



HAL
open science

Les outils de la veille pédagogique

Jean-Paul Pinte

► **To cite this version:**

Jean-Paul Pinte. Les outils de la veille pédagogique. XXII ème Congrès International de Pédagogie Universitaire, May 2004, Marrakech, Maroc. edutice-00000504

HAL Id: edutice-00000504

<https://edutice.archives-ouvertes.fr/edutice-00000504>

Submitted on 3 Jul 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les outils de la veille pédagogique

Jean-Paul Pinte, Enseignant en Informatique – Multimédia – Veilles

Adresse professionnelle : Université Catholique de Lille - FLSH ★ 60 Bd Vauban ★
F-59000 Lille

<http://www.icl-lille.fr>

Mail : Pintejp@aol.com, Tél.: 06.80.60.04.35

Doctorant Ecole Doctorale ICMS Marne La Vallée

Laboratoire CESD - IFIS

<http://www.polytechnicum.org/icms/>

Intervenant – Conférencier en recherche d'informations stratégiques

Collège pour l'Enseignement Pratique de l'Intelligence Economique

<http://www.cogefi.com/formations/intervenantsIE.pdf>

Conseil en Intelligence économique pour les Chambres de Commerce et d'Industrie.

Membre de l'Association Internationale de Pédagogie Universitaire

<http://www.ulg.ac.be/aipu/>

Membre du Comité scientifique de la Revue internationale des technologies en pédagogie universitaire dirigée par Thierry Karsenty

<http://www.profetic.org:16080/revue/sommaire.php3>

Préambule :

A la rentrée 2004, un nombre encore plus important d'étudiants arrivera en cours avec comme outil complémentaire d'accès au savoir « la grande Toile ». Dans les classes de la maternelle et du primaire ils seront dans les prochaines années peu nombreux à n'avoir connu d'autre monde que celui du Web. En effet, aujourd'hui l'acquisition du réseau Internet conforté par l'accès haut débit s'avère de plus en plus rapide avec les incitations au divertissement, les recherches éducatives, les échanges avec les amis, l'écoute de la musique, que facilite cet outil formidable à l'école et à la maison.

Les Nouvelles Technologies de l'Information et de la Communication constituent de plus en plus un atout qui va aider l'apprenant à effectuer des tâches en le plaçant dans une relation nouvelle par rapport à l'enseignant.

Le potentiel de ces outils est surtout lié à leur flexibilité, à leur capacité de véhiculer l'information de façon multimodale. Elles sont liées aussi à la fascination qu'elles exercent, surtout à l'époque de l'adolescence, pendant laquelle les représentations négatives ou faussées de soi-même et de son travail apparaissent fréquemment.

L'étudiant qui a des difficultés à suivre une leçon magistrale monomodale, qui repose tout entière sur l'écoute de la langue orale, peut modifier la représentation négative qu'il a de ses capacités à partir du moment où il suit la même leçon au moyen de fiches ou d'hypertextes, sur son écran.

Mais, le Web leur permet aussi de trouver des travaux faits pour eux, sur mesure, qu'ils pourront « copier coller » en deux secondes et ensuite présenter à leur enseignant comme étant un original. Pendant que les jeunes découvrent Internet (et ils sont très rapides!), les enseignants, eux, doivent commencer à se méfier de cette montée en charge de petits futés qui se servent du réseau et trouvent l'information par simple « sérendipité »¹ dans ce monde d'infobésité² et s'intégrer peu à peu dans l'accompagnement à la recherche de l'information pertinente et à son exploitation dans le cadre des cours qu'ils soient ou non à distance.

¹ Jean-yves Prax définit la sérendipité comme l'art de trouver ce qu'on ne cherche pas. (Le Manuel du Knowledge Management, Ed. Dunod, p.76, 2003).

² Il y a infobésité lorsque la quantité d'informations disponible est si importante qu'il devient impossible pour les individus ou les organisations d'en extraire rapidement et efficacement l'information pertinente. (Anglais : Information overload). Cf. Schéma 1.

Heureusement, il existe aujourd'hui des outils qui peuvent aider les apprenants comme les enseignants à faire face à ces nouvelles tâches et ces nouveaux défis. Il s'agit des outils de veille pédagogique dont le présent article est l'objet.

I. De l'utilité d'une veille pédagogique

Les Technologies de l'Information et de la Communication (TIC) sont perçues en ce début de siècle comme un défi dans la mesure où elles sont vouées à mettre en crise et à désarticuler le système éducatif. Elles constituent aussi un véritable pari, puisqu'elles peuvent contribuer à modifier l'approche de fond des acteurs éducatifs en ce qui concerne l'autonomie des apprenants.

