



A U C O U R A N T

Une source continue des tendances sociales dans l'environnement des technologies de l'information du secteur public

Mars 2004

Le 12 mars 2004

Plonger dans le « Web profond »

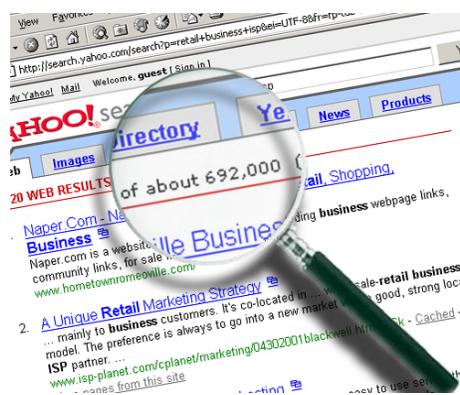
Nous savons tous qu'Internet renferme un trésor d'information. Toutefois, nous ne sommes pas tous d'habiles courtiers quand vient le temps d'accéder à cette richesse ou de la gérer. Si vous savez comment taper un mot ou deux et cliquer avec une souris, vous pouvez appeler Google et Yahoo! à l'aide, mais vous serez peut-être découragé en voyant les millions d'occurrences qu'ils vont vous présenter. Mais ce n'est pas tout : ces recherches n'explorent que la surface d'Internet. Et il existe de nombreux chercheurs qui ne demandent qu'à sonder les profondeurs.

Il y a 90 téraoctets de données que les outils de recherche traditionnels n'arrivent pas à trouver.

Faisons donc un plongeon exploratoire dans le « Web profond », c'est-à-dire dans les 99 % et plus du contenu connu d'Internet auxquels même les principaux moteurs de recherche n'ont pas accès. En dehors des transactions ciblées effectuées au moyen d'outils de recherches propres aux sites, il y a 90 téraoctets de données, stockées en grande partie dans des bases de données, que les outils de recherche traditionnels n'arrivent pas à trouver. Il se passe beaucoup de choses dans le domaine de la recherche dans le Web profond. On essaie de trouver des moyens d'insérer ces données dans les recherches générales, en sachant que si la proportion de contenu consultable augmente, la qualité des recherches augmentera aussi.

Voilà une bonne nouvelle pour la fonction publique! En permettant aux citoyens d'avoir un plus grand accès au contenu d'Internet, la transparence en direct du gouvernement s'en

trouvera accrue. De plus, les fonctionnaires, qu'ils soient employés de première ligne ou d'arrière-guichet, profiteront éventuellement d'un accès plus efficace à l'information dont ils ont besoin, par exemple pour aider un client ou pour analyser des données.



En ce qui concerne la peur d'être submergé par les millions de résultats qui déferlent lorsque vous cliquez sur le bouton « Recherche Google », il existe des moyens de transformer le raz-de-marée qui noie les résultats en une onde ciblée et facile à gérer. BrightPlanet offre sur son site Web un guide téléchargeable de 85 pages, intitulé Guide to Effective Searching of the Internet (Offert en anglais seulement). Le sommaire du Guide est en lui-même très utile. On suggère notamment d'effectuer les recherches en utilisant des noms plutôt que des verbes, et de préciser de six à huit mots dans l'interrogation.

Pourquoi le format RSS ?

Le Web est une excellente source d'information, mais quand il faut se tenir au fait des nouvelles qui affluent dans les sites d'actualités et les blogues, on est bien vite dépassé. C'est là que le format RSS entre en jeu et devient utile pour les fonctionnaires. Le format RSS (« Really Simple Syndication » ou « Rich Site Summary ») est un format qui aide à gérer diverses sources d'information en livrant dans votre

DANS CE NUMÉRO

Plonger dans le « Web profond »

Pourquoi le format RSS ?

Technologie de la traduction : rien ne vaut le cerveau humain

Les logiciels sociaux ou comment faire fructifier votre capital organisationnel

Pensez-y à deux fois avant d'envoyer ce courriel!

La syndication de contenu vraiment simple (RSS)

boîte de réception les nouvelles les plus récentes et des résumés d'articles. En plus d'afficher de façon commode l'information dont vous avez besoin, le texte en format RSS est personnalisable, libre de publicités agaçantes et mis à jour automatiquement à partir des sites de votre choix.

