

# Ce que le web 2.0 a changé dans notre manière de faire de la veille

Par Christophe Deschamps

Article d'introduction à l'étude SerdaLAB 2009, « Veille, moteurs de recherche et collaboratif : marché, besoins et tendances en 2009-2010 » - Accessible ici : <http://www.serdalab.com/etude/veille-moteurs-de-recherche-et-collaboratif-marche/>

Le terme "web 2.0" a été initialement utilisé en septembre 2005 par Timothy O'Reilly, éditeur bien connu d'ouvrages sur l'informatique, dans un article introduisant la conférence éponyme qui devait se tenir au mois d'octobre suivant. Il y décrivait une nouvelle vague d'outils et de services qui allait, selon lui, modifier notre rapport au web et à l'information. Dans un article rédigé quelques mois plus tard, j'écrivais que les technologies du web 2.0, en offrant de nouveaux moyens de découvrir et de partager l'information, tout en donnant à chacun les moyens de mieux se l'approprier, allaient plonger les veilleurs et documentalistes dans une ère nouvelle. Qu'en est-il quatre ans plus tard?

Nous aurons en partie répondu à cette question en détaillant deux innovations "2.0" qui justifient à elles seules l'idée d'une nouvelle génération d'outils et donc d'usages qui en découlent.

La première est le RSS, un acronyme pour Really Simple Syndication désignant une version simplifiée du langage XML qui permet de diffuser facilement le contenu d'un site ou d'un blog en

« l'encapsulant » entre des balises (titre, auteur, dates, corps du texte,...). Celles-ci sont alors susceptibles d'être reconnues par des logiciels ou services en ligne dits « agrégateurs ». Il suffit au lecteur de repérer sur un site web l'icône d'un flux RSS pour s'y. Dès lors il n'y a plus à le visiter puisque chaque nouvel article publié est automatiquement diffusé dans l'agrégateur.

Blogs, wikis, services de « *social networking* » (LinkedIn, Viadeo), de « *social bookmarking* » (Delicious, Diigo), de « *micro-blogging* » (Twitter), tous les outils 2.0 utilisent cette technologie pour diffuser leur contenu. Le RSS constitue donc une véritable "grammaire" qu'il s'avère essentiel de maîtriser pour qui veut mener une veille efficace sur le web.

La surveillance de flux RSS est toutefois suffisamment simple pour que toute personne désireuse de s'en servir puisse le faire sans compétences techniques particulières et cette simplicité d'usage doit être soulignée car il s'agit d'un trait spécifique aux outils 2.0. Plusieurs raisons à cela, d'une part ils ont très souvent été créés par des développeurs pour adresser leurs propres problématiques de gestion de l'information - ainsi en va-t-il du wiki créé par Ward Cunningham en 1994 ou du service de gestion des favoris (*social bookmarking*) Delicious, créé par Joshua Schachter en 2003 -, d'autre part, ils correspondent à une stratégie marketing élémentaire : les personnes connectées au web étant de plus en plus nombreuses, il faut adapter les services web au plus grand nombre pour les rendre utilisables par tous.

Aussi louable soit-il, cet objectif, en sert toutefois un second autrement plus important : faire utiliser ces services par le plus grand nombre possible d'internautes leur donne une véritable valeur ajoutée.

Nous abordons ainsi la seconde innovation forte du web 2.0, le *crowdsourcing*. Il s'agit d'un néologisme conçu en 2006 par Jeff Howe et Mark Robinson, du magazine Wired. Sa traduction littérale est « approvisionnement par la foule » auquel on préférera celle de « foules-ressources ». Il ne désigne pas une fonctionnalité particulière mais un concept qu'appliquent les services « 2.0 » en misant sur une grande quantité d'utilisateurs pour que la mécanique du nombre, canalisée par un outil technique et statistique, fasse émerger une valeur ajoutée dont tous bénéficieront en retour. Celle-ci pourra être explicite, comme lorsque nous aidons Google à améliorer une traduction automatique (Google outils linguistiques) ou, plus souvent, implicite, comme lorsqu'Amazon nous envoie des propositions d'ouvrages susceptibles de nous intéresser. Dans ce cas les algorithmes de filtrage collaboratif de ce service utilisent nos choix antérieurs pour nous présenter des recommandations automatiques selon un principe simple : « si comme xx % de nos clients, vous avez aimé tel produit, alors vous aimerez aussi tel autre ». Il est alors aisé de comprendre que plus Amazon aura mis en place des moyens de capter nos traces implicites et explicites pour les