Alors qu'elles ont acquis une position prééminente dans la société, les TIC tardent cependant à trouver leur place réelle dans le système éducatif et, alors que les jeunes sont en pleine immersion dans ce que l'on pourrait appeler les multimédias, les instances éducatives tendent encore à préserver le seul mode de l'écriture/lecture.

Les dernières avancées en termes de mise en place de plates-formes de ressources en ligne et de formation à distance orientent inévitablement nos systèmes éducatifs vers les autos apprentissages.

Deux problèmes se heurtent aujourd'hui au développement des TIC :

- la production actuelle des ressources n'est pas conduite par des critères didactiques et pédagogiques, mais par des critères liés aux intérêts techniques et au profit précise P. Moeglin³ en observant que les produits pédagogiques ne sont pas conçus dans une perspective pédagogique et qu'ils exigent en fait de l'apprenant que lui-même crée une convergence dans un produit qui n'en a pas (références souvent à des modèles pédagogiques sous-jacents contradictoires, taylorisme et cognitivisme), bref que l'apprenant soit autonome à priori⁴.
- le second problème, constaté surtout en Europe est lié à l'attitude des enseignants envers les TIC. Les enseignants ont des difficultés à comprendre le caractère structurel des changements qu'on leur demande de gérer et d'accompagner. Le problème auquel les systèmes éducatifs doivent se mesurer est celui de la vitesse des changements *extra muros*, qui semble

³ Moeglin P. (1996) Multimédias et éducation, le démon de la convergence, in Outils multimédias et stratégies d'apprentissages du FLE, « Cahiers de la Maison de la recherche », Lille III, Université Charles de Gaulle.

⁴ Il ne s'agit pas d'une autonomie positive, dans ce cas, mais de la capacité à s'adapter à un matériel d'apprentissage.

incompatible avec la durée nécessaire que demande un changement culturel et professionnel significatif.

Il faut aussi souligner le décalage qui existe entre les savoirs académiques transmis par les enseignants et ce que les étudiants en perçoivent aujourd'hui avec l'avènement des nouvelles technologies de l'information et de la communication et d'en définir les éléments favorisant une interaction entre l'enseignant et l'étudiant. Pour cette approche, il est nécessaire de prendre en compte les conceptions de l'étudiant et convenir que celui-ci réfléchisse sur celles de l'enseignant, d'où un processus d'apprentissage conçu comme un dialogue entre ces deux parties.

II. Internet et les multimédias au cœur du processus de veille pédagogique

Il y a seulement quelques années encore, une personne en quête d'information savait qu'il lui fallait se rendre dans un lieu documentaire pour satisfaire sa demande après avoir exploité la richesse du lieu. Dans ce contexte il lui fallait alors comprendre un minimum les techniques professionnelles conçues par des experts de l'information documentation et se les approprier pour retrouver le rangement ou le classement d'une ressource, d'un article ou d'un livre, ce qui n'était pas toujours une chose facile à l'époque.

Les années 80 ont fondé de grands espoirs sur la recherche documentaire informatisée avec l'arrivée des ordinateurs et la possibilité de formuler plus ouvertement ses recherches et en diversifiant les accès au signalement des ressources, ce qui a sensiblement modifié la relation de l'utilisateur au document. Depuis 1993, le réseau des réseaux a entraîné le concept de cyberspace qualifié de nouvel emblème de l'encyclopédisme universel par Morizo⁵. Son introduction dans les centres de ressources documentaires et à la maison a rapidement mis en évidence les raccourcis entre accès à l'information et construction de savoirs.

Quelques chiffres peuvent à cet effet nous donner aujourd'hui le vertige :

- Le cabinet ETForecasts estime que le nombre d'internautes dépassera le milliard d'ici 2005
- Selon International Data Corp, plus de 30 milliards d'e-mails seront échangés par jour fin 2005

⁵ Morizo C., La recherche d'information, Ed. ADBS Nathan Université, 2002.

- Online Computer Library Center Ind. (OCLC⁶) évalue à 9, 04 millions le nombre de sites uniques sur le Web.
- La société Cyveillance⁷ estime que le Web visible contient aujourd'hui plus de 8 milliards de pages.
- Selon la société Brightplanet⁸ le Web invisible⁹ contiendrait 550 milliards de documents à l'heure où cet article est rédigé...

On comprend aisément que devant un tel gisement informationnel, tout acteur du système éducatif soit si souvent insatisfait du résultat de ses recherches sur le Net. Aussi, en entrant dans ce nouveau millénaire le milieu éducatif s'interroge aujourd'hui plus particulièrement sur les relations entre la recherche documentaire, l'utilisation des nouvelles technologies et les apprentissages. Les différents modes de structuration des connaissances dans les documents numériques deviennent aussi des centres d'intérêt pour les chercheurs en psychologie cognitive. Internet et les multimédias¹⁰ doivent donc être perçus aujourd'hui par les enseignants comme par les étudiants comme des outils leur permettant de soutenir les préparations de cours, de trouver des outils pédagogiques concrets, d'effectuer des recherches sur une problématique ou un thème donné et ainsi, d'actualiser l'enseignement en fournissant des références additionnelles aux cours dispensés en face à face.

