

La Veille multilingue : outils et ressources pour les traducteurs

Jean-Paul PINTE
Institut Catholique de Lille

STITLE STATE



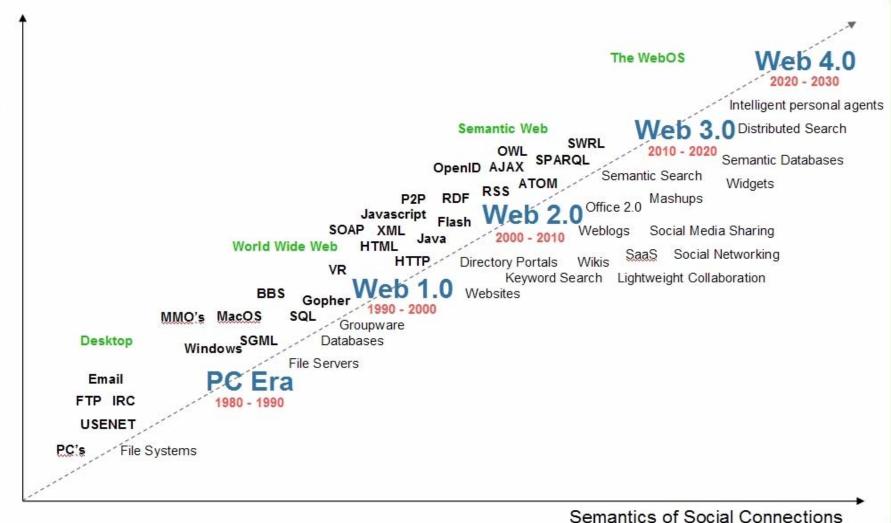
Nous verrons ensemble

- De nouveaux paradigmes
- Approche de la veille multilingue
- Un secteur toujours en chantier

Quelques avancées notables

En guise de conclusion

Quelle évolution pour le Web?









La recherche entre Sérendipité et Zemblanité

 Selon une des définitions les plus courantes, la **sérendipité** serait « *l'art de* faire des découvertes heureuses, inattendues et utiles par hasard » (Mark Raison) alors que la zemblanité désignerait « la faculté de faire de façon systématique des découvertes malheureuses, malchanceuses, attendues et n'apportant rien de nouveau [1]»

[1] Cf. Blog des automates intelligents : http://automatesintelligents.blog.lemonde.fr



En général au niveau de l'information l'internaute :

- oublie les ressources « papier », les banques de données, les ressources ouvertes
- ignore les portails spécifiques, confond les documents électroniques offerts par les bibliothèques et les ressources Internet en accès libre pour tous
- ne connait pas l'existence du Web invisible
- ne sait pas comment citer les documents numériques, ni les bibliographies
- manque de repères critiques : Google et Wikipédia deviennent ses principales encyclopédies !
- Et surtout ne pense pas ou ne sait pas chercher en dehors de sa langue!



Une approche de la veille multilingue

- Les principales tâches reliées à la recherche d'information multilingue sont le filtrage, la sélection et le classement des documents qui pourraient être pertinents pour l'usager.
- Le principal objectif de la recherche d'information multilingue est de fournir des outils à l'utilisateur qui ne serait pas familier avec une langue particulière, mais qui serait quand même intéressé à obtenir des documents dans une autre langue ou plusieurs autres langues.
- Aujourd'hui, beaucoup de travail reste à faire pour obtenir le même niveau de performance de recherche obtenu avec les systèmes de recherche monolingues.

Et pourtant ils sont nombreux sur le marché!



PulSarNotrino

-Synomia

albert :)

SYSTRAN



Les moteurs de recherche basés sur la traduction automatique

- Traduction automatique de la requête ou du corpus de document
- Moins de précision que la collection de documents qui contiennent un contexte d'information nettement plus important
- C'est donc la traduction de la requête qui est retenue lorsqu'on parle de recherche d'information multilingue
- Cette méthode est en cours de consolidation et les résultats qu'elle donne ne sont pas très satisfaisants, car elle se base sur la traduction automatique dont les résultats ne sont pas toujours probants, surtout lorsque les requêtes sont courtes. En effet, plus les requêtes sont courtes, moins les résultats seront bons.