transformer en données statistiques et plus la société sera en mesure d'affiner les propositions qu'elle fait à ses clients, c'est à dire de mieux les servir afin de générer plus de profit...

Ces deux innovations ont été le moteur de l'évolution de pratiques informationnelles dont les professionnels de l'information doivent au mieux être les acteurs et au pire les témoins. Quelles sont-elles?

### **Fragmentation des données, fragmentation des moteurs de recherche**

Nous sommes plus nombreux que jamais à tenir un blog<sup>1</sup>, commenter des articles de journaux, donner notre avis ou converser via Twitter<sup>2</sup>, mais aussi à publier des vidéos sur Youtube et partager nos documents numériques sur Scribd ou Slideshare. Le simple fait que tout le monde soit désormais en mesure de s'exprimer sur tout à d'ailleurs généré un nouveau type de veille. Que vous soyez une entreprise ou un particulier il est désormais indispensable de surveiller votre réputation numérique, c'est à dire ce qui se dit de vous sur les différents territoires du web, afin de détecter d'éventuels risques (attaque de vos produits, diffamation) et d'y répondre.

Ces nouvelles pratiques entraînent donc un accroissement de l'information numérique sur lequel nous ne possédons malheureusement pas de chiffres mais qui est, à n'en pas douter, exponentiel. C'est précisément ce franchissement d'un seuil quantitatif qui est la source d'une rupture dont nous commençons tout juste à prendre conscience. Les services cités ci-dessus deviennent en effet des silos dans lesquels l'information vient se déposer et si jusqu'à maintenant Google permettait d'adresser la plupart des questions susceptibles de trouver une réponse sur le web, ce n'est désormais plus le cas. Si par exemple vous recherchez un document PDF il vous faudra certes interroger Google, Yahoo! ou Bing avec l'opérateur adéquat mais il vous faudra aussi, dans un objectif d'exhaustivité, utiliser (et donc connaître!) un second moteur pour Scribd, Docstoc, Issuu, et un troisième indexant les services de partage de fichiers de type Rapidshare ou Megaupload (car vous pourriez bien y faire des découvertes étonnantes ...), etc.

Si mener une veille exhaustive sur le web a toujours été illusoire, nous pouvions avoir l'impression qu'en maîtrisant bien les moteurs de recherche généralistes nous n'étions pas loin d'y parvenir. Avec le web 2.0 cet idéal s'éloigne à grands pas. Non seulement nous en sommes très loin, mais il n'existe tout simplement pas d'outil de veille, aussi avancé soit-il, susceptible de nous aider à l'atteindre. Cela nécessite donc de la part des professionnels de l'information une position paradoxale qui nous semble toutefois la seule possible : nous devons, d'une part, mener une veille permanente sur tous les services et moteurs susceptibles de nous aider à rechercher l'information dans ces nouveaux silos et, d'autre part, faire notre deuil d'une exhaustivité inatteignable, ce qui nécessite un changement de posture qui tient de la mini-révolution copernicienne. Accepter cela c'est en effet tourner le dos à une conception sur laquelle reposait jusqu'à maintenant tout travail documentaire digne de ce nom et que symbolise bien les classifications que nous utilisons (thésaurus, taxonomies), à la fois reflets de nos connaissances et promesses que tout ce que nous découvrirons trouvera sa place dans l'ordre du monde. Ce territoire n'est définitivement pas celui des veilleurs dont le devoir est de lutter en permanence contre la fausse sécurité du plan de classement pour laisser la place à l'information apparemment incohérente ou aux signaux faibles.

