

N° 219 - Septembre 2005

BASES

Connaître et bien utiliser
les banques de données

Agents

- Weblogs : demain ... la veille, pp.1-4

Web invisible

- Nouveautés chez Dafsaliens, p.6
- Urbadoc.com : un "mini-serveur" spécialisé sur l'urbanisme, pp.8-9

Actualités

- Cedrom-SNI propose en exclusivité la nouvelle offre de l'Agefi, p.7

A Lire

- "The Skeptical Business Searcher", pp.10-11

Agenda

- Veille sur le Net : des outils simples au service de l'internaute, p.12
- Optimisez vos recherches sur le Web visible et invisible, p.12

Weblogs : demain ... la veille

Olivier Ertzscheid

Un nouveau « Web »

Après l'avènement du "Web", après celui de l'e-mail et des listes et groupes de discussion, après l'essor et l'engouement relatif aux technologies de la messagerie instantanée et conjointement au développement des technologies WiFi et de l'Internet sans fil, les "weblogs" ou "blogs" ou "carnets Web" s'affirment comme la dernière évolution marquante liée aux technologies de l'information et de la communication.

Même si plusieurs déclinaisons sont possibles, un weblog désigne habituellement une page personnelle, généralement alimentée par un seul individu et constituée des courts messages affichés et archivés automatiquement selon un ordre ante-chronologique.

Ils font progressivement leur entrée dans le monde de l'entreprise en complément ou parfois en remplacement d'applications de CRM (outil de gestion de la relation client), d'outils collaboratifs, voire d'intranets complets, grâce à des fonctionnalités de gestion de projet.

Mais c'est pour les veilleurs et particulièrement dans des contextes d'intelligence économique que leur impact est le plus mesurable : ils s'avèrent être des sources d'information aussi inédites que pertinentes.

Une croissance exponentielle

Les weblogs sont en France – et depuis déjà quelques années outre-Atlantique –, un vrai phénomène de masse : 4,1 millions en septembre 2003, 6,5 millions en décembre 2004, 7,8 en mars 2005 et une courbe de développement exponentielle. Des chiffres certes impressionnants mais à modérer, car beaucoup ne sont que très peu ou pas du tout alimentés. Néanmoins, l'engouement pour leur utilisation est comparable à celui rencontré pour l'usage du courrier électronique il y a de cela quelques années, et leur taux de croissance à celui du Web à ses débuts. "Potential for growth within the enterprise similar to the growth of instant messaging"⁽¹⁾.

(1) Mayfield R., "Blogs bubble into business", <http://www.socialtext.com/weblog/040128enterprise.html>, consulté le 31/08/2005



Weblogs *suite*

Olivier Ertzscheid

Les raisons du succès : des outils gratuits, faciles, efficaces

Les raisons de ce succès sont liées à plusieurs facteurs convergents. D'abord **leur coût minimal ou nul** grâce à l'arrivée sur le marché, dès juillet 1999, d'outils grands publics gratuits et faciles d'emploi, pour lesquels il n'est plus nécessaire d'avoir une connaissance même minimale d'HTML pour publier des pages sur le Net.

Le plus utilisé de ces outils demeure Blogger (www.blogger.com), racheté par Google.

Ensuite, **leur mode de publication simple et intuitif** : quel que soit l'outil, tout se déroule à partir d'une interface semblable à celle de l'envoi d'un courrier électronique.

Enfin, leurs **fonctionnalités redoutablement efficaces** en terme de gestion automatique de l'information :

- les informations "postées" sont automatiquement archivées et l'utilisateur peut choisir la fréquence de cet archivage (mensuel, hebdomadaire, quotidien) ;
- chaque information se voit automatiquement attribuée un URL permanent ;
- chaque information est "indexée" à l'aide de balises "méta" minimales : auteur, date de publication, titre (d'autres champs d'indexation peuvent être ajoutés).

Contribuent également à cette efficacité l'ensemble des services collaboratifs qu'offrent ces weblogs :

- chaque information postée peut faire l'objet de commentaires, liés au message et également archivés ;
- gestion des droits : il est possible d'ouvrir des droits différents en lecture et en écriture et donc de "faire tourner" ces outils sur de larges communautés d'utilisateurs.

Comme le résume rapidement cet inventaire non exhaustif, les weblogs offrent ainsi au plus grand nombre les attendus techniques minimaux de tout système d'information "élaboré" : indexation, gestion des droits, archivage, recherche.

C'est d'ailleurs à ce titre qu'ils remplacent d'ores et déjà dans certains grands groupes les architectures lourdes et trop souvent sous-utilisées de type Intranet.