Les moteurs de recherche basés sur des corpus d'apprentissage

- Cette méthode s'appuie, pour la traduction des requêtes, sur un thésaurus créé (grâce à un algorithme) à partir d'un corpus d'apprentissage aligné au niveau de la phrase pour trouver des cooccurrences de termes en contextes (Huthins, 2005).
- Les résultats de cette méthode sont généralement assez satisfaisants mais ont l'inconvénient d'être limités : comme pour toute méthode statistique, on constate un arrêt de la progression des résultats à un certain seuil.



Les moteurs de recherche basés sur des dictionnaires de reformulation

- Cette méthode, appelée en anglais *«machine-readable dictionaries-based method»* se base sur l'expansion de la requête. Cela consiste à formuler la requête autrement en remplaçant les mots qui la composent par des variantes afin de récupérer des documents pertinents dans lesquels les termes saisis ne sont pas toujours présents.
- Cette reformulation se fait à l'aide de dictionnaires monolingues qui permettent la reformulation dans une même langue (synonymes, antonymes, etc.) et les dictionnaires bilingues qui permettent la reformulation dans des langues différentes (Aljlayl et al., 2001).



Les moteurs utilisant une langue pivot (l'interlingua)

Cette méthode consiste à extraire la sémantique des textes de la langue source grâce à un langage pivot.

Il s'agit d'un langage unifié permettant de représenter la sémantique des différentes langues et basé sur le concept des graphes représentant les phrases.

Cette méthode est la moins consolidée de toutes celles que nous avons citées et peu de résultats satisfaisants ont été constatés (Hahn et *al.*,2004).



Ex: Un dispositif unique outillant l'ensemble du processus de veille

Stocker

- 3 modules types pouvant être clonés et paramétrés pour créer autant d'applications métier spécifiques
- Paramétrage « sur mesure »
- Rapidité de déploiement
- Maintenance simple

Collecter / créer

- On line (browser)
- Tous formats de fichier
- Pluri-médias
- Intégration bureautique
- Formulaires
- Catégorisation métier
- Workflows de validation
- Agrégation de flux externes
- Ajouts de commentaires
- Contribution par mail

Gérer

Moteur de règles :

- Workflows paramétrables
- Versioning
- suivi des modifications
- Check in / check out
- Règles documentaires

Organiser







Bourses d'échanges

Piloter

- Tableau de bord d'utilisation metrics
- Paramétrage intégral
- Gestion décentralisée
- Archivage semi-automatique

Diffuser

Diffusion pro-active et ciblée :

- Profil d'intérêt individuel
- · Sollicitation sur compétences
- Alertes générales ou ciblées
- Croisement offres / demandes
- Exports

Diffuser

Portail Web

Portail

Interface graphique personnalisée

Publier

- Droits et rôles individuels
- Accès Intranet / Extranet / Internet
- · Menus dynamiques
- Multilinguisme intégral et simultané
- API pour intégration

Consulter

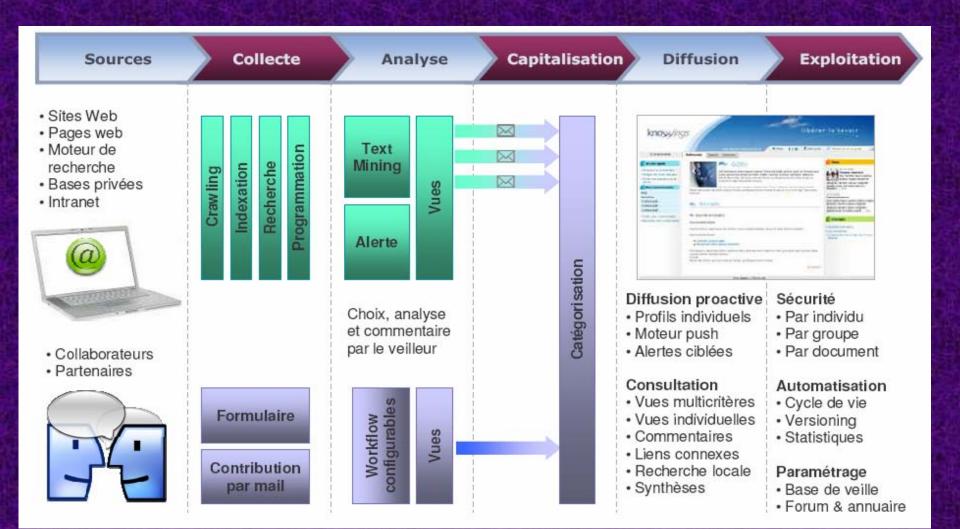
- Accès browser
- Vues multi-dimensions
 Arborescences métier
- Vues personnalisées
- Dossier personnels

Rechercher

- Plein texte, plein doc.
- Multicritères
- Concepts et synonymes métier
- Multi-bases + file system
- Multilingue



Global Finder



GlobalFinder

Avec Global Finder, la capacité à accéder et à exploiter l'information en multilingue est maîtrisée : les requêtes sont automatiquement traduites (en anglais, français, italien, allemand ou espagnol) en fonction de la langue de chacune des sources sollicitées.