La recherche d'information ne doit pas être uniquement une affaire de machine, mais une affaire d' « homme », produisant du savoir, le mettant en scène, le conservant, le communiquant, et se l'appropriant pour être et pour agir.

Depuis la mise en place d'outils d'apprentissage en ligne et de plates-formes de ressources pédagogiques en ligne on voit poindre ici et là l'évolution de plusieurs paramètres :

- Les centres de ressources permettent de développer le travail coopératif et garantissent à l'apprenant non plus de se fondre dans le groupe, mais de se construire grâce à des relations avec autrui et de développer sa personnalité globale et sa relation à l'altérité.

⁶ <http://Wcp.oclc.org>

⁷ <http://www.cyveillance.Com>

⁸ <http://www.brightplanet.com>

⁹ Le « Web visible » est l'ensemble des pages Web qui peuvent être indexées par les moteurs de recherche, par opposition aux pages du « Web invisible » que les robots des moteurs ne peuvent atteindre.

¹⁰ Les multimédias sont composés des outils de communication comme les Cd Rom, la presse, la radio, la télévision, etc.

- L'apprenant acquiert des compétences d'apprentissage transférables qui lui assurent la possibilité de travailler en autonomie le moment venu.
- L'intégration des multimédias offre, de plus, une gamme d'exploitations individuelles variées où l'enseignant se doit de préparer des tâches ou scénarios pédagogiques.
- L'apprenant au travers de l'intégration des multimédias dans l'apprentissage à la liberté de choisir ses parcours, d'effectuer des découvertes. L'enseignant est là pour écouter, encourager et permettre aux apprenants de confronter leurs observations et leurs découvertes.

La transition vers l'accès aux nouveaux dispositifs cités ci-dessus doit se préparer par des étapes où l'apprenant apprend à utiliser les multimédias avec le guidage souhaité de l'enseignant.

« Peut-être que nous devrions utiliser les technologies pour fournir une base de données riche en information et démonstrations. Fournir un laboratoire d'apprentissage dans lequel les étudiants font de l'exploration et résolvent des problèmes qui sont proposés par les enseignants.

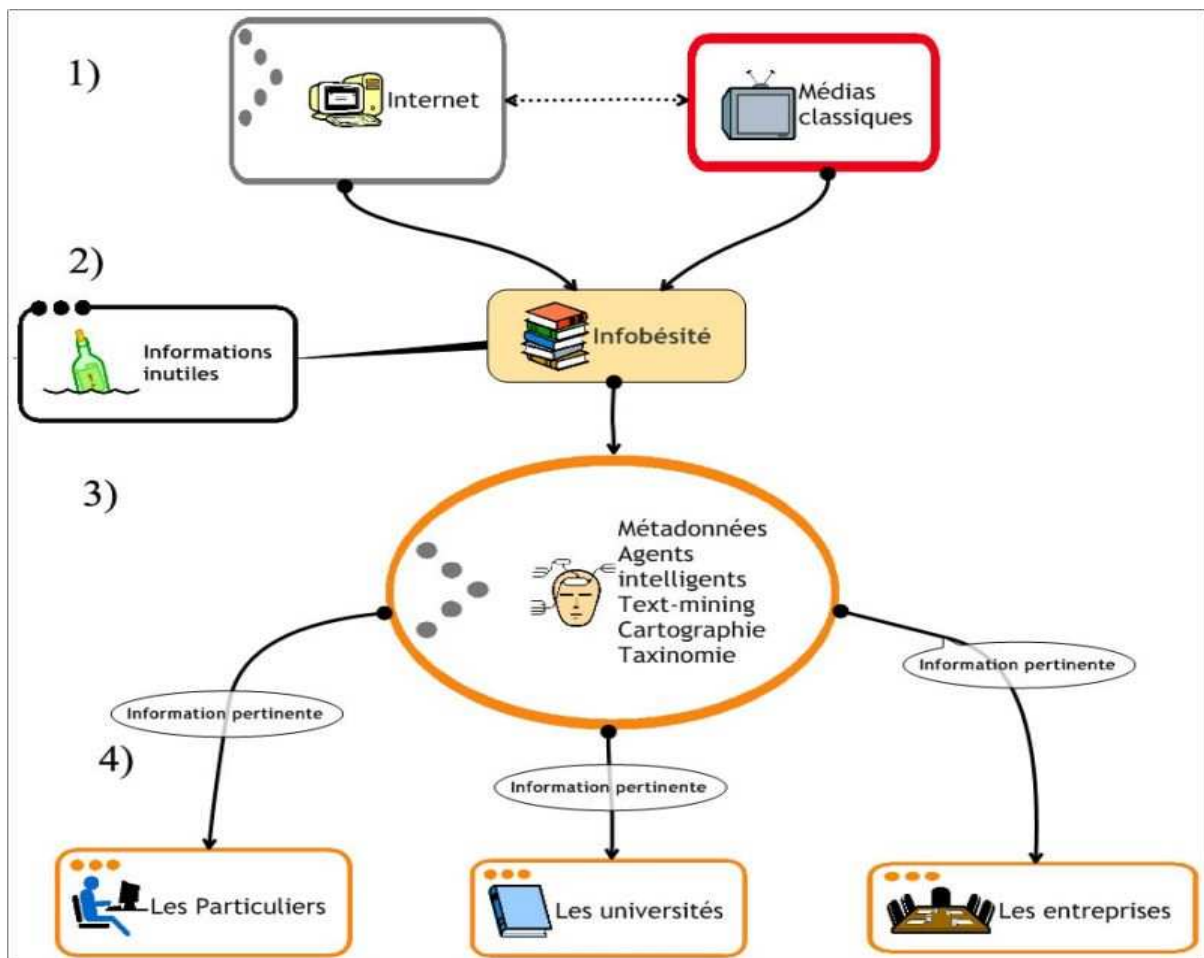
De cette façon, les professeurs deviennent des assistants dans la découverte des connaissances, des guides dans l'exploration, la structuration de la compréhension de l'étudiant.¹¹ »

