

Quelle veille pratiquer sur un site web ?

« On veille, on pense à tout, à rien
On écrit des vers, de la prose
On doit trafiquer quelque chose
En attendant le jour qui vient. »
Aragon (*Blues*)

Le terme de « veille » est omniprésent aujourd'hui dans les revues spécialisées. Il apparaît partout, depuis son introduction en France en 1988, au moment où le ministère de l'Enseignement supérieur et de la Recherche mettait en place un Comité d'orientation stratégique de l'information scientifique et technique et de la veille technologique.

Claire Dartois

Bibliothèque publique
d'information
Service des documents
électroniques

dartois@bpi.fr

Il est souvent accompagné d'équivalents, comme « *intelligence économique* », « *intelligence des affaires* », « *veille stratégique* » dans la littérature destinée au milieu des affaires. Citons, entre autres, l'article récent intitulé « Distinguez l'info stratégique du bruit de fond¹ », ou le dossier réalisé sur le web par la société Multimedium² qui en explicite les stratégies pour l'entreprise. Avant d'en voir les applications au monde des bibliothèques, dans le cas de la veille appliquée à un corpus de liens déjà constitué, attardons-nous un instant sur l'étymologie même du mot.

L'*Encyclopédie Hachette* indique, à la suite du sens premier, évident (s'abstenir de dormir la nuit), celui de « *monter la garde, faire preuve de vigilance* ». « Veiller à » consiste à « *s'occuper attentivement de* » et « Veiller sur » à « *surveiller avec vigilance* ». Dans le monde de la mer, on dira « *faire attention à, surveiller : veiller la mer, les voiles* ». Ces diffé-

rents sens, assez peu éloignés les uns des autres d'ailleurs, convergent vers l'idée d'attention soutenue, d'alerte face à un danger, de surveillance d'un bien ou d'un territoire. Le sens migre alors vers l'aspect dynamique de l'activité. François Jakobiak, dans un ouvrage qui fit référence en 1991³, analyse ce terme en disant qu'il n'est pas en adéquation avec la pratique « *qui comprend les opérations de recherche, de collecte, de diffusion, de traitement, de validation et d'utilisation de l'information à caractère stratégique. Celles qui sont nécessaires à la prise de décisions importantes* ». Un article de fond, signé de Paul-Dominique Pomart, dans le *Dictionnaire encyclopédique de l'information et de la documentation*⁴ fait la synthèse des acceptions qui circulent aujourd'hui en soulignant deux traits essentiels : la veille est un processus continu, inscrit dans la durée,

3. François Jakobiak, *Pratique de la veille technologique*, Paris, Éd. d'Organisation, 1991.

4. *Dictionnaire encyclopédique de l'information et de la documentation*, Paris, Nathan, 1997, coll. « Nathan-Université. Réf. ».

1. *Netsurf*, n° 57, décembre 2000, p. 80-84.
2. <http://www.mmedium.com/veille/>

Claire Dartois, conservateur de bibliothèque, a été directrice adjointe à la bibliothèque du Conservatoire national des arts et métiers, avant de rejoindre le Service des documents électroniques à la Bibliothèque publique d'information. Titulaire d'une maîtrise et d'un DEA en sciences de l'éducation, elle est aussi expert auprès de la Communauté européenne, DG XIII, à la section des publications électroniques et des bibliothèques, et chargée de cours à l'université Paris VII dans le module sur la recherche documentaire. Elle a publié de nombreux articles, notamment sur le rôle d'Internet dans les bibliothèques.

et qui vise à une transmission personnalisée des données recueillies. Un autre document de référence en décline les caractéristiques en insistant sur l'importance de suivre, voire d'anticiper les innovations, et de travailler à la satisfaction des besoins exprimés ou latents des usagers⁵.

La veille documentaire à la BPI

Comment ces concepts peuvent-ils être transportés et adaptés à la réalité d'un corpus de liens déjà constitué ?