De même, pour exploiter les résultats, l'utilisateur peut bénéficier d'une fonction de traduction automatique de mots et d'expressions.

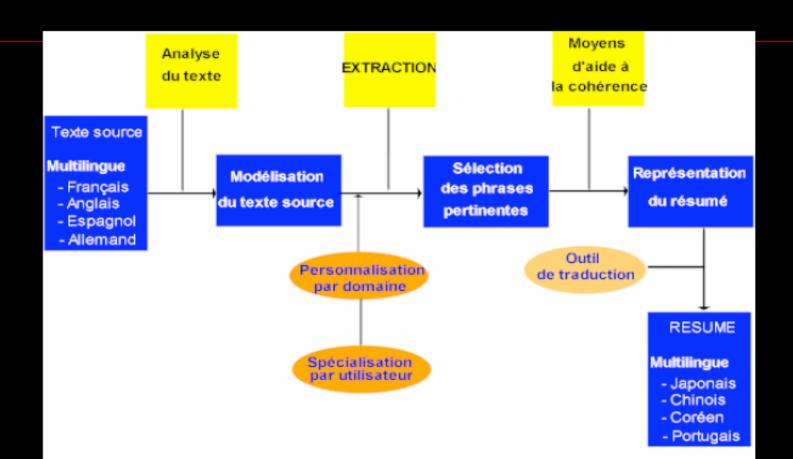
Pour s'adapter parfaitement au contexte métier de chaque entreprise, les dictionnaires de traduction automatique peuvent être enrichis de terminologies métier «sur mesure».



Technologies textuelles multilingues

Aspects multilingues

Langues traitées : Français, Allemand, Anglais, Arabe, Chinois, Coréen, Espagnol, Grec, Italien, Japonais, Néerlandais, Norvégien, Portugais, Russe





http://www.lingway.com/

Principales fonctionnalités de LINGWAY KM 3.6 :

- Acquisition de données multi-formats
- Structuration automatique de documents en XML
- Indexation et extraction de métadonnées
- Recherche sémantique multilingue
- Outils de navigation avancés (cartographie)
- Interface documentaliste
- Paramétrage et adaptation de dictionnaires métiers

Des outils qui vont vers la mobilité



World-class dictionaries, grammar checkers, and language software

English Definitions Go Search Add this to your site!

M Shopping Cart Volume Discounts English V

Mac OS X | Windows | Palm | Window

ort

Dictionaries	
English Dictionaries	
French-English	
Spanish-English	
German-English	
Portuguese-English	
Other Languages	
Online Dictionaries	
Collins Dictionaries	
Grammar Checkers	
English	
French	
Spanish	
German	

Ultralingua Online Dictionary Dictionary Conjugation Numbers References Other Resources Subscribe Login Definitions & Translations Stemmed search ✓ English definitions ✓ Go Search help Type or paste a URL to reproduce the page with dictionary-enabling. More information Spanish → English ✓ http:// Email updates on sales and special offers	About Ultralingua	
Stemmed search Search help Type or paste a URL to reproduce the page with dictionary-enabling. More information Spanish> English http://		
Search help Type or paste a URL to reproduce the page with dictionary-enabling. More information Spanish> English http://		
Type or paste a URL to reproduce the page with dictionary-enabling. More information Spanish> English w http://		
Spanish> English http://		
Email updates on sales and special offers	Submit	
Enter Email Here Submit	al	
View Ultralingua's Privacy Policy		

Babelplex

Babelplex Bilingual Search

Culture et veille informationnelle

Ce mariage déjà défendu par Jean-Paul Pinte dans sa thèse sur la Veille et la culture informationnelle en éducation (Décembre 2006) soulève aussi le ... lewebpedagogique.com/cultureinformationnelle/

Culture et veille informationnelle » Culture informationnelle

Culture informationnelle et droit à l'éducation au XXI° siècle - Culture et veille informationnelle. Jean-Paul Pinte, Auteur : Jean-Paul Pinte Niveau : ... lewebpedagogique.com/cultureinformationnelle/category/lire/culture-informationnelle/

