

BLOG ACTULLIGENCE

KB Crawl SA lance KB Crawl V4

Écrit par Frédéric Martinet
15-03-2009

Vendredi 13 je me suis rendu sur [Search 2009](#) histoire de voir un peu les dernières évolutions des outils de recherche d'information, d'y recroiser quelques connaissances et de [distribuer mes nouvelles cartes de visite](#) tout juste sorties de la presse.

J'ai profité de mon passage pour assister à la conférence de presse de [KB Crawl](#) qui présentait la quatrième version de son logiciel de veille en avant première à sa commercialisation qui est annoncée à la fin du mois, lors de [Documentation](#).

Quelles nouveautés pour KB Crawl 4 ?

Tout d'abord il s'agit d'une toute nouvelle version! A part les principes de fonctionnement de KB Crawl 3 et les principes de fonctionnement inhérents à tout logiciel de veille qui restent inchangés, sous le capot, tout le code informatique a été redéveloppé pour cette nouvelle version. On passe ainsi du langage Delphi à du Microsoft .Net ce qui permettra de faciliter les évolutions du produit et d'avoir aussi une meilleure stabilité, une architecture distribuée permettant d'améliorer les temps de crawl et une meilleure montée en charge le tout s'appuyant sur une base de données SQL.

Ce logiciel a par ailleurs été restructuré afin d'être aussi bien un logiciel monoposte qu'un logiciel collaboratif, ce qui faisait cruellement défaut à la V3 qui par exemple gérait difficilement les accès concurrents à une source lors d'une utilisation collaborative.

KB Crawl permettra désormais le multi-threading, c'est à dire la possibilité de crawler plusieurs pages à la fois et aussi la possibilité de travailler sur le soft pendant qu'il crawle.

Côté fonctionnalités, KB Crawl 4 facilitera le paramétrage des sources derrière un formulaire (authentification, et interrogation) qui pouvait s'avérer difficile sur la V3 et nécessitait parfois le recours à un module complémentaire, KB Nav Record.

L'autre amélioration fonctionnelle, c'est la découpe d'une page web de façon graphique afin de créer des fichiers Excel ou XML qui pourront par exemple comporter les titres, contenu, auteur et date de tous les articles présents au sein d'une page d'actualité ou bien encore de faire une veille tarifaire sur le site d'un cyber marchand.

Enfin, KB Crawl 4 permettra le crawl et l'indexation des caractères non latins (arabes, russe, ...) grâce au moteur d'indexation Lucène.

Côté "nouveautés" marketing : nouvelles plaquettes, annonce de la gamme de produits : KB Crawl 4, KB Crawl suite (modules complémentaires) et KB Platform.

KB Platform est la plateforme de veille de KB Crawl SA, plateforme basée sur le CMS open source Joomla, qjui permet de stocker, de partager et d'enrichir les informations collectées via KB Crawl 4.

Voilà pour ce qui en est des messages passés pendant la conférence de presse.

En guise de conclusion :

Le produit semble prometteur grâce au module de découpe de pages web, grâce l'amélioration de la capacité de crawl et à la gestion du multi-threading.

Les prérequis pour faire tourner KB Crawl semblent importantes si vous l'installez en configuration monoposte puisqu'il vous faudra pas mal de mémoire vive, le framework .Net 3.5, IE 7, et d'autres choses ce qui est loin d'être le cas aujourd'hui dans la plupart des entreprises que je visite. La configuration client-serveur serait donc sans doute à privilégier.

Il est dommage qu'à 15 jours de la commercialisation nous n'ayons eu droit qu'à un seul screenshot du logiciel, et à de simples screenshots de sites web, en dehors de l'interface de KB Crawl 4, qui peuvent être mis en surveillance (Facebook, LinkedIn, ...) Malheureusement, pas la plus petite démonstration live du produit. Pourtant, moi je dis qu'une petite démonstration vaut mieux qu'un long discours surtout lorsqu'il s'agit de logiciel de veille. Une version d'essai à destination des journalistes, qui j'en suis sûr sont désireux de faire leur travail consciencieusement et de se faire leur propre idée du logiciel, serait un plus.

Aller plus loin :

[Voir le communiqué de presse](#)