



Automatiser sa veille sur Internet

Par Julien Flandrois, Directeur Commercial chez KB CRAWL SAS, éditeur français d'outils de veille et d'intelligence économique sur Internet depuis 1995.

Pourquoi automatiser sa veille sur Internet ?

La collecte de l'information disponible sur Internet constitue un atout pour celui qui peut l'exploiter : lancement d'un produit concurrent, parution d'un brevet, publication d'un appel d'offre ou d'une nouvelle réglementation. Mais la surveillance « manuelle » des sites Web est chronophage et il est aujourd'hui possible de l'automatiser.

Comment automatiser sa veille sur Internet ?

Pour accompagner une automatisation de la veille sur Internet, il est essentiel d'adopter une démarche méthodique répondant au cycle de la veille. Dans ce cycle, le logiciel de veille KB Crawl assure la mission de surveillance de sites et la collecte d'informations pertinentes.

L'automatisation accroît considérablement l'efficacité de la veille sur Internet par rapport à une veille « manuelle » du Web.

Surveiller le Web visible et invisible

Outre les possibilités de surveillance de sites, KB Crawl est capable de surveiller le Web invisible qui constitue la plus grande partie des informations disponibles sur Internet, ouvrant un champ d'application considérable : veille concurrentielle, réglementaire, technologique, commerciale, etc.



Automatiser sa veille sur Internet

Pourquoi automatiser sa veille sur Internet ?

Internet regorge d'informations en accès libre, diversifiées et récentes. La collecte de cette information « fraîche » constitue un atout pour celui qui peut l'exploiter : lancement d'un produit concurrent, parution d'un brevet, publication d'un appel d'offre ou d'une nouvelle réglementation.

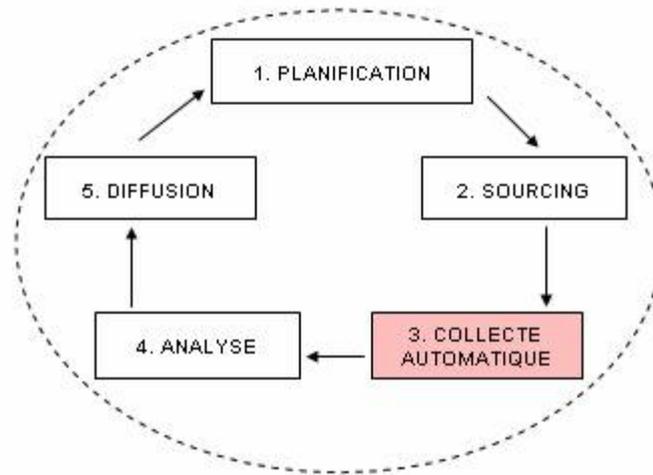
On constate que les internautes retournent fréquemment sur les mêmes sites pour voir si une nouveauté est apparue. Ces sites sont rassemblés dans leurs carnets d'adresses (bookmarks ou favoris). Mais sur les quantités de pages passées en revue quotidiennement, très peu sont mises à jour régulièrement. Ce qui amène les internautes à parcourir les sites la plupart du temps pour rien. Cette tâche est chronophage et à faible valeur ajoutée.

Les internautes doivent alors se résoudre au mieux à consulter manuellement quelques pages, voire finalement à délaisser le Web. Mais, même s'il est difficile de surveiller ce flux continu d'informations se déversant sur le Web, doit-on se résoudre à laisser passer des opportunités commerciales et ignorer les avancées de la concurrence ?

Sans aucun doute non. D'autant plus qu'il est aujourd'hui possible d'automatiser sa démarche de veille sur Internet. Accompagné d'une approche méthodique et rationnelle, le logiciel de veille automatique KB Crawl propose un outil capable de surveiller les sites Web, c'est-à-dire de détecter des nouveautés pertinentes sur des sites prédéfinis et d'alerter les responsables intéressés.

Comment automatiser sa veille sur Internet ?

Afin d'éviter de se perdre dans les profondeurs du Web, il est essentiel d'adopter une démarche méthodique de veille qui est celle des veilleurs, les professionnels de l'information. Cette démarche constitue le cycle de la veille.



Le cycle de la veille sur Internet

1. Planification de la veille sur Internet en fonction de la problématique retenue
Ex : Veille réglementaire

2. Définition des sources pertinentes à surveiller
Ex : www.legifrance.fr et www.europa.eu.int

3. Surveillance automatique des sites avec KB Crawl
Ex : Mise en place d'une surveillance quotidienne fondée sur des mots-clés

4. Analyse et traitement des informations brutes
Ex : Parution d'un projet de directive européenne

5. Diffusion de l'information
Ex : Publication de l'information sur l'Intranet de l'entreprise



Automatiser sa veille sur Internet

Contrairement aux recherches ponctuelles effectuées habituellement avec les moteurs de recherche comme Google ou Yahoo!, la démarche de veille doit être persévérante. Elle s'enrichit au fil du temps et doit s'adapter continuellement aux problématiques soumise : veilles concurrentielle, réglementaire, commerciale, technologique, etc.