ABRC - Culture informationnelle

Selon la Déclaration de Prague publiée à la réunion internationale d'experts en culture informationnelle en 2003, la culture informationnelle se définit ... www.carl-abrc.ca/projects/information_literacy/IL_policy_statement-f.html

Urfist Info: La culture informationnelle mondialisée

23 août 2006 ... La culture informationnelle est une notion complexe et souvent mal ... Étant donné l'importance reconnue de la culture informationnelle, ... urfistinfo.blogs.com/urfist_info/2006/08/la_culture_info.html

Information literacy - Wikipedia, the free encyclopedia

Encyclopedia article on **information literacy**. Includes definition, history, and list of references.

en.wikipedia.org/wiki/Information literacy

ACRL - Information Literacy

Information literacy resources from the Association of College & Research Libraries.

www.ala.org/ala/acrl/acrlissues/acrlinfolit/informationliteracy.cfm

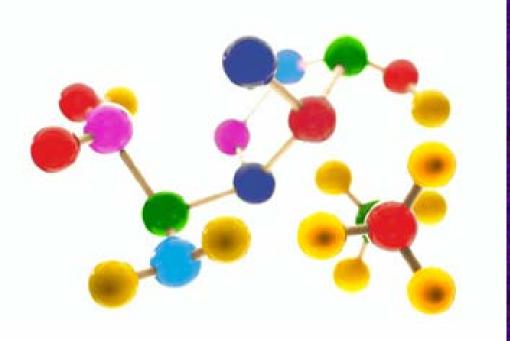
ACRL - Information Literacy Competency Standards for Higher

Describes how to a new **information literacy** program or assessing an existing program. These standards, performance indicators, and examples are approved by ...

www.ala.org/acrl/ilcomstan.cfm

The Information Literacy Website

This website has been designed and developed by information professionals from key UK organisations actively involved in the field of **information literacy**.



SINEQUA CS

http://www.sinequa.com

 Pertinence et navigation : Sinequa CS utilise la combinaison de sa technologie brevetée d'analyse sémantique et linguistique, avec l'analyse statistique, structurée et morphosyntaxique pour retrouver, organiser et naviguer à travers les éléments de connaissance les plus pertinents par rapport à une recherche.



http://www.arisem.com

PROFESSIONAL

Information Mining Solutions

Caractéristiques techniques

Contenu

- Agents intelligents pour la collecte automatisée de l'information accessible sur Internet
- Formats de documents supportés : HTML, XML, Texte, RTF, MS Office, Adobe Acrobat PDF...

Langues

- Analyse sémantique en français, anglais, allemand, italien, espagnol
- Recherche et traitement crosslingues
- Kits de contenu disponibles en français et en anglais