La Bibliothèque publique d'information, comme bon nombre d'établissements (la Bibliothèque nationale de France, les bibliothèques universitaires, les bibliothèques municipales, les musées...) a progressivement mis en place, à l'intérieur du site web de l'institution⁶, une sélection de liens, organisée thématiquement, et accessible tant de l'intérieur que de l'extérieur. Elle est alimentée par une équipe de neuf personnes, qui se partagent les domaines de travail sur Internet, comme les cédéroms bibliographiques et multimédias dépendant de ces disciplines, en liaison avec le service des Imprimés. Le site de la BPI, et les problèmes posés par l'introduction du réseau, ont fait l'objet d'une présentation détaillée dans

un article de 1999⁷ auquel le lecteur intéressé peut se reporter...

Nous nous proposons maintenant d'examiner les objectifs de ce type particulier de veille documentaire, les moyens techniques et humains mis en œuvre, pour nous interroger enfin sur ses perspectives et ses limites. L'expérience décrite ici est marquée par l'empirisme : s'il est possible aujourd'hui de tracer un premier bilan, le processus, par définition, est dynamique, et doit évoluer au contact d'expériences menées ailleurs, relativisant le travail entrepris. Entretenir de manière satisfaisante une collection de signets est une préoccupation qui apparaît nettement dans les listes de diffusion professionnelles francophones (adbs-info, biblio-fr, la *Lettre du bibliothécaire québécois*⁸...), et plus encore dans les publications étrangères (*Library Trends*, *Library Abstracts*, etc.), voire dans les bases de données professionnelles comme LISA (Library Information Science Abstracts).

Des logiciels de vérification des liens

Le premier problème qui se pose, le plus évident peut-être, est celui de la mise à jour des adresses URL (Uniform Resource Locator) qui ont été sélectionnées. Par définition, les adresses Internet sont volatiles, disparaissent, les contenus changent, ne répondant plus aux critères de choix. Chaque personne du service des Documents électroniques est responsable des liens de ses domaines, qui doivent être valides, si l'on ne veut pas que le lecteur soit confronté à une fin de non-recevoir, la fatale « erreur 404 » indiquant que la page ne peut être chargée.

Comme dans les collections papier, le désherbage est de mise : adresses caduques, rubriques obsolètes doivent être éliminées. Pour ce faire, l'interrogation « manuelle » de chaque lien est une tâche bien trop lourde ; des outils automatisés permettent une interrogation de tout le corpus pour le repérage des adresses à problèmes. À la BPI, c'est le logiciel Pagemill qui remplit cette fonction, sans disposer de tous les raffinements techniques utiles (les messages du type « *une erreur réseau est survenue* », ou « *système trop occupé* » sont équivoques et demandent une seconde vérification). Ces moyens techniques se développent rapidement ; dès 1998, la liste biblio-fr publiait une synthèse de vérification d'URL⁹, articulée en plusieurs rubriques : sites d'information, outils sur Mac, sous Unix, sur PC (ou Linux), et plates-formes non précisées. Comme il est d'usage dans cette liste alimentée par les utilisateurs, ceux-ci commentent les prestations offertes et précisent s'il s'agit de logiciels gratuits (*free-ware*) ou payants (*shareware*).

Olivier Andrieu, responsable du remarquable site *Abondance*, et de la lettre hebdomadaire d'information « *Actu-Moteurs* » dédiée aux outils de recherche, propose également une sélection de sites web et de logiciels pour la détection des liens rompus¹⁰, avec, à chaque fois, une fiche descriptive. Ceux-ci sont malheureusement encore des outils américains, à l'interface peu ou pas traduite en français. Si l'on examine en détail, à titre d'exemple, l'un de ces logiciels, *Xenu's Link Sleuth*¹¹, on peut parcourir (et imprimer) une documentation technique d'une dizaine de pages composée d'une description, de la manière de télécharger le logiciel, des questions couramment posées (FAQ), des « bogues » possibles, des

5. *Dictionnaire encyclopédique des sciences de l'information et de la communication*, sous la dir. de Bernard Lamizet et Ahmed Silem, Paris, Ellipses-Marketing, 1997.

6. Consultable à l'adresse <http://www.bpi.fr>

7. Dominique Baude, Claire Dartois, « Internet à la Bibliothèque publique d'information », *Libre-Accès*, octobre 1999, n° 20, p. 40-65.