Les différentes étapes décrites ci-dessus ne peuvent pas être automatisées dans leur intégralité. Le veilleur reste au coeur du système. KB Crawl joue un rôle important sur les tâches à moindre valeur ajoutée pour le veilleur telles que la surveillance récurrente de sites et la détection de nouveautés. KB Crawl doit être considéré comme « un pilote automatique ».

Le veilleur est en charge des missions de « décollage » et d'« atterrissage ».

Le fonctionnement de KB Crawl se décline en quatre étapes :

1. Le veilleur indique à KB Crawl les sites à surveiller et définit un périmètre de surveillance à l'intérieur des sites les plus importants. Il lui précise les types d'alertes qu'il souhaite recevoir (mots-clés, définition des requêtes, etc.).
2. À partir d'une première version enregistrée, KB Crawl visite à la fréquence souhaitée (allant de la minute à la semaine) ces sites et détecte les nouveautés.
3. Lorsqu'une nouveauté est pertinente, KB Crawl envoie automatiquement par courrier électronique un bulletin d'alerte aux personnes concernées.
4. KB Crawl indexe et archive les nouveautés des sites dans une base de données afin que les veilleurs puissent faire consulter à tout moment les résultats de veille.

Avec KB Crawl, l'efficacité de la veille sur Internet est multipliée par un facteur 3 environ. L'investissement du logiciel est rentabilisé en moins de deux mois par rapport au temps passé par les veilleurs à explorer manuellement les sites. L'outil peut s'avérer vital pour l'entreprise quand il permet de détecter une information commerciale cruciale qui aurait échappé à une veille manuelle.



Automatiser sa veille sur Internet

Surveiller le Web visible et invisible

Les moteurs de recherche actuels ne référencent dans leurs résultats qu'une toute petite partie des informations disponibles sur le Web (moins de 10%). Le reste constitue le Web dit « invisible » car les robots des moteurs de recherche ne repèrent pas ces pages et ne permettent donc pas aux internautes de les retrouver via Google ou Yahoo ! Ce sont, par exemple, les pages sans lien, isolées sur la toile, les pages nécessitant une authentification, les pages dynamiques renvoyant des résultats de recherche, etc.

Cependant, le Web invisible contient une information riche dont il est difficile de faire abstraction. L'un des points forts de KB Crawl est de surveiller le Web invisible, notamment grâce à la gestion des « formulaires ». En effet, de plus en plus d'informations sont accessibles via des formulaires, c'est-à-dire la saisie de données dans une page pour obtenir des résultats, comme pour des annonces immobilières, des brevets, des textes réglementaires, nécessitant l'acquittement d'une redevance d'accès.

KB Crawl est alors capable d'enregistrer tout type de formulaire puis de répéter cette opération autant de fois qu'il le faut afin de détecter les nouveautés. C'est ainsi que l'on peut automatiser des recherches quotidiennes « manuelles » itératives comme la détection de nouveaux appels d'offres, par exemple.

Maîtriser l'information stratégique de son environnement

Tout le monde fait de la veille. Depuis quelques années, les grands groupes industriels ont institué le poste de veilleur. Mais pour beaucoup de structures, les fonctions traditionnelles de l'entreprise demeurent concernées : la documentation, les études, la R&D ou encore les services de marketing et communication.

Chacun peut aujourd'hui automatiser sa surveillance du Web en fonction de ses centres d'intérêt. Le tableau ci-dessous recense quelques cas concrets de veille sur Internet en fonction des problématiques soulevées.

Fonctions	Type de veille	Intérêts	Exemples de sites
<ul style="list-style-type: none"> - Documentation - Etudes/Stratégie - Com & Marketing 	CONCURRENTIELLE	<ul style="list-style-type: none"> - Surveiller les dernières publications de ses concurrents - Détecter le lancement de nouveaux produits 	Ceux des concurrents essentiellement
<ul style="list-style-type: none"> - Juridique - R&D - Documentation 	REGLEMENTAIRE	<ul style="list-style-type: none"> - Détecter les nouvelles réglementations - Agir auprès des institutions avant la mise en application des lois 	www.legifrance.fr www.europa.int.eu
<ul style="list-style-type: none"> - R&D - Documentation 	TECHNOLOGIQUE	<ul style="list-style-type: none"> - Surveiller la publication de nouveaux brevets - Collecter les nouvelles études - Suivre les travaux des chercheurs 	ep.espacenet.com www.inist.fr
<ul style="list-style-type: none"> - Documentation - Etudes/Stratégie 	SECTORIELLE	<ul style="list-style-type: none"> - Surveiller l'actualité de son secteur - Suivre ses fournisseurs 	Les sites de news généralistes et spécialisés (CCI)
<ul style="list-style-type: none"> - Etudes/Stratégie - Com & Marketing 	SOCIETALE	<ul style="list-style-type: none"> - Détecter les nouvelles tendances - Evaluer l'image de l'entreprise 	<ul style="list-style-type: none"> - Forums de discussion - Newsgroups
<ul style="list-style-type: none"> - Com & Marketing - Documentation - Commercial 	COMMERCIALE	<ul style="list-style-type: none"> - Surveiller les appels d'offre - Suivre l'actualité des clients - Détecter de nouvelles opportunités dans la presse 	djo.journalofficiel.gouv.fr Sites des clients