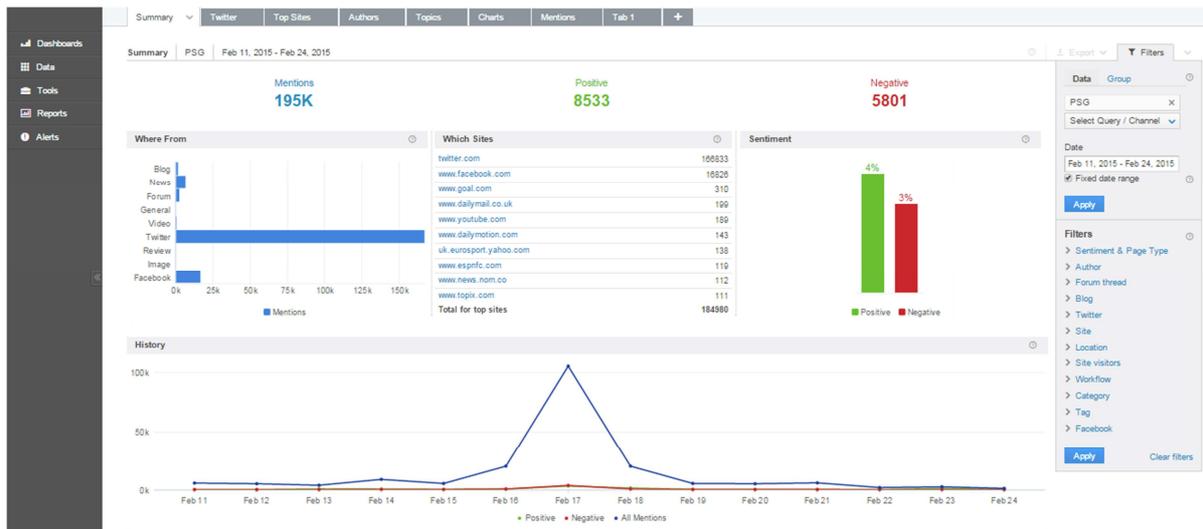


Fig. 3. Vue d'ensemble du logiciel Brandwatch

Nous allons montrer comment l'utilisateur est en fait invité à agencer entre eux ces objets hétérogènes, mais également à les réinterpréter, voire à en contourner le sens. La réputation en ligne n'émerge que de l'articulation entre ces artefacts, finalement plus agnostiques qu'univoques, et les épistémologies professionnelles de l'utilisateur.

Définir ses communautés de marque : labellisation manuelle et calcul de l'influence

Nous allons ici nous concentrer sur le module « communauté » proposé par Buzz, le logiciel de l'éditeur Meltwater (San Francisco). Bien que le terme de communauté prenne un grand nombre de sens sur ce marché, il renvoie dans ce cas précis à une acception élargie des « communautés de marque » décrites par le marketing, c'est-à-dire les clients affectivement attachés et investis dans la marque, relais potentiel pour des opérations de communication (Mellet 2012). Pour les identifier, la plateforme Buzz propose à l'utilisateur de connecter les comptes institutionnels de la marque sur Twitter et Facebook ; les *followers* et fans de la marque sur les réseaux sociaux sont alors listés dans un onglet spécifique en tant que « communauté ». A partir de cette page, il devient possible d'écouter spécifiquement ce « premier cercle » des consommateurs, dont il est ainsi présumé qu'il joue un rôle crucial dans la construction de la réputation en ligne. En soi, ce module communautaire ne donne pourtant pas assez de « prises » pour le calcul de ce public particulier (Gillespie 2014), et donc pour l'action. Parlant des pratiques d'un client, lui-même éditeur de logiciels, une CSM de Meltwater explique :

« Leur logiciel est principalement utilisé par des ingénieurs, mais les ingénieurs, ça peut travailler dans beaucoup de secteurs : ils peuvent être dans l'aéronautique, dans l'éducation, et ainsi de suite. Donc, ce qu'ils font c'est qu'ils marquent les membres de leur communauté en fonction de leur industrie. Si une personne publie à propos de [nom du client] et travaille dans l'aéronautique, elle sera marquée 'aéronautique'. Disons que vous avez cinq catégories, du coup vous pouvez... Vous avez les gens triés par catégorie, et vous pouvez cliquer sur l'étiquette 'aéronautique', pour voir tous les gens que vous avez sélectionné dedans » (Belén, CSM, Meltwater).