8. <http://listes.cru.fr/www/info/adbs-info>
<http://listes.cru.fr/www/arc/biblio-fr>
<http://www.sciencepresse.qc.ca/lbq/lbq.html>

9. <http://listes.cru.r/www/arc/biblio-fr/1998-03/msg00124.html>

10. <http://www.abondance.com/ressources/verifliens.html>

11. <http://home.snafu.de/tilman/xenulink.html>

aménagements prévus par le concepteur et de l'historique du produit. L'information semble exhaustive, mais est celle du producteur.

Il n'est donc pas toujours aisé de se repérer pour choisir l'outil le plus adapté aux besoins dans une offre à croissance exponentielle. Un tableau comparatif des outils de vérification des liens pour Windows 95-98¹² permet un regard plus large sur l'offre, indiquant pour chaque produit la date et le numéro de la version, son « poids », son prix (en dollars), ses modalités d'installation avec une analyse des fonctionnalités, et l'accès direct au téléchargement de l'application.

Le web invisible

À ce stade, quel que soit le logiciel installé, il faut, après élimination des liens caducs, rectifier l'information destinée à l'utilisateur, à savoir supprimer les descriptifs de sites, et, si l'adresse a migré sur la toile, partir à sa recherche, en utilisant moteurs, méta-moteurs, voire agents intelligents. Une connaissance très pointue des potentialités respectives de ces différents outils est indispensable au veilleur, ou au « guetteur », comme disent joliment nos confrères québécois. Certains sont plus adaptés à la recherche de littérature grise, d'autres à la recherche d'images ; la syntaxe avancée varie (utilisation ou non des guillemets, exclusion ou inclusion de certains termes, recherche sur le titre, l'URL, le domaine, dédoublonnage des résultats et classements selon différents critères...). La littérature professionnelle fournit des articles très fouillés, comme par exemple une comparaison point à point des métamoteurs ou une évaluation des nouveaux produits¹³.

N'oublions pas cependant que le web invisible (que les Anglo-Saxons

nomment plus judicieusement « *Deep Web* »), dont la taille se situerait, si l'on en croit l'étude de *Netsources* de juillet-août 2000, à 550 milliards de documents, représente les masses d'information non indexées par les moteurs de recherche, et qui sont, selon le jargon documentaire, de l'ordre du « silence » dans les résultats de recherche. Cet iceberg, dont la taille est variable, est donc un obstacle considérable à la navigation...

Développer les rubriques, trouver et valider de nouvelles ressources est un axe essentiel pour que la collection de signets soit vivante et attractive

Le travail d'élagage n'est, bien entendu, que la première étape de la surveillance globale des liens. Souvent long, parfois fastidieux, il est pourtant une condition *sine qua non* de la fraîcheur et de la fiabilité des informations proposées. Parfois même, le site est notre seul contact avec des utilisateurs distants, qui ne fréquentent pas les collections, et qui se font une image instantanée des approximations et des manques de la sélection de liens de la bibliothèque.

Un partage de ressources

Développer les rubriques, trouver et valider de nouvelles ressources est un axe essentiel pour que la collection de signets soit vivante et attractive, car si « *une bibliothèque est un organisme vivant* » (Ranganathan), elle doit impérativement se développer.

Le travail de prospection est varié : il passe par le dépouillement de la

presse spécialisée, des cahiers multimédias des quotidiens, qui offrent de plus en plus de sélections commentées de sites, la consultation des listes de diffusion, et un regard attentif sur les choix faits par d'autres qui peuvent compléter les nôtres. Le supplément annuel de la revue américaine *Choice*¹⁴ propose des analyses conséquentes, faites par des spécialistes des domaines, des sites web existants : il est à regretter que son équivalent français n'existe pas encore... l'offre est en effet exponentielle, touffue, difficilement cernable, et le bibliothécaire dans une situation (déstabilisante) « d'hyper-choix », pour reprendre la formule de Toffler.