La mission de la Veille pédagogique vise dans ce cadre à promouvoir et soutenir l'intégration des TIC (Technologies de l'Information et de la Communication) dans l'enseignement et cette mission peut se traduire par trois axes fondamentaux:

- Informer les intervenants dans l'éducation sur les ressources didactiques disponibles et les pratiques pédagogiques innovantes en matière de TIC.
- Appuyer les acteurs de l'enseignement dans leur démarche d'appropriation pédagogique des TIC par le développement d'habiletés telles la recherche efficace de ressources disciplinaires voire encore la création de scénarios pédagogiques intégrant les TIC.
- Traquer, traiter, partager et capitaliser l'information pour une meilleure gestion des connaissances.

¹¹ Norman D.A. (1991), Cognitive Artifacts in Designing interaction: Psychology at the Human-Computer Interface, New York, Cambridge University Press.

Schéma 1 : L'homme face à l'infobésité



1- Les principales sources d'information viennent de l'ensemble des médias c'est-à-dire de la télé, de la radio, des journaux etc., et bien sûr d'Internet. Ces deux grandes sources d'information se complètent souvent entre elles aujourd'hui par le jeu de l'hypertexte et donc du renvoi vers une ressource sur la toile.

2- Face à cette massification de l'information, l'homme arrive vite à une surcharge informationnelle, qualifiée aujourd'hui d'infobésité. De cette masse d'informations il lui faut alors en sélectionner et en extraire l'essentiel.

3- Il existe différentes méthodes et moyens pour sérier l'information, comme l'utilisation de méta données (données sur des données qui décrivent le contenu et la fonctionnalité de sources), les agents intelligents pour automatiser et filtrer l'information, le text-mining ou l'analyse de textes pour en retirer l'essentiel, la cartographie qui permet une représentation claire et synthétique de l'information sous forme de carte, enfin la taxinomie pour classer et répertorier l'information afin de se rendre directement à l'essentiel.

4- Ces méthodes vont permettre d'extraire de la masse d'informations celles qui sont pertinentes pour n'importe quel quidam qui recherche sur Internet ; pour l'université dans le cadre d'une gestion de la connaissance ; pour l'entreprise, afin de réaliser des gains de temps dans la recherche d'une information stratégique dont elle a besoin (sur un produit, sur un concurrent, sur un marché, etc.).

A ces effets, tous les acteurs de l'éducation de l'apprenant à l'enseignant en passant par l'organisation se doivent de se constituer une boîte à outils ou mallette de veille pédagogique qui va des compétences relatives à l'analyse du contexte cyberculturel, au décodage de la communication et à la structuration de la relation pédagogique avec les NTIC.

III Les outils de la veille pédagogique

Le filtrage d'information en ligne plus communément appelée de nos jours recherche d'informations en ligne grâce à des moteurs de recherche et à des logiciels documentaires n'est pas nouveau et remonte à une quarantaine d'années et, comme le signale Perriault¹² l'histoire de cette technique est intéressante car, après avoir tendu vers le « tout automatique », on est en train de réinsérer l'intervention humaine dans son dispositif.