Malgré la présence d'un module identifiant automatiquement la « communauté » de la marque cliente, la plateforme offre de surcroît la possibilité de labelliser des sous-segments de cette communauté. Cependant, à la différence de la classification effectuée lors de la captation des données en temps réel, celle-ci s'applique aux comptes, aux « personnes », et non aux *verbatim* ; mais surtout, le codage doit cette fois être réalisé « à la main ». Le chargé de veille

cliquera sur chaque profil individuel pour lire les informations qu'il donne sur lui-même, et ses publications récentes ; de ces informations, il infèrera le domaine ou secteur d'activité pertinent pour chaque profil – déduction impossible pour un algorithme. On mesure donc la difficulté d'une telle opération lorsque la « communauté » identifiée comporte plusieurs milliers de membres. Pour y remédier, le module communautaire peut alors être combiné avec la mesure de l'influence individuelle des sources :

« [Nom du client] tend à se focaliser sur les gens qui ont beaucoup de *followers*, qu'ils vont considérer comme des influenceurs. Même si moi je les suivais, ils ne me considéreraient pas comme une influenceuse parce que je ne... Ce que je dis ne va toucher personne. Vous pouvez classer les résultats, les membres de la communauté, en fonction de leur influence. C'est une bonne manière de se dire : ok, je vais catégoriser ces personnes-là ».

La plasticité, et donc le potentiel heuristique du tableau de bord, repose sur le grand nombre de métadonnées associées à chaque source et chaque *verbatim*. Dans le cas de Meltwater, le logiciel calcule pour chaque membre de la communauté un score d'influence, essentiellement fonction de son propre nombre d'abonnés. De cette façon, l'utilisateur construit un scénario d'usage où seuls les comptes les plus « influents » se voient reconnaître un rôle dans la formation de la réputation en ligne, qui justifie le coûteux effort humain de labellisation sectorielle ; identification communautaire et mesure d'influence sont ici combinées de façon ponctuelle afin de trier le bon grain des influenceurs de l'ivraie des fans anonymes⁷. Les fonctions analytiques sont construites de telle manière qu'elles présentent entre elles des aspérités qui permettent leur association : dans ce cas précis, il s'agit de la possibilité de classer les *followers* de la marque par influence, et de leur associer manuellement des *tags* de façon à les distinguer entre eux. Ces trois fonctionnalités (communauté, influence, *tagging*) ne sont pas en soi destinées à être associées ; prises séparément, elles ne « disent » que peu de choses de la réputation de la marque. C'est seulement par leur combinaison ponctuelle, laissée libre par les développeurs du logiciel, que peut être isolé un public spécifique de la réputation de marque, à la fois influent et spécialisé par secteur.

L'originalité de ces logiciels de veille consiste donc à ne pas fixer *a priori* le script de l'usage (Akrich 1987). Au contraire, les divers modules sont livrés comme autant d'objets librement associables et révocables dans des scénarios d'utilisation dont la construction appartient à l'usager, assisté des recommandations de son conseiller. Nous rejoignons donc ici B. Rieder (2012) lorsqu'il décrit l'évolution des systèmes de données contemporains vers des modèles dits « relationnels », producteurs d'objets « *uniformes et faiblement connectés, pouvant être ordonnés selon un nombre potentiellement illimité de règles* »⁸. Il semble que pour les éditeurs de tableaux de bord, la meilleure manière de prendre en compte l'hétérogénéité des données consiste à ne déterminer ni les calculs, ni les usages qui en sont faits. Au contraire, les plateformes étudiées suggèrent, orientent, proposent, sans jamais enfermer les données dans des ontologies rigides ; l'usager y est enjoint à y inscrire ses catégories, ses objectifs, ses savoirs spécifiques.