Par ailleurs, et ceci est plus préoccupant, cette offre est extrêmement inégale : des sites personnels farfelus côtoient des travaux universitaires, des annonces commerciales, des sites pornographiques, violents, que les moteurs de recherche sélectionnent de la même façon. Borges décrit ainsi un univers sans limites : « *Son livre s'appelait le livre de sable, parce que ni ce livre ni le sable n'ont de commencement ni de fin.* » Cette image vaut pour la profusion du cybermonde, et son aspect à la fois immatériel et illimité. Il écrit plus loin¹⁵ : « *Le nombre des pages de ce livre est exactement infini. Aucune n'est la première, aucune n'est la dernière.* »

À la BPI, il s'agit à la fois d'une veille individuelle (les secteurs sont partagés) et collective, puisque c'est un travail d'équipe, ce qui implique de redistribuer une information glanée dans une revue ou sur une liste de diffusion aux personnes intéressées. Sans ce partage des ressources, bien connu des documentalistes, la prospection efficace serait impossible. Différents sites nous apparaissent comme des références : les unités régionales de formation à l'information scientifique et technique

12. <http://winfiles.cnet.com/apps/98/htmlverify.html>

13. *Netsources*, n° 18, janvier-février 1999.

14. Web IV, *Current Reviews for Academic Libraries*, 2000, vol. 37. Supplement.

15. Jorge Luis Borgès, *Le livre de sable*, Paris, Gallimard, trad. française 1978.

Quelle veille pratiquer sur un site web ?

(URFIST), avec leur remarquable travail de sélection et de présentation des liens, l'Association des professionnels de l'information et de la documentation (ADBS), l'École nationale supérieure des sciences de l'information et des bibliothèques (Enssib), l'Institut national des sciences appliquées de Lyon (INSA), entre autres, dans leur effort de transmission « pédagogique » des découvertes. Les guides annuaires réalisés par des experts (*About.com*, *The Argus Clearinghouse*, *Bubl Link*, *The www Virtual Library*...) proposent des choix raisonnés, commentés, et régulièrement mis à jour. La BnF dans ses signets, propose en outre l'accès à une rubrique intitulée « Réalisation de répertoires sélectifs »¹⁶, qui est « une amorce de veille documentaire ». Le projet « Desire » (Development of a European Service for Information on Research and Education), pour n'en citer qu'un, favorise la création de répertoires par domaine, destinés aux chercheurs, et est impulsé par quatre pays de l'Union¹⁷.

Un éventail structuré de propositions

Bien entendu, chaque ressource nouvelle est ensuite évaluée et testée ; peu à peu, à partir de la littérature existante et surtout de l'expérience de l'équipe, a été construite une grille d'évaluation des sites Internet (cf. encadré). Les ressources retenues sont ensuite présentées selon un cadre spécifique, qui s'applique à tous les domaines, mais qui autorise une certaine souplesse : l'art ou l'ethnologie, par exemple, ont une rubrique « musées », qui n'apparaît pas en économie ni en mathématiques. Cette formalisation permet de même au lecteur de retrouver une homogénéité dans les présentations, et de faire le lien avec les collections sur

16. <http://www.bnf.fr/web-bnf/liens/moteurs/realisation.html>
17. <http://www.desire.org>

GRILLE D'ÉVALUATION D'UN SITE INTERNET

1. Le contenu

- Le thème du site est-il clairement défini ?
- Quelle est la précision de l'information ?
- L'information est-elle datée ? Quelles sont les fréquences de mises à jour ?
- Les documents anciens sont-ils archivés ?
- Y a-t-il des liens utiles vers d'autres sites ? Sont-ils en cohérence avec le contenu ?
- Ces liens sont-ils décrits et évalués ? Sont-ils à jour ?

2. Les auteurs

- S'agit-il d'une institution connue (université, ministère, etc.) ? et spécialiste de ce type d'information ?
- Est-ce au contraire un site personnel ? L'auteur est-il reconnu dans la discipline ?
- Peut-on contacter ces auteurs (mél, adresse postale) ?

3. L'organisation de l'information

- Y a-t-il un sommaire ou un plan du site ?
- Peut-on remonter à la « racine » du site dans l'adresse URL ?
- La navigation est-elle aisée (touches haut de page, retours, etc.) ?
- Le site contient-il des éléments bibliographiques ?
- Connaît-on la taille du site (nombre d'éléments indexés, de pages) ?