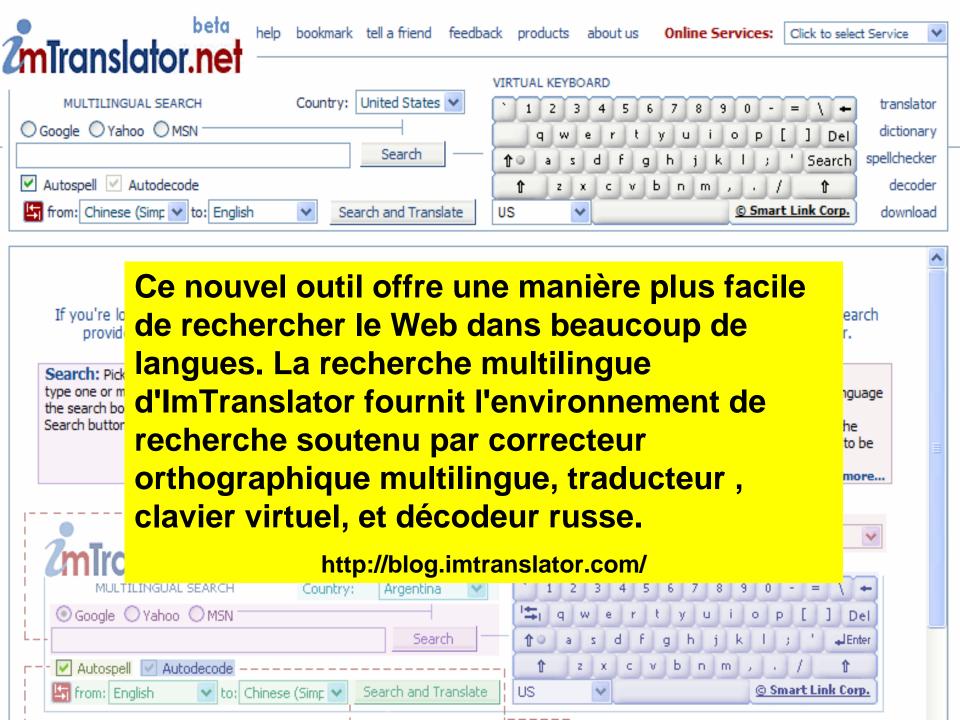
Interopérabilité

- Kaliwatch Professional peut être utilisé seul.
- Il peut également être connecté à Kaliwatch™ Server, la solution de gestion et organisation des connaissances de votre entreprise

http://www.pertimm.com



- Des requêtes expertes sans thésaurus ni langage d'interrogation complexe
- Les ontologies dynamiques (concepts)
- Facile à mettre en oeuvre et à utiliser
- Multilingue et cross-lingue
- Aucune limitation en capacité, pas de facteur d'échelle





XRCE a mis en place un **moteur de recherche multilingue** appelé **LIRIX** pour vous aider a trouver les informations qui vous intéressent sur ce site.

LIRIX permet d'améliorer considérablement la pertinence des résultats par rapport au moteur de recherche Jakarta-Lucene sur lequel il se base.

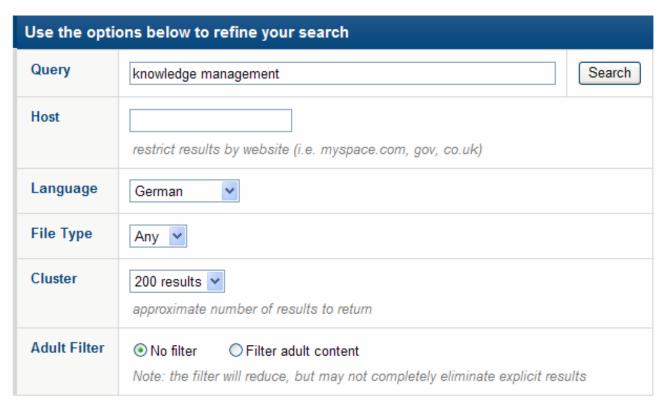
Il a également la capacité unique d'effectuer des recherches multilingues: tapez quelques mots en Français, et vous obtiendrez des documents à la fois en Français et en Anglais

Ouverture du concept multilingue vers les outils de cartographie



web news images wikipedia blogs jobs more »

advanced search



simple search

preferences

Le cas de Clusty (Ex Vivisimo)



web news images wikipedia blogs jobs more »

knowledge management

Search

advanced preferences



Font size: A A A

Clustering about 200 results in German with adult filter off (simple search)

Top 232 results of at least 103,100 retrieved for the query knowledge management (details)

Search Results

1. Knowledge Management Master, Pumacy Technologies AG :: Startseite :: 🖻 🔍 🛞

Überblick über die Basistechnologie KMmaster zum systematischen Entwickeln, Erfassen, Verteilen, Aufbewahren, Anwenden und Bewerten von Wissen.

www.kmmaster.de - [cache] - Live, Ask

2. Customer Knowledge Management 🖻 🔍 🛞

Das Forschungsprojekt des Fraunhofer Instituts zum Thema Customer Knowledge Management ... Artikel "Vom Management der Kundenbeziehungen zum Customer Knowledge Management"

www.customer-knowledge-management.de - [cache] - Live, Ask

GfWM - Gesellschaft für Wissensmanagement e.V. □ Q ⊗

Die Bedeutung dieser Ressource ist längst erkannt – das richtige **Management** und der Nutzen daraus ist jedoch nach wie vor Gegenstand intensiver Diskussionen.