Anticiper, enseigner et compenser les failles de l'algorithme : le cas du sentiment analysis

L'analyse de tonalité (*sentiment analysis*) est peut-être la fonctionnalité la plus emblématique des tableaux de bord d'e-réputation, cristallisant enthousiasme et controverses. Au moyen

⁷ Sur le rôle spécifique de la mesure d'influence comme critère d'échantillonnage sur le web, voir Kotras (2015).

⁸ Publié sur son blog *The Politics of Systems* (notre traduction).

d'algorithmes issus du *natural language processing*⁹, celle-ci classe chaque *verbatim* comme positif, négatif, ou neutre. De par l'ampleur de la réduction ainsi pratiquée sur le matériau verbal, ces algorithmes ont donc été hautement décriés (Fouetillou 2009), mais aussi vendus comme la panacée en matière d'e-réputation : avec ces outils, il devenait possible de déterminer instantanément la polarité de chaque prise de parole, et par agrégation celle de la réputation conférée aux marques sur le web. Aujourd'hui, les acteurs du marché entretiennent une position plus mesurée vis-à-vis de ces technologies, puisque clients comme prestataires se montrent régulièrement critiques envers une méthode non seulement binaire, mais aussi largement faillible¹⁰. Pour autant, tous les logiciels analysés présentent un module de *sentiment analysis*, qui étiquette les opinions collectées et présente des scores agrégés. L'omniprésence du *sentiment analysis* dans les tableaux de bord, *a priori* surprenante au vu de la confiance toute relative dont il fait l'objet, ne peut se comprendre que par la capacité de ce module à devenir « autre chose » que ce pour quoi il est théoriquement conçu. L'étude des situations concrètes d'usage de cet instrument montre en effet comment cette boîte noire par excellence est en fait sujette à l'interprétation, voire à la contradiction, de la part d'acteurs peu enclins à lui déléguer entièrement leur réflexivité. Comme nous allons le montrer, cette distance critique envers l'indicateur est là encore permise et encouragée, à la fois par l'architecture du logiciel et par ses promoteurs.

En effet, nous avons pu constater comment, face à des clients prompts à s'inquiéter au vu des pourcentages annoncés de billets « négatifs », leurs interlocuteurs déploient une véritable pédagogie de la faillibilité, invitant à considérer non pas la proportion absolue de positifs et de négatifs, toujours imparfaite, mais sa variation dans le temps. Travis, *sales manager* pour Sysomos, explique : « *tout le monde sait que ce n'est jamais à 100% exact. Mais si, pour une raison ou pour une autre, on a un pic qui passe des 5% habituels à 40% de négatifs, ça veut dire quelque chose. Ça vaut au moins le coup de vérifier* ». Tous les prestataires rencontrés, en France et aux Etats-Unis, suggèrent ce type d'usage du *sentiment analysis*, qui ne constitue pas une « astuce » livrée par le conseiller à son client pour son rapport au logiciel. Au contraire, les développeurs eux-mêmes revendiquent ce qui relève d'une forme alternative d'objectivité, identifiée par Desrosières (1995 : 25), reposant sur la constance des mesures dans le temps, et non sur leur réalisme extrinsèque. Face à un objet difficilement mesurable, les prestataires enseignent donc au client à mesurer leur réputation non pas à l'aune des scores bruts, mais de leur variation. Certains éditeurs logiciels vont plus loin dans cette logique : Synthesio propose aussi de désactiver le module d'analyse de tonalité, dans le cas où celui-ci fournirait de manière récurrente des résultats inadaptés. Cette possibilité fut salutaire pour une grande ONG humanitaire, cliente du logiciel :

« On s'est aperçu que pour [nom de l'ONG], la machine donnait énormément de messages négatifs ; parce qu'en fait, les sujets qu'ils traitent sont des sujets anxigènes, la maladie, machin... Donc la machine dit que c'est négatif, parce qu'il y a le mot 'désastre', le mot 'catastrophe', etc. » (Thomas, dirigeant, Synthesio – France).