"Les blogs et leur écosystème, immiscés entre le courrier électronique et le Web, sont le maillon manquant dans la chaîne de communication..."

Ils se rapprochent de la promesse originelle du Web, dans la mesure où ils sont le premier outil qui a rendu facile – ou au moins simplifié – l'action de publier sur le Web⁽²⁾.

Des weblogs pour les veilleurs

Parmi l'ensemble des weblogs se développant sur le Net, il est possible d'isoler trois macro-catégories⁽³⁾ :

1. des journaux personnels, simples "consignes" d'états d'âme quotidiens d'anonymes ou de célébrités (Pierre Bilger ancien PDG d'Alstom, Alain Juppé, Dominique Strauss-Khan, etc.) ;
2. des carnets de note ("notebooks") qui proposent un contenu sous forme d'approches et de synthèses ("focused essays") ;
3. et enfin des weblogs-filtres ("filters") annonçant colloques, séminaires, parutions, manifestations scientifiques, etc.

Même si de nombreux weblogs sont une hybridation des catégories listées ci-dessus, elles n'en demeurent pas moins un filtre pertinent pour l'analyse.

A ce titre, les deux dernières catégories constituent des sources intéressantes pour les actions de veille, sources qui pourraient, à terme et sous réserve que l'explosion quantitative du phénomène se confirme, devenir incontournables – voire remplacer la totalité des sources Web habituellement consultées (sites Web, forums...) : nombre de diffuseurs d'information ont

en effet d'ores et déjà fait le choix du blog en remplacement des formules de type newsgroup ou liste de diffusion.

A titre d'exemple, voici quelques weblogs francophones et anglo-saxons labellisés "Intelligence économique" ou "Business Intelligence" et appartenant aux deux dernières catégories :

- **Carnets** :
 - <http://totem-consult.blogspot.com>
 - <http://infoguerre.blogspot.com/>
 - <http://www.aurorawdc.com/ci/>
- **Filtres**
 - <http://www.itfacts.biz/> (ce site ne délivre qu'une information strictement factuelle) ;
- **Hybrides** (à la fois filtre et carnet)
 - <http://deanie.ultimteam.com/>
 - www.outilsfroids.net
 - www.urfist.info

RSS : Une technologie au service de la veille

Les weblogs se distinguent des forums et des communautés en ligne de type Usenet en cela qu'il sont initialement des espaces de publication et non de discussion.

(2) Gillmor D., *We the Media (hard-back) : Grassroots Journalism by the People, for the People*, July 2004, O'Reilly Media, Inc., CA.

(3) Blood R., *The Weblog Handbook: Practical Advice on Creating and Maintaining your Blog*, Cambridge, MA: Perseus Publishing, 2002.



L'information publiée, même si son statut varie considérablement d'un site à un autre, est ainsi en général immédiatement capitalisable.

Cela est notamment rendu possible par l'utilisation du format RSS qui, s'il n'est pas l'apanage exclusif des weblogs, y a trouvé l'un de ses usages les plus pertinents en terme de diffusion et de traitement d'information, via le processus de syndication.

"RSS is an extensible metadata description and syndication format, currently used for a number of applications, including news and other headline syndication, weblog syndication, etc."⁽⁴⁾

Très concrètement, la syndication de contenu rend possible sur des logiciels clients de type "agrégateurs", l'affichage structuré en XML des informations publiées sur un site. Il est ainsi possible de recevoir uniquement les derniers contenus d'un site et d'en garder une trace, un archivage chronologique.

La différence avec les outils classiques de monitoring est de taille, puisque ce type de diffusion permet d'éviter nombre des biais traditionnellement associés au processus de veille (fausses mises à jour, localisation de la partie effectivement mise à jour, ...) tout en conservant la plupart des fonctionnalités classiques (cycles programmables, gestion d'arborescence, etc.)

Si l'on reprend le découpage canonique d'un cycle de veille (ciblage des informations, traque, collecte, traitement, diffusion), indépendamment de sa nature (économique, technologique, documentaire) et de sa finalité (stratégique, anticipative, commerciale, etc.), l'exploitation des weblogs permet un double gain, à la fois qualitatif et quantitatif.

Prenons à titre d'exemple la phase de repérage et de ciblage des sources.

Il existe dans la "blogosphère", sur une thématique, un secteur, une industrie ou un produit donné, ce que l'on appelle des "pundit", experts auto-proclamés et dont les sites bénéficient d'un nombre de liens entrants (backlinks) qui tendent à asseoir cette expertise lorsqu'elle semble justifiée.