### **Démocratisation des pratiques de veille**

Grâce aux outils du web 2.0, la veille est devenue l'affaire d'un grand nombre de personnes dont une bonne part n'a jamais suivi de formation initiale. Une carence généralement vite compensée par une

---

<sup>1</sup> Technorati en répertoriait 133 millions en 2008, sans même prendre en compte la Chine et l'Inde.

<sup>2</sup> Ce service serait actuellement utilisé par 5 millions de personnes.

curiosité que motive un besoin personnel ou professionnel de se tenir informé, mais aussi par une dose de bon sens et une bonne connaissance des outils disponibles. Ainsi trouve-t-on maintenant régulièrement sur le web francophone des articles de blogueurs expliquant comment ils ont mis en place leur propre dispositif de veille à coup de RSS, de moteurs de recherche et d'agrégateurs. Une question "poil à gratter" surgit alors inévitablement : est-il normal que, quatre ans après le "lancement" du web 2.0, des professionnels de l'information ne sachent toujours pas utiliser ces technologies? N'est-ce pas prendre le risque de se voir déborder par des utilisateurs avancés ainsi que par une "Google génération" qui ne se pose même pas la question de "faire de la veille" mais se contente de surveiller, via son agrégateur favori, l'arrivée de telle séquelle d'un jeu vidéo populaire?

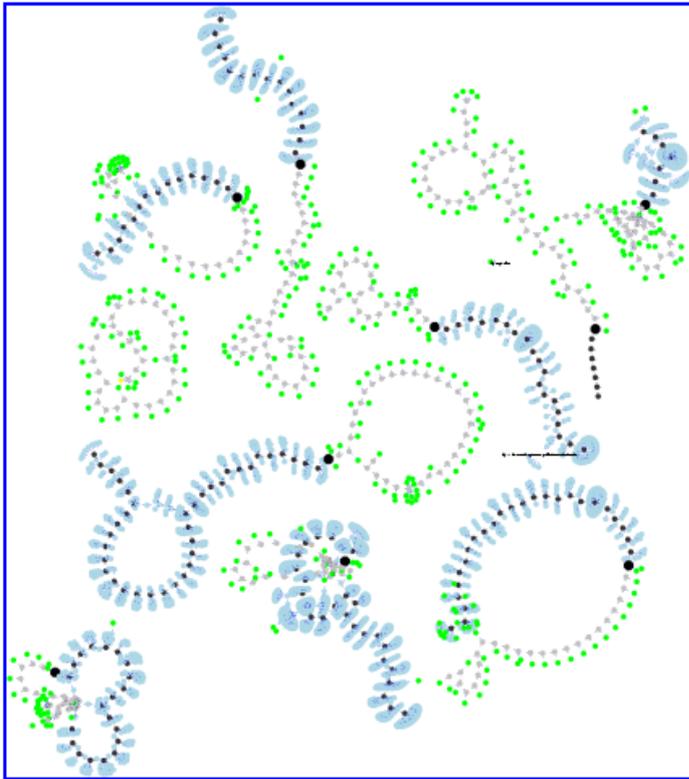
### **Veille sociale et en temps réel**

Le nouvel écosystème engendré par le web 2.0 a toutefois généré ses propres réponses aux problèmes d'accroissement et de fragmentation de l'information. En effet, ayant été pensé pour servir l'individu ancré au cœur de son réseau, il permet de mener une veille en utilisant les membres constituant ce dernier comme autant de filtres à l'information. Ainsi lorsque vous *followez* (suivez) quelqu'un sur Twitter ou Friendfeed vous bénéficiez des découvertes qu'il fait à titre personnel et souhaite ainsi partager. Si vous avez bien choisi les personnes à suivre (*sourcing*) vous finirez par voire émerger de leur "conversation" les informations les plus intéressantes du moment sur votre domaine de prédilection et ce n'est donc plus un site web ou un blog qui sera estampillée "source pertinente" mais bien une personne.