Ici, c'est la limite même de son approche ontologique du langage, associant systématiquement une polarité à un terme donné, qui est prise en compte par le logiciel à travers la possibilité de désactiver le *sentiment analysis*. Les termes de « désastre » ou de « catastrophe » ne qualifient pas la réputation de l'ONG, mais désignent bien le cœur de son activité quotidienne. La

⁹ Ces algorithmes linguistiques consistent à attribuer une tonalité à un *verbatim*, sur la base des termes qu'il comporte (approche morphosyntaxique), ou de sa ressemblance avec des *verbatim* pré-classés manuellement (*machine-learning*). Voir l'article fondateur de Pang et al. (2002).

¹⁰ Exemple le plus connu, le traitement de l'ironie est notoirement impossible, dans la mesure où celle-ci inverse la polarité préétablie pour chaque terme par le dictionnaire algorithmique.

faillibilité de l'algorithme est donc non seulement enseignée par les chargés de clientèle, mais aussi inscrite dans la faculté laissée au client de dessaisir cet outil de l'analyse, en fonction de son contexte réputationnel particulier. Mais les éditeurs ne se contentent pas de minorer le poids de l'algorithme de tonalité : nombre de tableaux de bord offrent à l'utilisateur de contredire son verdict. Synthesio et Radarly, notamment, permettent de modifier manuellement le codage de chaque *verbatim*, travail relativement fastidieux mais permettant d'obtenir, sur des corpus filtrés et ciblés notamment, des scores jugés plus fiables que ceux fournis par la machine. Enfin, c'est le logiciel de Sysomos qui laisse à son client le plus de latitude critique envers son module d'analyse de tonalité, en l'autorisant à modifier ponctuellement l'ontologie de référence du logiciel.

« On a aussi une option qui permet de personnaliser le sentiment, en changeant certains mots. [...] Si vous vous regardez une marque de soupe, en soi, 'sick' va presque toujours être un terme négatif. Alors que pour un fabricant de planches de surf, 'sick' sera presque toujours un terme positif. Il est impossible pour un outil de savoir ça, donc vous pouvez en fait aller directement à l'intérieur pour ajuster ça et dire : 'je vends des soupes, donc je veux que 'sick' soit négatif pour l'algorithme' » (Travis, *pre-sales manager* Sysomos).

L'originalité de Sysomos est de permettre à l'utilisateur de « rentrer » dans l'arrière-cuisine algorithmique du logiciel, et d'y effectuer des modifications substantielles. L'ontologie proposée constitue de ce fait une sorte de référence *par défaut*, ne revendiquant pas de statut de vérité particulier autre que celui de convenir à une majorité des clients du logiciel. Le fait que ses développeurs y laissent une porte dérobée vers l'ontologie, modifiable à souhait, témoigne en soi de l'agnosticisme militant de l'analyse de tonalité. Les concepteurs du module d'analyse du sentiment se refusent ainsi à polariser de manière irrévocable l'univers sémantique du client, et manifestent par une architecture ouverte la prise en compte de la polysémie de nombreux termes, et la variété des contextes. Contrairement à ce que l'on pourrait attendre, l'outil-boîte noire par excellence du tableau de bord ne fournit que peu, ou pas de résultats immédiats. Il ne réalise son potentiel explicatif que dans la mesure où il est mis en perspective, réinterprété, voire contredit par son usager qui lui impose ses propres normes et classements. Un tel constat invite à relativiser l'ampleur des controverses recensées par Boullier et Lohard (2012, ch.2) autour du *sentiment analysis* : comme le souligne Van Couvering (2007), en contexte commercial, la validation du produit de l'algorithme repose moins sur son objectivité clairement établie que sur sa capacité à satisfaire des critères d'ordre opérationnel et marchand.