4. La recherche

- Un moteur de recherche interne est-il disponible ?
- Sur quels critères peut-on chercher (auteur, mots-clés, date...) ? Ces critères sont-ils explicités (exemples) ?
- Des opérateurs de recherche sont-ils proposés ?
- Du texte intégral est-il disponible ?
- Les résultats de la recherche semblent-ils pertinents ? Comment sont-ils présentés (classement) ? Quels sont les temps de réponse ?

5. L'accessibilité

- Le texte est-il facilement imprimable ?
- Le temps de chargement des images et du texte est-il long ?
- Le site contient-il des parties payantes ?

d'autres supports dans les espaces de lecture. Élargir ainsi l'éventail des propositions obéit à deux objectifs, qui ne sont pas contradictoires : gagner de nouveaux publics, améliorer l'accès à l'information, et valoriser les collections existantes. Une revue papier dépourvue d'index sera mieux consultée si l'on propose cet outil de repérage sur Internet ; si la bibliothèque n'a pas vocation à la conservation - c'est le cas de la BPI -, permettre un accès aux archives en texte intégral sur le réseau facilite le travail du lecteur. D'où la nécessité d'une réflexion sur l'offre proposée sur différents supports (papier, cédéroms, microformes, sites web...), visant à un équilibre global et à une satisfaction correcte des besoins.

Des rubriques particulières peuvent être créées en fonction de l'actualité (résultats du bac, du re-

censement, festivals de l'été, etc.) - thèmes ponctuels à durée de vie réduite -, comme peuvent se construire des thématiques plus larges, inspirées d'une demande, illustrant le principe de la veille comme satisfaction des besoins exprimés ou tacites. La rubrique aidant aux recherches d'emploi¹⁸, créée fin 1996, dans l'esprit du « rôle social » de l'établissement, est très largement sollicitée par les usagers.

La connaissance des besoins du public en matière de nouveaux signets suppose le recueil direct aux bureaux d'information des suggestions (ainsi

18. Claire Dartois, « Le site Internet sur l'emploi à la BPI », *Bibliothèques et emploi : le rôle des bibliothèques dans l'information sur l'emploi*, actes du séminaire européen organisé par la BPI, avec la DLL, en collab. avec le Goethe Institut et le British Council, Lille, 27 et 28 novembre 1997, Paris, BPI, 1998, coll. « La BPI en actes », p. 233-237.

que dans le cahier des lecteurs), les remarques des collègues, et le dépouillement des messages envoyés au responsable du site : nécessaire rétroaction pour l'amélioration globale du produit, tant sur le fond que sur la forme, pour l'arborescence, la navigation et la mise en page. Des enquêtes sociologiques, quantitatives et qualitatives permettent, quand la bibliothèque a les moyens de les organiser, le recueil d'un matériel de premier choix.

Les outils de mesure d'audience d'un site¹⁹ sont abondamment décrits dans la presse spécialisée, et quantifient les connexions des utilisateurs, en donnant aussi d'autres indications précieuses : temps de connexion (une visite d'une minute ne veut rien dire), refus de connexion (sites trop occupés, adresses erronées), permettant de dresser un tableau des liens les plus fréquentés.

Cependant, il n'est pas question dans une bibliothèque à vocation encyclopédique, de privilégier tel aspect du savoir et d'en ignorer d'autres, moins légitimes ou moins plébiscités. Ainsi, les mesures, quand elles sont possibles, permettent une vision globale des fréquentations. Il est bon cependant de garder à l'esprit cette pique : « *Il y a la vérité, il y a le mensonge, et il y a les statistiques* »...

Les moyens humains

Après cet aperçu des moyens techniques employés dans la veille, évoquons les moyens humains mis en œuvre, en tout cas dans le contexte décrit. La Bibliothèque publique d'information a la chance de disposer d'une équipe chargée des documents électroniques, Internet et cédéroms, ce qui facilite la synergie. Le gestionnaire des signets, membre de l'équipe, coordonne les mises à jour, signale les problèmes, opère le transfert des

fichiers Word en HTML, et veille à la bonne circulation à l'intérieur des données qui sont en expansion, comme à l'homogénéité des rubriques. Pour tous les membres du service, le travail de veille est lourd, comme celui d'actualisation des connaissances. Compte tenu des