Petit guide du format RSS

Vous aimez l'idée, certes. Mais comment faut-il s'y prendre?

1. D'abord, téléchargez et installez un lecteur RSS comme Sharpreader ou Bloglines.
2. Allez à un site qui vous intéresse, puis cherchez le lien « RSS » ou « XML ».
3. Cliquez sur le lien et copiez l'URL de la page dans votre lecteur RSS.
4. Une liste de titres devrait s'afficher. Cliquez sur le bouton d'inscription pour ajouter la source à votre liste.
5. Les titres devraient être envoyés dans votre boîte de réception.

Pour une explication détaillée du processus de configuration des

sources RSS, vous pouvez lire les instructions du magazine Slate de MSN dans l'article « How to Speed-Read the Net » et celles du Washington Post dans l'article « Refining Paperless News ».



Le 5 mars 2004

Technologie de la traduction : rien ne vaut le cerveau humain

Que penser de cette technologie de pointe? L'aéroport de Narita, à Tokyo, procède au lancement d'un traducteur de poche. Cette initiative s'inscrit dans un effort visant à devenir l'aéroport le plus moderne au monde. Les visiteurs pourront louer cet ordinateur de poche capable de traduire instantanément l'anglais parlé vers le japonais et vice versa. L'appareil comprend sans problèmes les interférences linguistiques, les expressions familières, les accents et les formes linguistiques.

Toutefois, lorsqu'il est question de traduction, aucun ordinateur ne peut se comparer au cerveau humain. La technologie ne peut pas remplacer les traducteurs mais elle peut toutefois les aider. On a tenté, pendant de nombreuses années, de construire un superordinateur capable de traduire couramment n'importe quelles langues. Par après, l'accent a plutôt été mis sur la recherche d'outils de pointe qui aideront les traducteurs.

Cette approche a été adoptée par les gouvernements du Canada et des États-Unis. Les É.-U. procèdent au lancement d'un nouveau bureau fédéral, connu sous le nom de National Virtual Translation Center qui regorge de nouvelles technologies d'aide à la traduction. Au coeur des mesures prises par le centre, on retrouve des outils capables de convertir, d'analyser et de traiter de la matière brute. La mémoire de traduction est un autre nouvel outil d'aide qui permet de sauver du temps et de l'énergie en associant de nouveaux segments de textes à traduire avec des segments de textes correspondants qui sont déjà traduits.



Améliorer l'alphabétisme en ligne

La principale préoccupation qui ressort, lorsqu'on envisage d'utiliser Internet comme une plate-forme pour communiquer avec les jeunes Canadiens, est que cet environnement en ligne ne rejoint que les jeunes qui connaissent le Web et qui sont alphabètes. Un projet innovateur conçu par les bibliothèques Dumfries and Galloway au Royaume-Uni aide à faire tomber ces barrières. Le projet, intitulé Get a Life, est un programme interactif en direct qui enseigne l'écriture créative et l'utilisation de websmarts dans un réseau fermé. Ce programme s'adresse aux jeunes âgés entre 10 et 14 ans et il a fait l'objet d'éloges de la part des jeunes qui y ont participé. Get a Life figure sur la liste du prix CILIP/LiS Libraries Change Lives 2004.



Analyser une épidémie

Vous êtes vous déjà demandé comment un contenu affiché en direct peut « infecter » autant de personnes? À cet effet, des études récentes démontrent que tout est une question d'emplacement. Habituellement, l'épidémie frappe lorsqu'une idée arrive à se tracer un chemin au travers d'une chronique Web dont le volume du trafic est très important.

Les auteurs d'un carnet Web empruntent régulièrement le contenu d'autres auteurs moins connus.

Des chercheurs du laboratoire de Hewlett-Packard ont développé un algorithme similaire à celui de Google qui classe les sites Web en fonction de leur capacité à introduire des idées dans le World Wide Web dans son ensemble. Des résultats intéressants sont ressortis de leurs recherches :

Les auteurs d'un carnet Web empruntent régulièrement le contenu d'autres auteurs moins connus.

Lorsqu'une idée « infecte » 10 chroniques Web ou plus, 70 % des auteurs ne fournissent pas le lien menant vers la chronique où ils ont trouvé l'information.

Il est possible d'avoir une démonstration visuelle interactive des recherches du laboratoire d'Hewlett-

Packard en utilisant l'outil Blog Epidemic Analyzer.