Le nombre de liens, très dense, autour desquels se déploient les weblogs publics les rend aisément localisables.

Une fois repéré ces incontournables, une méthodologie classique de type "fil d'Ariane" permet, souvent simplement en consultant la "blogroll" (liste des sites consultés par l'auteur du blog d'origine), de rendre ce ciblage quasi-exhaustif, en repérant des blogs tout aussi pertinents mais de moindre notoriété. Ce processus reste pour l'heure incontournable, étant donné qu'il n'existe pas encore de "Google des blogs"⁽⁵⁾ permettant d'en recenser la totalité, même si des indices permettent de supposer que Yahoo! déploiera bientôt sur

son moteur un onglet "blog", au même titre que ceux existant déjà (annuaire, web, images ...)

Une technologie incitative qui rend bavard

Du fait de leur facilité de mise en place, d'usage et d'alimentation, les weblogs sont un facteur d'incitation qui met à disposition des veilleurs de nouvelles sources, pour la seule raison que de nouveaux individus se mettent à publier des informations qui réclamaient auparavant de passer des filtres techniques démotivants (connaissance du code HTML, gestion d'un espace d'hébergement, gabarit du site et de la page, etc).

Quelle que soit leur provenance, et à tous les échelons hiérarchiques de l'entreprise, ces "nouveaux" acteurs de la diffusion d'information publient et communiquent avec, la plupart du temps, une liberté de ton que rend possible l'absence de processus de validation ou de chaîne éditoriale.

Ainsi PDGs, CEOs et autres experts sectoriels deviennent de nouvelles sources disponibles pour les veilleurs. Des gens jusqu'ici détenteurs d'un savoir opérationnel ont la possibilité de le rendre disponible à leur communauté ou au plus grand nombre.

Apports des weblogs dans un processus de veille

Grâce à la conjugaison d'un outil d'édition (blog) et d'une technologie de diffusion (RSS), il est possible de mettre en place de la syndication de contenu, c'est-à-dire de lire dans un logiciel client dédié toutes les dernières informations postées sur tel ou tel weblog.

Un bond quantitatif

Un individu peut "manuellement", chaque jour depuis son poste de travail, vérifier et éditer le contenu mis à jour d'un nombre de sites considérable. Sachant que seuls 20 % des weblogs existants publient une nouvelle information dans un délai inférieur à 24 heures, le suivi quotidien ou bi-quotidien d'un immense corpus devient possible et il est protégé des biais classiques (spam, fausses mises à jour ...).

Un bond qualitatif

Une fois cette information "monitorée", la souplesse et la modularité que l'on peut appliquer au traitement de l'information collectée est là encore remarquable et déroutante de facilité.

En termes de contenus, en sus des fils d'information (ou "flux RSS" ou encore "fils de syndication") standards fournis par les

(4) "RSS est un format extensible de description et de syndication de métadonnées, couramment utilisé par nombre d'applications parmi lesquelles la syndication d'actualités, de titres de presse, de weblogs, etc." Nanno T. (et al.), "Automatically collecting, monitoring, and mining japanese weblogs" <http://lr-www.pi.titech.ac.jp/blogwatcher/paper/www2004-838.pdf>, consulté le 31/08/2005.

(5) Précision ultime, au moment de mettre sous presse : Google a finalement pris Yahoo! de vitesse et a lancé son "Google des blogs" le 14 septembre 2005 (<http://blogsearch.google.fr>)



sites Web, il est possible de mettre en place des filtres internes au contenu même du site. Pour un site donné (prenons à titre d'exemple le site du magazine américain *Wired*), il est possible de récupérer via un fil d'information spécifique :

- uniquement les articles de l'éditorialiste Meg Ryan (spécialiste de la question des weblogs) ou d'un(e) de ses confrères(sœurs) ;

- l'intégralité des billets postés (syndication longue) ou le titre et les premières lignes seulement (syndication courte), et ce avec ou sans les commentaires postés ;

- de définir ses propres mots-clés et de ne récupérer que les billets publiés n'importe où sur le site de *Wired* et contenant le mot-clé en question.

L'information ainsi récupérée n'arrivant pas "brute" mais déjà élaborée selon une grammaire XML, il est alors possible de la coupler avec des applications déjà existantes dans l'entreprise (revue de presse, notes d'étonnement, rapports, etc)

Veille mais aussi guerre économique

Du fait de la densité des liens qui constituent la blogosphère, de l'accession à ce mode de diffusion de gens jusqu'ici silencieux et de la visibilité presque immédiate de l'information ainsi éditée dans les moteurs de recherche, les weblogs constituent également de redoutables outils au service de pratiques relevant de la désinformation ou de la guerre économique.