Sans rentrer dans l'historique complet de la recherche automatisée de documents, il convient néanmoins de signaler les travaux qui ont suivis ceux de Paul Otlet en 1904 avec sa classification décimale universelle. Dès 1960 les travaux de Grolier et notamment de Gardin avec son équipe du CNRS réalisent Syntol¹³ (Syntagmatic Organisational Language) qui permet de construire des requêtes (questions posées à un système documentaire) très fouillées pour consulter une base de données contenant des données indexées permettant d'effectuer des recherches à l'aide de mots clés, que l'on utilise encore aujourd'hui encore dans les moteurs de recherche et de relations sémantiques qui indiquaient le rapport de sens entre les descripteurs. (Ex : on pouvait distinguer une question concernant le chien de mon père de celle qui concernait le père de mon chien pour reprendre l'exemple de Perriault).

¹² Perriault J., L'accès au savoir en ligne, Ed. Odile Jacob, 2002, p.64.

¹³ R. C. Cros, J.-C Gardin et F.Lévy, l'Automatisation des recherches documentaires. Un modèle général. Le Syntol, Paris, Gauthier-Villars, 1964.

Notons que les moteurs actuels n'utilisent plus cette fonctionnalité et se contentent de Et, de OU et de SAUF que l'on peut interposer entre des termes ou des groupes de termes, encadrés par des parenthèses.

A tout ceci s'ajoute la volonté des acteurs de vouloir accéder à des informations de plus en plus pertinentes en caressant de près l'idée de les obtenir si possible directement dans leur courriel sans en faire la demande ...

A la vue de toutes ces attentes, de l'évolution d'Internet et du déluge informationnel engendré, cette analyse a pour objectif maintenant de donner les bases nécessaires et suffisantes à tous les acteurs de l'éducation pour trouver de manière efficace l'information sur le réseau. Une première partie abordera une typologie des informations présentes sur le Web tout en posant la faisabilité de leur collecte, une seconde sera consacrée aux grandes familles d'outils manuels et automatisés de recherche. Pour conclure, Jean-Paul Pinte propose de consulter en ligne sur Internet une liste synthétique de liens et de ressources utiles voire indispensables sélectionnés à votre intention, pour vous aider à mieux appréhender le concept de veille pédagogique.

III-1 De l'information blanche à l'information noire.

La définition d'une typologie de l'information est un préalable à toute activité de collecte et de traitement dans le processus de veille. Celle présentée dans le schéma 2 ci-dessous reprend les travaux de divers auteurs (Martre, 1994), (Dou, 1995), (Achard et Bernat, 1998), (Bloch, 1999).

Schéma 2 : Typologie de l'information

(Le code couleur est adopté par l'Afnor) (Hermes, 2000)

<i>INFORMATION</i>	<i>BLANCHE</i>	<i>GRISE</i>	<i>NOIRE</i>
Type	Scientifique, technologique, commerciale, juridique, financière, stratégique, personnelle		
Niveau	Tactique, opérationnel, stratégique		
Domaine opératoire	Documentaire, de situation, d'alerte		
Intérêt	Fatale, utile, pertinente	Pertinente, critique	Critique
Accès	Public	Restreint	Strictement limité
Classification	Non Protégé	Diffusion restreinte	Confidentiel - Secret
Disponibilité	80%	15%	5%
Acquisition-Exploitation	Légale sous réserve de respecter les droits de propriété	Domaine juridique non clairement défini. Risques d'ordre jurisprudentiel	Illégal. L'acquisition relève de l'espionnage. Risques très élevés.
Forme	Formelle (texte) ou informelle (conversation, rumeur)		
Sources	Ouvertes	Autorisées - Fermées	Clandestines
Coût	Faible	Faible	Elevé
Rentabilité	Elevée	Très élevée	Faible

L'activité de veille pédagogique repose sur le concept d'autonomie Informationnelle. Ceci suppose de définir par avance la base à partir de laquelle sera développé le processus, autrement dit quel type d'information entre dans cette problématique ? A quel niveau décisionnel peut elle être utilisée ? Quel est son domaine opératoire ? En quelque sorte quel niveau d'attention temporelle requiert elle ? Quel est son intérêt, son accès, sa forme, ses sources, etc. ?

Dans ce tableau, seules l'information blanche et l'information grise concernent la veille, le troisième, celui de l'information secrète relève de l'activité des services spéciaux.

Le principe essentiel de la veille menée par un étudiant ou un enseignant est fondé sur le recueil de l'information blanche, l'information grise marquant une zone de non droit séparant d'une manière relativement floue les territoires autorisés des territoires interdits.

Globalement nous pouvons définir les sources d'informations, du point de vue de la structuration, puis du point de vue des supports comme présenté dans le schéma 3 ci-dessous.