Cette possibilité d'une information en temps-réel est un sous-produit du web 2.0 qui a montré sa pertinence lors d'événements récents (atterrissage d'un A320 sur l'Hudson, attentats de Bombay). Signe de ce succès, on voit depuis quelques mois apparaître de nombreux moteurs permettant d'exploiter cette manne informationnelle : Yauba, Pulse, Scoopler, Spkyed, All140,... A tel point que certains se demandent s'il est encore utile d'avoir un agrégateur de flux RSS alors qu'ils font maintenant l'essentiel de leur veille "en temps réel". De fait la réponse à cette question est en soi un critère qui doit distinguer le professionnel de l'amateur : oui, utiliser un agrégateur de flux RSS est essentiel pour, d'une part, surveiller ce qui n'est pas "populaire" (et risque donc de contenir de l'information de type "signal faible") et pour, d'autre part, disposer d'un historique qui formera la base d'un travail ultérieure d'analyse.

### **Analyse statistique et modélisation prédictive**

Le flot ininterrompu de données qui nous arrive désormais via le web a entraîné la mise en ligne de solutions d'aide à l'analyse de données jusqu'alors réservées à des professionnels des statistiques (Manyeyes d'IBM, Swivel, Chartle, Verifiable,...). Si leur usage est adapté au plus grand nombre, il s'agit toutefois bien d'une seconde réponse aux problèmes évoqués plus haut. Un service comme Bscopes par exemple donne une bonne idée du type d'outils d'aide à l'analyse qui pourrait se développer dans les années à venir. Il permet de dresser une cartographie des flux RSS d'un utilisateur et en fait émerger jour après jour les billets les plus populaires. Chaque blog ou ensemble de blogs thématiques dispose ainsi d'une identité visuelle évolutive



Représentation graphique des flux RSS de l'auteur par le service Bscopes ([www.bscopes.com](http://www.bscopes.com))

L'utilisateur est alors en mesure de traiter beaucoup plus d'informations entrantes grâce à une vision d'ensemble qui lui permet de détecter des schémas (*patterns*) ou des configurations révélant par exemple qu'il se passe quelque chose dans telle partie de la carte. Dans un deuxième temps, et à l'instar de ce que permettent les systèmes de prédictions météorologiques, on pourra envisager de modéliser cet univers mouvant afin de prévoir de quelle « région » pourrait venir la prochaine information de rupture, celle permettant par exemple de détecter avant les autres l'innovation qui fera la différence avec les produits concurrents. Il s'agira en quelque sorte de mener une veille par anticipation, dans le but de mieux anticiper.

### **Veille et analyse collaborative pour les organisations**

Il n'aura pas fallu attendre longtemps pour voir l'apparition de solutions inspirées du web 2.0 destinées aux organisations. Déployés sur les serveurs d'une entreprise ou distribués "on the cloud", des services comme BlueKiwi, Knowledge Plaza, Jamespot Pro ou encore Xwiki permettent à chacun de partager très simplement sa veille quotidienne avec son équipe. Mais là où ces solutions apportent une valeur ajoutée encore plus grande au cycle de l'information c'est dans leur capacité à permettre une analyse collaborative de celle-ci. En effet, la logique de l'entreprise qui amène à toujours plus de rationalisation se trouve, avec l'information, confrontée à une limite qui est celle de la subjectivité de ceux qui la reçoivent et l'interprètent. Les mécanismes intégrés aux technologies 2.0 vont permettre d'ajouter ces différentes subjectivités les unes aux autres dans le but de co-construire une réalité toujours relative mais enrichie des connaissances de chacun. C'est sur ce principe que la communauté du renseignement américain a lancé en septembre 2007 Intellipedia, une plateforme qui intègre plusieurs technologies « 2.0 », avec pour objectif d'aider à consolider des informations grâce aux validations fournies par les analystes des différentes agences de renseignement américaines.

Toutes les évolutions que nous venons d'évoquer sont plus que des pistes à suivre, elles sont la

réalité que nombre d'entre nous vivent déjà quotidiennement et elles modifient très concrètement le rapport que nous entretenons à l'information. A la question de savoir si le web 2.0 à modifié nos pratiques de veille la réponse est donc oui, évidemment, en profondeur et pour longtemps.

Schéma récapitulatif de l'article :