Conclusion

Les logiciels d'e-réputation se caractérisent, on l'a vu, par leur grande plasticité : l'utilisateur est continuellement enjoint à définir son univers réputationnel par la circonscription d'entités sémantiques et cognitives, et par l'agencement ponctuel des scores et des modules analytiques. Cette perméabilité aux épistémologies professionnelles permet d'isoler ponctuellement des sous-corpus dans le flux continu des données conversationnelles ; vendus comme panoptiques et génériques, les tableaux de bord – comme bien d'autres outils automatisés (Benedetto-Meyer 2014) – nécessitent en fait spécification et interprétation locale. Ainsi, le tableau de bord contribue à performer l'utopie managériale de l'entreprise « orientée client » en réalisant un contact permanent entre celle-ci et ses consommateurs. En matérialisant la voix du consommateur, les outils de veille nourris par les catégories de leurs usagers tendent à agir en retour sur celles-ci : nous avons ainsi pu constater que la rémunération de certains managers « *social media* », est d'ores et déjà indexée sur l'évolution de scores spécifiques calculés par ces instruments.

Plus encore, le déploiement de ces tableaux de bord contribue également à dessiner un nouveau régime de l'opinion, mobilisant des algorithmes *a priori* agnostiques et des centres de calcul librement combinables afin de réduire la variété et la spontanéité d'un matériau préexistant à tout questionnement. C'est là l'innovation décisive de ce régime de connaissance : renonçant à susciter des opinions pré-calibrées – procédé qui faisait la force épistémique du sondage (Blondiaux 1998) –, les développeurs de tableaux de bord semblent réaliser le projet d'une approche praxéologique de l'opinion, fondée sur ses manifestations singulières dans l'espace public, et dont L. Quéré regrettait à l'époque qu'elle « [requière] *bien sûr de disposer de données conversationnelles représentant des processus réels* » (1990 : 56). Parce qu'elle rend disponibles ces données, tout en les rendant perméables aux épistémologies professionnelles des usagers, la forme « tableau de bord » commercialisée aujourd'hui en e-réputation parvient à transformer le bruit en signal, c'est-à-dire à faire coïncider les deux univers, potentiellement infinis, des réponses déjà écrites et des questions informulées.

Références

- Akrich, Madeleine. 1987. "Comment décrire les objets techniques ?", *Techniques et Cultures*, 9, 49–64.
- Beauvisage, Thomas, Jean-Samuel Beuscart, Vincent Cardon, Kevin Mellet, Marie Trespeuch. 2013. "Notes et avis des consommateurs sur le web: les marchés à l'épreuve de l'évaluation profane", *Réseaux*, 177 (1), 131–161.
- Benedetto-Meyer, Marie. 2014. "Du datamining aux outils de gestion. Enjeux et usages des 'scores' dans la relation client", in Alexandre Mallard et Emmanuel Kessous (dir.), *La fabrique de la vente. Le travail commercial dans les télécommunications*, Paris, Presse de l'Ecole des Mines, 147–169.
- Beuscart, Jean-Samuel. 2008. "Sociabilité en ligne, notoriété virtuelle et carrière artistique : Les usages de MySpace par les musiciens autoproduits", *Réseaux*, 152 (6), 139–168.
- Blondiaux, Loïc. 1998. *La fabrique de l'opinion. Une histoire sociale des sondages*. Paris, Seuil.
- Boullier, Dominique, Audrey Lohard. 2012. *Opinion mining et sentiment analysis*. Open Edition. <http://press.openedition.org/198>.
- Bourdieu. 1973. "L'opinion publique n'existe pas", *Les Temps Modernes*, 318, 1292–1309.
- Bowker, Geoffrey C., Susan Leigh Star. 1999. *Sorting Things out: Classification and Its Consequences*. Cambridge, Mass, MIT Press.
- Callon, Michel, Fabian Muniesa. 2003. "Les marchés économiques comme dispositifs collectifs de calcul." *Réseaux*, 122 (6), 189–233.
- Campbell-Kelly, Martin, Daniel D. Garcia-Swartz. 2007. "From Products to Services: The Software Industry in the Internet Era." *Business History Review*, 81 (4): 735–64.
- Cardon, Dominique. 2013. "Dans l'esprit du PageRank. Une enquête sur l'algorithme de Google." *Réseaux*, 177 (1): 63–95.
- Cardon, Dominique, Julien Levrel. 2009. "La vigilance participative. Une interprétation de la gouvernance de Wikipédia", *Réseaux*, 154 (2), 51-89.
- Chauvin, Pierre-Marie. 2013. "La sociologie des réputations: Une définition et cinq questions", *Communications* 93 (2), 131–45.
- Dagiral, Éric, and Sylvain Parasie. 2015. "La « Science Des Données » À La Conquête Des Mondes Sociaux. Ce Que Le « Big Data » Doit Aux Épistémologies Locales", in Pierre-Michel Menger et Simon Paye (dir.), *Big Data, Entreprises et Sciences Sociales*, Paris, Editions du Collège de France.