**Vouloir constituer
de nos jours
une bibliothèque
encyclopédique apparaît
comme une sorte
de gageure,
d'idéal inaccessible,
du moins si l'on prétend
viser à l'exhaustif**

autres tâches, dont le service public, l'impression de travail « en mosaïque » est parfois décourageante, comme l'est l'inquiétude devant les (nombreux) fichiers en attente. Percec²⁰ le dit avec force : « *Comme les bibliothécaires borgeiens de Babel qui cherchent le livre qui leur donnera la clé de tous les autres, nous oscillons entre l'illusion de l'achevé et le vertige de l'insaisissable.* »

La mise en commun des informations recueillies est aussi une condition indispensable à l'économie de temps et d'énergie, contre les tentations de la thésaurisation ou de l'isolement. La formation continue, l'entretien des connaissances représentent en effet un investissement non négligeable. La floraison des stages organisés par les instituts de formation des bibliothécaires sur les ressources thématiques du web en est un signe.

Enfin, un travail suivi avec les responsables des acquisitions est la base d'une politique documentaire de coordination des ressources, dans le cadre d'un espace (les locaux de la bibliothèque), d'un budget, d'équipes qui œuvrent aux différents points de la chaîne documentaire, d'une carte documentaire (l'offre des autres établissements), au service de publics qui structurent à la fois offre et demande. Sans cet effort de synergie, un ensemble de liens web ne sera qu'un ajout mal intégré, mal utilisé, tant par les collègues aux bureaux d'information, que par des usagers, perdus dans un labyrinthe virtuel.

Limites et perspectives

Quelles sont enfin les limites et les perspectives de cette veille un peu particulière ? Henri-Jean Martin, dans un article intitulé « Éloge de la perfection »²¹, rappelait que « *vouloir constituer de nos jours une bibliothèque encyclopédique apparaît comme une sorte de gageure, d'idéal inaccessible, du moins si l'on prétend viser à l'exhaustif* ». Remarque à méditer, plus encore dans le monde virtuel, où les données sont labiles, où la mouvance est reine, et où la fiabilité est parfois sujette à caution. Une collection de liens est toujours « en train de se faire », et peut difficilement, à un instant *t*, être irréprochable. Les principes de la « *qualification de l'information* », selon Jean-Pierre Bernat²², qualités intrinsèques, comme l'exactitude, la fraîcheur, l'environnement (contexte) s'ajoutent aux critères formels (rapidité, présentation, prix raisonnable) devraient raisonnablement s'appliquer à cette offre particulière... Limites humaines enfin, puisque le temps nous est compté et que la tâche s'apparente au remplissage d'un nouveau tonneau des Danaïdes.

19. <http://www.abondance.com/ressources/audience.html>

20. Georges Percec, *Penser/Classer*, Paris, Hachette, 1985, p. 42.

21. « La Bibliothèque, miroir de l'âme, mémoire du monde », *Autrement*, n° 121, avril 1991.

22. *Bases*, n° 146, janvier 1999.

Quelle veille pratiquer sur un site web ?

Une veille intelligente sous-entend, de plus en plus, *une coopération entre les institutions*. Le projet SiteBib, comme le groupe « Enrichi/Groupe de travail signalement partagé des ressources en ligne » amorcé plus récemment à Lyon 3 par Thierry Samain²³, est bâti sur un partage de la prospection, de la sélection

et de l'analyse des ressources. Les bibliothèques universitaires françaises présentent leur sélection dans un portail commun²⁴ : évolution donc dans l'approche partagée... Les petits établissements, qui n'ont qu'une équipe réduite (parfois à une personne), n'ont ni le temps ni les moyens d'alimenter une offre volumi-

neuse ; puisqu'il n'y a pas de *copyright* sur les liens, les choix faits par des institutions mieux dotées peuvent servir de base, pour être ensuite adaptés à la réalité locale. Sans ces nouvelles habitudes de travail intellectuel et de partage, il semblera bien futile, chacun pour soi, de... veiller au grain.

Janvier 2001

23. <http://www.abf.asso.fr/enrichi>
Voir aussi, dans ce numéro, l'article de Thierry Samain, « Du catalogue enrichi au métacatalogue », p. 82-86.

24. <http://www-scd.univ-lyon3.fr/busignets>