www.wissensmanagement-gesellschaft.de - [cache] - Ask

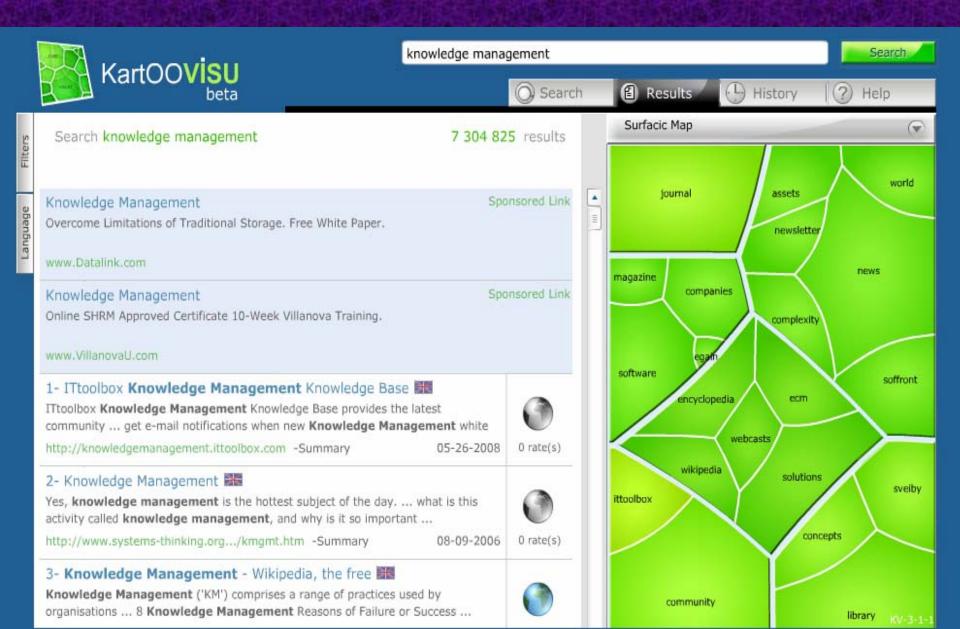
powerDOCUMENTS - structure, archive and find your knowledge! - ... □ Q ⊗

powerDOCUMENTS is a solution for small and mid-size companies to retend and manage the corporate knowledge to to the avoidance of knowledge loss.

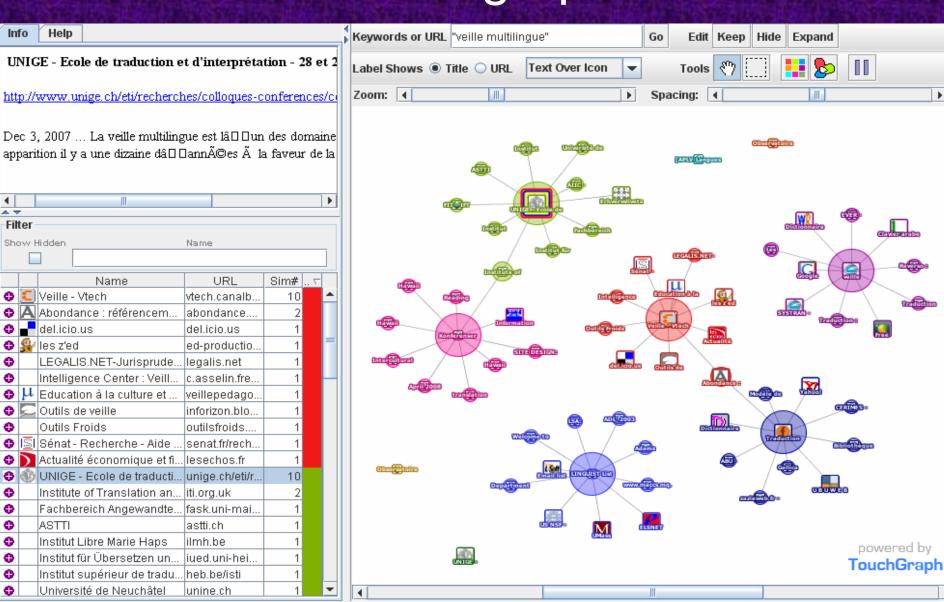
www.powerdocuments.de - [cache] - Ask

Wissensmanagement [-,mænidʒmənt] (englisch knowledge management) ist ein zusammenfassender Begriff für alle Managementpraktiken, die darauf abzielen, in Organisationen Wissen ...