Schéma 3 : Sources et supports d'information

<i>Information Support</i>	<i>Structurée</i>	<i>Non structurée</i>
Documentaire	<i>Rapports d'étonnement Enquêtes Formulaires Questionnaires Annuaire téléphoniques Tarifs Normes</i>	<i>Presse (journaux, revues) Ouvrages, livres Rapports d'ambassades (PEE) Publications des cabinets spécialisés, des organismes consulaires Documents légaux (bilans, rapports annuels) Courrier Journaux d'entreprise, tracts Rapports de stages, thèses Plaquettes commerciales Manuels d'utilisation Revues et documents en ligne</i>
Electronique	<i>Bases de données Brevets</i>	<i>Sites Internet Forums de discussion Listes de diffusion Emails</i>
Multimédia		<i>Enregistrements sonores Films, documentaires, reportages Photographies</i>
Relationnel		<i>Expertise interne Echanges clients/fournisseurs Colloques, séminaires, foires Réseaux et chambres consulaires, clubs d'entreprises</i>
Informel		<i>Conversations Indiscrétions Rumeurs</i>

Dans sa thèse de Doctorat Franck Bulinge¹⁴ définit l'information structurée comme une information textuelle dont le traitement peut être assuré automatiquement et en totalité par des outils informatiques (Exemple des bases de données relationnelles). L'analyse du schéma 3 nous permet de constater que l'information non structurée est beaucoup plus abondante et diversifiée que l'information structurée.

La problématique de la veille est donc complexe : il s'agit de collecter et d'exploiter de l'information ouverte dont la diversité est telle qu'il est impossible d'envisager un système de traitement homogène.

A cet effet, il convient avant tout de se lancer dans l'utilisation d'outils de veille automatisés, de bien connaître les familles d'outils auxquels l'acteur de l'éducation a accès pour aboutir dans sa demande d'information.

III-2 Les familles d'outils ¹⁵

- Les répertoires ou les annuaires

Ils sont parmi les premiers outils de recherche à avoir fait leur apparition sur le Web, le plus connu étant Yahoo créé en 1994 par David Filo et Jerry Yang, deux étudiants de l'Université de Stanford aux Etats-Unis.

¹⁴ Bulinge. F. *Thèse de Doctorat* « Pour une culture de l'information dans les petites et moyennes organisations : Un modèle incrémental d'Intelligence économique ». Décembre 2002

Le but des annuaires est de répertorier les sites Web et de les classer dans des catégories thématiques pour faciliter leur identification par l'internaute. A partir d'un thème, une série de sous thèmes est proposée à l'utilisateur pour aboutir enfin à une liste de sites répondant à sa question.

On peut distinguer trois types d'annuaires que sont les annuaires généralistes, les annuaires sélectifs, les annuaires et portails thématiques.

Les annuaires généralistes sont les plus connus et recense les sites grand public et professionnels, dans tous les domaines.

Les annuaires sélectifs cherchent à avoir la couverture la plus complète possible du Web. Plus anciens que les répertoires généralistes ces annuaires ont été réalisés par des professionnels de l'information, bibliothécaires ou documentalistes, qui sélectionnent uniquement les sites les plus à même de répondre aux diverses questions des chercheurs, enseignants, étudiants et chercheurs. Ces annuaires couvrent le plus souvent toutes les matières enseignées dans les universités, et ne référencent dans chaque catégorie que les sites les plus riches¹⁶.

Les annuaires et portails thématiques sont construits sur le même principe que les annuaires généralistes mais ont une approche thématique : leur objectif est de recenser les ressources dans un domaine précis. Les mêmes raisons qui ont conduit de nombreux annuaires généralistes à se transformer en portails ont fait évoluer nombre d'annuaires généralistes à se transformer en portail ont fait évoluer nombre d'annuaires thématiques vers le portail thématique appelée parfois « vortail¹⁷ ».

- Les moteurs de recherche

Apparus sur le Web dans les années quatre-vingt-dix, la vocation des moteurs de recherche est d'indexer le plus grand nombre de pages Web. Ces moteurs sont accessibles via des requêtes par mots-clé sur le texte intégral des pages.

Trois éléments composent les moteurs de recherche : un robot, un index, un serveur Web.

Le robot également *crawler* ou *spider* parcourt le Web de façon automatique, de liens en liens. Le robot de Google s'appelle par exemple *GoogleBot*. En partant d'un certain nombre d'adresses définies au préalable, le robot va tester tous les liens hypertextes qu'il rencontre et va rapatrier le contenu des pages dans sa base. Il va

¹⁵ Pour veiller à l'évolution des outils de recherche sur le WEB et découvrir leur fonctionnement, il convient de découvrir le site référence en ce domaine : <http://outils.abondance.com/>

¹⁶ Une description des principaux annuaires sélectifs est donnée dans la rubrique du même nom sur le site de la BnF (www.bnf.fr/pages/liens/)

¹⁷ Francisation du terme anglais vortal, employé pour vertical portal.