Ainsi, la veille du lancement international de l'Airbus A 380, Randy Baseler, Vice-Président du Marketing de Boeing, a choisi d'ouvrir son blog personnel⁽⁵⁾ stigmatisant (selon lui) les "erreurs de jugement" associées à ce gros porteur et espérant probablement contrebalancer une couverture médiatique essentiellement élogieuse.

Les blogs : demain le grand soir ?

L'adhésion du grand public au phénomène est déjà actée. De plus en plus d'entreprises et/ou de sites d'information les utilisent, en remplacement de leurs "anciens" modes de diffusion (forums, mailing-lists...). Des pans entiers de l'Internet basculent dans la blogosphère, comme par exemple la totalité des messages postés sur des forums de discussion du service Yahoogroups⁽⁶⁾, qui sont désormais accessibles au format RSS. Les **Intranets changent** de physionomie et emportent (enfin !) l'adhésion des utilisateurs, en encourageant au lieu de décourager la mutualisation de contenus.

Les **moteurs de recherche** proposent des services d'alerte sur mot-clé en RSS, voire même de syndiquer le résultat d'une recherche.

Les **outils clients** de type "agrégateurs" sont de plus en plus inclus directement dans les logiciels de messagerie électronique ou les navigateurs, mettant sur un même plan le monitoring de pages Web, de fils de syndication, de newsgroups ou bien encore de messageries électroniques.

Les weblogs remplacent progressivement les anciennes **pages personnelles**, bien souvent bricolées dans un HTML balbutiant.

Au vu de ces quelques éléments, là encore non exhaustifs, on voit mal ce qui pourrait remettre en cause un tel changement.

Pour autant, il ne s'agit pas de sombrer dans une trop enthousiaste technophilie. Certaines résistances et difficultés doivent être soulignées :

- le déploiement de weblogs en entreprise n'est pas si naturel que cela et se heurte à nombre de **résistances** essentiellement "culturelles", souvent liées aux **modes de management** en vigueur ;

- sur le web "public", le référencement et l'indexation des blogs par les moteurs de recherche est encore loin d'être optimal et exhaustif.

Enfin, au-delà de certaines dérives (employé remercié pour avoir dans son blog diffusé des informations non autorisées ou non conformes), le potentiel des blogs semble encore

largement sous-évalué, aussi bien dans le cadre des pratiques de veille que dans celui d'autres applications, dans lesquelles ils font chaque jour la preuve de leur efficacité et au rang desquelles on peut notamment citer la communication de crise⁽⁷⁾, la gestion de la relation client, le marketing viral, mais également la gestion des connaissances⁽⁸⁾ (KM).

Dans tous ces secteurs, les weblogs donnent lieu chaque jour à de nouvelles expérimentations⁽⁹⁾ et innovations, marquant là encore un changement d'importance et une valeur ajoutée souvent immédiatement mesurable.

(5) www.boeing.com/randy/

(6) <http://groups.yahoo.com>

(7) "Crisis Communications Blogs Use Them Cautiously"

<http://tinyurl.com/43ozk>

(8) www.lgi.ecp.fr/~dudezert/images/CER0501.PDF

(9) nombre d'entreprises utilisent ainsi des K-Blogs (Knowledge blogs) pour animer des communautés de pratique ou faire de la localisation d'expertise. D'autres (les mêmes le plus souvent) se servent également de P-Blogs (Project Blogs) qui prennent en charge toutes les étapes d'une gestion de projet.

L'auteur

Après avoir été Directeur des Etudes du Master d'intelligence économique de l'Université Toulouse 1, Olivier Ertzscheid est aujourd'hui Maître de Conférences en Sciences de l'information à l'Université de Nantes (IUT de la Roche sur Yon). Il a mis en place et collaboré au weblog du réseau Urfist (www.urfist.info) avant d'ouvrir il y a peu son propre weblog : <http://affordance.typepad.com>

Cet article reprend les points abordés dans l'article « *Weblogs : un nouveau paradigme pour les systèmes d'information et la diffusion de connaissances ? Applications et cas d'usage en contexte de veille et d'intelligence économique.* » présenté lors du Colloque ISKO 2005. La version présentée au colloque est disponible en ligne à l'adresse http://archivesic.ccsd.cnrs.fr/sic_00001433.html