Le cas de Kartoo



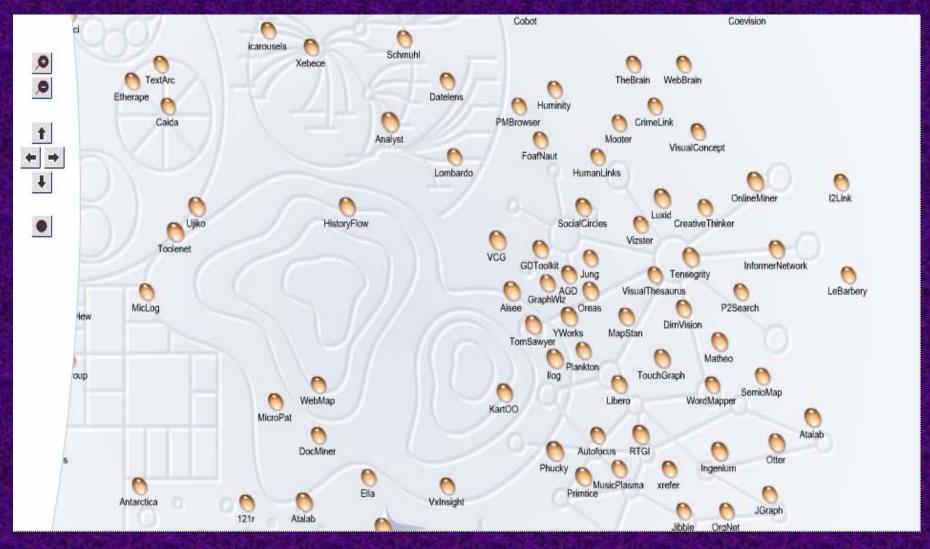
Ils y viendront rapidement ... Touchgraph

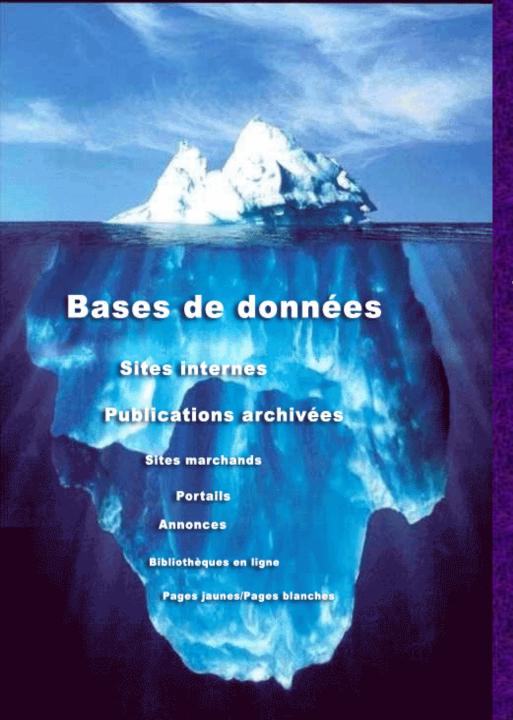


UJIKO



La cartographie des outils de cartographie avec Mapdream.com





Sans oublier l'accès au Web invisible

 il représente 95% du contenu total accessible publiquement et gratuitement en ligne

- de nombreux sites présents dans le web visible peuvent constituer des portes d'entrée sur le web abyssal.

Exemples: Clusty.com

Exalead.fr

Completeplanet

Quelques ressources pour se lancer dans la veille multilingue



machine, et vous permet de supprimer des logiciels indésirables en toute sécurité. [>>]



Recherche

Plan du site

Newsletter

Abondance.fr



Référencement Suisse

Nous pouvons placer votre site en tête des moteurs de recherche

Un nouvel emploi?

Test d'anglais immédiat et gratuit Et gagnez encore un mois offert!

Rechercher sur Abondance :							
			Chercher	Recherche avancée			
(e)	Sur tout le site		Sur l'actu d	lanuis 1999			