également visiter toutes les pages soumises par les éditeurs de sites, dans le cadre du référencement. Ce parcours de liens en liens à travers le Web lui demande en moyenne quatre semaines, ce délai pouvant varier de une à six semaines.

Une fois le « tour du WEB » achevé, le crawler repart de sa base pour la mettre à jour et pour identifier de nouvelles pages.

L'index contient tous les mots de toutes les pages rapatriées par le robot. Le plus souvent, le texte intégral de la page ainsi que ses différentes balises Meta¹⁸ (titre, mots-clés...) sont indexés, mais il peut y avoir des variantes selon les moteurs.

Le serveur Web offre l'interface de recherche à l'utilisateur. Celle-ci lui permet de lancer une requête par mots sur l'index du moteur, avec des possibilités plus ou moins sophistiquées.

- Les métag moteurs

C'est à partir de ces outils que l'on peut évoquer le début d'une veille digne de ce nom. Comme aucun outil de recherche n'assure une couverture complète du réseau Internet, il est indispensable d'en interroger plusieurs si l'on souhaite avoir un panorama de ce qui existe sur le net sur un sujet, ou tout simplement pour augmenter ses chances d'identifier des pages pertinentes.

C'est principalement le rôle des méta moteurs désignés aussi sous le nom de « *méta moteurs on-line* » que d'interroger plusieurs outils, successivement ou simultanément selon les cas.

Certains méta moteurs sont des logiciels qu'il faut télécharger préalablement et qui se connectent à l'Internet le temps de la recherche. Ils entrent dans la catégorie des « agents pour la recherche », mais fonctionnent en fait comme du méta moteur *off line*.

Parmi les nombreux méta moteurs existants l'Open Directory en recense plus de 185 dans sa catégorie « Metasearch Tools », les plus sophistiqués enregistrent la requête de l'Internaute, l'envoient simultanément à différents moteurs et annuaires (que l'on peut choisir dans une liste), puis rapatrient les réponses, dédoublonnent les résultats, les classent (par pertinence, thème..) et offrent parfois la possibilité de vérifier la validité des liens.

- Les agents intelligents

¹⁸ Une balise Meta est une information insérée par le Webmaster qui ne s'affiche pas à l'écran et qui donne des indications sur la page (titre, mots-clés).

Ils sont classés dans la catégorie des « agents pour la recherche » et sont parfois aussi appelés méta moteurs off line.

Parmi ces agents intelligents on distingue deux catégories dont la première est celle des agents de recherche d'informations qui sont des logiciels intermédiaires entre les moteurs de recherche et les agents intelligents. Même s'ils ne remplissent pas toutes les caractéristiques des agents intelligents, ils sont pour la plupart, beaucoup plus évolués que de simples moteurs de recherche. D'ailleurs de nombreux efforts sont entrepris par les éditeurs de logiciels pour se rapprocher des agents intelligents.

Un agent intelligent doit pouvoir prendre des initiatives et agir sans intervention de l'utilisateur final. Dans le contexte du Web il doit pouvoir agir alors que l'utilisateur est déconnecté. Les éditeurs de logiciels contournent le problème de l'autonomie en programmant leurs logiciels afin qu'ils puissent, à intervalles réguliers, se connecter automatiquement au Web pour y effectuer les tâches confiées par les utilisateurs. Dans ce cas là, on ne peut véritablement parler d'autonomie car l'agent n'est pas actif en permanence et ne fait que reproduire comme un automate la requête programmée par l'utilisateur.

Les agents de recherche d'informations sont d'une grande diversité et ils remplissent rarement tous les mêmes tâches. Cependant, on peut identifier cinq fonctionnalités principales :

- La recherche ponctuelle d'information

Il s'agit de pouvoir effectuer des recherches sur un thème précis. Cette recherche peut être plus ou moins approfondie et complète.

- La mise en place d'une veille thématique

Cette fonctionnalité permet à l'utilisateur de suivre en permanence l'évolution de l'information sur des thèmes qu'il aura lui-même définis. Le logiciel se charge de recherche de manière automatique et périodique l'information souhaitée.

- La gestion de l'information récoltée

Il s'agit d'éditer, d'archiver, de détruire et de mettre à jour les résultats des recherches.

- L'analyse des documents rapatriés Cette analyse va par exemple permettre de mettre en relation les différents documents de manière à faire ressortir de la masse des informations pertinentes. Il s'agit également de créer des résumés automatiques.

- La navigation off line

Cette fonctionnalité permet de naviguer parmi les informations trouvées sur Internet par le système sans nécessairement être connecté.

Il est important de noter que ces fonctionnalités peuvent être présentes dans de nombreux logiciels sans qu'elles mettent en jeu pour autant les mêmes technologies.