- Dagiral, Éric, and Ashveen Peerbaye. 2012. “Les mains dans les bases de données : Connaître et faire reconnaître le travail invisible”, *Revue d’anthropologie des connaissances*, 6 (1), 191–216.
- Desrosières, Alain. 1995. “Classer et mesurer : les deux faces de l’argument statistique”, *Réseaux*, 71 (13), 11–29.
- . 2008. *L’argument Statistique (I) - Pour Une Sociologie Historique de La Quantification*. Paris, Presses de l’Ecole des Mines.
- Fouetillou, Guilhem. 2009. *Blog de Linkfluence*. Adresse URL : <http://fr.linkfluence.net/2009/09/03/la-tarte-a-la-creme-du-sentiment-analysis/> (consulté le 17/11/2011).
- Gillespie, Tarleton. 2014. “The Relevance of Algorithms”, in Pablo Boczkowski, Tarleton Gillespie, et K.A. Foot (dir.), *Media Technologies. Essays on Communication, Materiality and Society*, Boston, MIT Press, 167–93.
- Knorr-Cetina, Karin, Barbara Grimpe. 2008. “Global Financial Technologies: Scoping Systems That Raise the World”, in Trevor Pinch et Richard Swedberg (dir.), *Living in a Material World: Economic Sociology Meets Science and Technology Studies*, Cambridge, Mass, MIT Press, 161–90.
- Kotras, Baptiste. 2015. “Des opinions qui (se) comptent. Influence, visibilité et hiérarchisation des opinions sur le web”, in Christophe Alcantara (dir.), *E-réputation. Regards croisés sur une notion émergente*, Toulouse, Lextenso.
- Mellet, Kevin. 2012. “Contagion, influence, communauté. Petite socio-économie des agences de social media marketing”, in Franck Cochoy (dir.), *Du lien marchand : comment le marché fait société. Essai(s) de sociologie économique relationniste*, Toulouse, Presses Universitaires du Mirail, 151-173.
- Orléan, André. 1999. *Le pouvoir de la finance*. Paris, Odile Jacob.
- Pang, Bo, Lillian Lee, Shivakumar Vaithyanathan. 2002. “Thumbs up? Sentiment Classification Using Machine Learning Techniques”, in *Proceedings of EMNLP’02*, 79–86.
- Power, M., T. Scheytt, K. Soin, K. Sahlin. 2009. “Reputational Risk as a Logic of Organizing in Late Modernity”, *Organization Studies*, 30 (2-3), 301–324.
- Quéré, Louis. 1990. “Opinion : l’économie du vraisemblable. Introduction à une approche praxéologique de l’opinion publique”, *Réseaux*, 43 (8), 33–58.
- Rieder, Bernhard. 2012. “ORDER BY column_name. The Relational Database as Pervasive Cultural Form”, Présentation au workshop *The Lived Logics of Database Machinery*, Londres.
- Van Couvering, Elizabeth. 2007. “Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality”, *Journal of Computer-Mediated Communication*, 12 (3), 866–87.