ABONDANCE						
Pour tout savoir sur les	s annuaires et moteurs de recherche : Recherche d'information et référencement Mercredi 28/05/2008					
Vous débutez ?	Accueil > Recherche >					
Actualité	veille multilingue Recherche dans Tout le site (17)					
Blog	veille multilingue Recherche dans Tout le site (17)					
L'actu sur votre site	Technologie Synomia					
Dossiers / Articles	17 documents trouvés pour: veille ET multilingue					
Méthodologies	· · · · · · · · · · · · · · · · · · ·					
Outils de recherche	Pour affiner votre recherche Aide					
Audits	recherche d'information >> valeur ajoutée >>					
Forums / Chat	cahier des charges >> Gestion de projet >>					
Lettres d'information						
Études						
FAQ	Trier par date / pertinence Montrer / Cacher les extraits					
Livres	 article pour les abonnés au site Pratiquez la veille collaborative en ligne avec Rivalmap () Le référencement de sites multilingues http://www.abondance.com/article-abonnes.html 					
Emploi						
Tribune						
Humour	2. Archives de la lettre "Recherche & Référencement " - Classement des articles par thèmes InfoCodex, une solution de veille entre sémantique et réseaux neuronaux (5 pages) Nous voulons associer la					
Liens						
Offre commerciale	reconnaissance de formes à l'image et à la vidéo " (5 pages). Directeur du Laboratoire d'Ingénierie de la Connaissance Multimedia Multilingue (CEA) et à l'origine notamment du moteur de recherche					
Boutique en ligne	http://abonnes.abondance.com/archives/theme-moteurs.html					
Zone Abonnés						
Réseau Abondance	3. articles de la lettre d'information Recherche et Référencement - année 2006					
	Le référencement de sites multilingues (1ère partie) (6 pages) Le second a opté pour la recherche intra-site et la veille concurrentielle ()					
Annonces Google	http://lettres.abondance.com/articles-rechref-2006.html					
1	4. Namella saurea d'information : Banart Buyan (Bacharcha d'information et veille					
Vatua alta mala	4. Nouvelle source d'information : Report Buyer (Recherche d'information et veille - Forum Abondance) Enfin, Report Buyer dispose également d'un catalogue mis à jour régulièrement proposant des fonctionnalités de recherche					
Votre site web	avancées et du support de l'équine de service clientèle multilinque acqueillante et disponible					

Des étudiants rentrent dans le flux

DOMOTIQUE NU JAZON



Contactez l'auteur

CATÉGORIES

Démarche et outils de veille

Introduction à la domotique et état du marché

La presse sur le secteur

La robotique domestique

Les acteurs de la domotique au Japon

Les industries annexes

Les salons dans le monde

Présentation: sujet, cadre et auteurs

ARCHIVES

août 2007

juin 2007

avril 2007

19 décembre 2006

Les outils de veille utilisée et la démarche

L'orientation de nos recherches

Le but de ce blog n'était pas de créer un catalogue ou un inventaire des derniers gadgets ou nouveautés nouvellement commercialisés ou en développement au Japon.

Les articles sélectionnés devaient permettre d'illustrer les multiples applications de la domotique aujourd'hui au Japon : les produits choisis ne sont qu'un prétexte qui nous permettent de montrer comment ce secteur intervient aujourd'hui dans le quotidien des Japonais, que ce soit à la maison ou dans leur environnement professionnel.

Même si de plus en plus de commerces européens proposent aujourd'hui des solutions domotiques à leurs clients, le consommateur reste à "éduquer "et "sensibiliser" sur les usages et avantages de la domotique. Une étude du marché japonais, plus avancé que le notre, peut y aider.

Notre stratégie

Notre stratégie a été de nous baser sur les recherches japonaises actuelles en domotiques. Chacun s'est assigné un rôle : qui de se

« MAI 2008

dim lun mar mer jeu ven sam 25 26 27 28 29 30 31

BLOGS AMIS DU MASTER CMAI DE LILLE 1

Japon : laboratoire du marché des séniors

La publicité en Chine

Le marché chinois des cosmétiques

Relations transatlantiques

Situation du textile français

LES AUTEURS

Hélène Jacquet

Mouloud Hammadou

PRESSE, BLOGS SUR CE SECTEUR

A voir aussi ...

 La page de JP Lardy « Tous les outils en une page » http://urfist.univ-lyon1.fr/risi/outils.htm

 Une typologie des outils de recherche sur la toile : http://www.adbs.fr/

En guise de conclusion

- Trouver l'information multilingue utile à l'aide d'automates ou d'outils de recherche reste encore un vœu pieux pour beaucoup
- Il est donc urgent de former les traducteurs aux habiletés informationnelles et leur rôle est aujourd'hui primordial à l'heure de la société de l'information dans un contexte d'«infobésité»
- Il faut sortir aussi de l'«alphanétisation» qui voudrait que l'on s'exprime et que l'on apprenne en toute autonomie avec les nouveaux outils issus de la culture informationnelle et que l'on comprenne les enjeux et les implications professionnelles et personnelles des évolutions technologiques et sociologiques de notre époque.



Merci de votre attention ...

Contact: pinte.jp@gmail.com

http://veillepedagogique.blog.lemonde.fr

http://lewebpedagogique.com/cultureinformationnelle