Pour la seconde catégorie, celle des agents intelligents, la recherche ne se fait pas depuis un site Web. Ces agents sont en effet des logiciels qui doivent être téléchargés et installés sur le disque dur de l'utilisateur dans sa version définitive ou dans sa version d'évaluation.

Dans une étude de Gilles Balmisse sur les agents en septembre 2002, Jean Ferbe en donne la définition suivante :

« On appelle agent intelligent une entité réelle ou abstraite qui est capable d'agir sur elle-même et sur son environnement, qui dispose d'une représentation partielle de cet environnement, qui, dans un univers multi agent, peut communiquer avec d'autres agents et dont le comportement est la conséquence de ses observations, de sa connaissance et des interactions avec les autres agents. »

Si le principe de fonctionnement des agents est le même que celui des méta moteurs sur le Web, leurs possibilités plus étendues et en constante évolution en font toutefois des outils de recherche plus performants.

En dehors du téléchargement des résultats qui pourront être enregistrés pour un examen ultérieur, les agents intelligents proposent de plus en plus d'éditer un rapport de recherche reprenant la liste des résultats, sous la forme d'une page HTML.

Enfin, et surtout en ce qui nous concerne, ces agents intelligents se transforment de plus en plus en véritable automate de veille que l'on peut paramétrer pour se déconnecter une fois le travail terminé, envoyer un rapport de recherche signalant les nouvelles pages, télécharger des documents sur le disque, éliminer les liens non valides, ... etc.

IV Conclusion : Vers une écologie informationnelle

L'espace informationnel est devenue complexe et le devient chaque jour davantage. A la diversité des documents dans leur nature, leur forme, leur intention de communication, s'ajoute aujourd'hui une diversité des accès technologiques, des circuits de production et de diffusion. Les nouvelles mises en scène de l'information, induites par des changements de supports, impliquent cependant de nouveaux modes de lecture et de prise d'information dont l'impact fait encore l'objet de recherches. La recherche d'information pour l'apprenant comme pour l'enseignant en vue d'atteindre une connaissance reste à ce jour une aventure où chacun parte à la découverte de l'inconnu, et il faut donc que le voyage soit tentant.

L'accès à l'information et à des documents implique encore plus qu'hier une écologie informationnelle nécessitant une compréhension de leurs modes d'élaboration et de diffusion.

La diversité des outils décrits dans cet article montre que les techniques électroniques de traitement de l'information et de communication jouent les rôles les plus divers dans la formation à l'activité intellectuelle. Dans ce contexte, l'enseignant s'il veut rester au cœur de l'organisation et de la diffusion des savoirs doit s'efforcer de montrer à ses étudiants qu'un même message peut être véhiculé par diverses technologies dont il se doit de maîtriser les composantes pour caractériser l'espace de ressources dans lequel se situe la classe, le groupe ou encore le département d'université.

La base proposée en ligne¹⁹ par Jean-Paul Pinte comporte une sélection de liens et de ressources utiles retenus par l'auteur depuis 2002. Elle complète la plate-forme de partage de connaissances et de veille Commun@utice²⁰ » mise en place avec près de 600 étudiants.

A l'aide de cette mallette pédagogique, chaque acteur de l'éducation fait ses premiers pas dans le monde de la veille et apprend à apprendre avec les nouvelles technologies de l'information et de la communication.

¹⁹ <http://www.ifrance.com/laveillepedagogique>

²⁰ <http://communautice.icl-lille.fr/annu>

Références bibliographiques :

Barbot M. J., Camatarri G., (1999), Autonomie et apprentissage, l'innovation dans la formation, Education et Formation, Ed. PUF, 242 p.

Guir R., (2002), Pratiquer les TICE, former les enseignants et les formateurs à de nouveaux usages, Ed. De Boeck , 300 p.

Authier M., (1998), Pays de connaissances, Ed. Du Rocher, 249 p.

Prax. J-Y., (2003), Le manuel du Knowledge Management, une approche de 2^{ème} génération, Ed. Dunod, 473 p.

Morizio C., La recherche d'information, Ed. ADBS Nathan Université, 2002, 126 p.

Perriault J., L'accès au savoir en ligne, Ed. Odile Jacob, 2002, 266 p.

Perriault.J., Education et nouvelles technologies, théorie et pratiques, Nathan Université, 2002, 125 p.

B.Foenix-Riou, Recherche et veille sur le Web visible et invisible, Paris, Ed. Tech et Doc, 2001.

B.Foenix-Riou, Guide de recherche sur Internet, outils et méthodes, ADBS, Nathan Université, 2002, 127 p.

Carlo Revelli, Intelligence stratégique sur Internet : Comment développer des activités de veille et d'intelligence économique sur le web ?, Ed. Dunod 2000, 220 p.