

DESS Réseaux d'information
et document électronique

Mise en place d'une méthodologie de veille à partir de la collection-test Sibel

Irénée Jacques LUMANISA NGOY-BILA

Sous la direction de Omar LAROUK

Maître de conférence à l'Enssib

Sommaire

RÉSUMÉ	4
ABSTRACT	5
REMERCIEMENTS	6
INTRODUCTION	7
PARTIE 1 : PRÉSENTATION DU DÉPARTEMENT DES RESSOURCES DOCUMENTAIRES	11
1. LES MISSIONS DU DÉPARTEMENT DES RESSOURCES DOCUMENTAIRES	11
2. DÉFINITION DU PROBLÈME : OBJET DU STAGE.....	13
PARTIE 2 : PRÉSENTATION DE SIBEL	14
1. APERÇU HISTORIQUE DE SIBEL	14
2. MODE D'EMPLOI DE SIBEL.....	15
2.1. <i>Le mode administration</i>	15
2.2. <i>Le mode utilisateur</i>	17
PARTIE 3 : MÉTHODOLOGIE DE LA VEILLE	19
1. LA VEILLE	19
2. DÉFINITION DES CONCEPTS RELATIFS À LA VEILLE	20
2.1. <i>Veille environnementale</i>	20
2.2. <i>Veille technologique</i>	20
2.3. <i>Veille concurrentielle</i>	21
2.4. <i>La veille commerciale</i>	21
3. MÉTHODOLOGIE DE VEILLE	22
4. LES OUTILS DE VEILLE.....	24
4.1. <i>Typologie d'outils de veille</i>	24
4.2. <i>Les outils de veille on line : technologie pull</i>	24
4.3. <i>Les outils de veille : technologie push</i>	25
5. QUELQUES FONCTIONNALITÉS DES OUTILS DE VEILLE.....	26

5.1.	<i>Fonction d'alerte</i>	26
5.2.	<i>Fonction d'analyse</i>	26
5.3.	<i>Fonction filtrage</i>	27
5.4.	<i>Conservation des requêtes</i>	27
5.5.	<i>Surveillance des modifications de contenus</i>	28
5.6.	<i>Ciblage des sources interrogées</i>	28

PARTIE 4 : OPTIMISATION DE LA GESTION DE L'ANNUAIRE SIBEL.29

1.	DES OUTILS DE VÉRIFICATION DES LIENS HYPERTEXTES.....	29
1.1.	<i>Les logiciels de vérification des liens</i>	30
1.1.1	Choix des logiciels.....	30
1.1.2	Check Web Links.....	31
1.1.3	Xenu's Link Sleuth.....	32
1.1.4	Résultat de vérification des liens.....	34
1.1.5	Check Web Links.....	35
1.1.6	Xenu's Link Sleuth.....	36
1.1.7	Tableau des codes erreurs Internet enregistrés dans le site Sibel.....	37
2.	DES OUTILS DE VEILLE.....	39
2.1.	<i>Choix des requêtes</i>	39
2.2.	<i>Quelques notions de pertinence de résultats par rapport à Sibel</i>	40
2.3.	<i>Grille d'évaluation des résultats d'interrogation des outils de veille</i>	41
2.4.	<i>Résultat de recherche des outils de veille</i>	43
2.5.	<i>Les outils de veille sélectionnés</i>	46
2.5.1	Copernic.....	46
2.5.2	BullsEye Pro.....	47
2.6.	<i>Etude comparative des outils de veille sélectionnés</i>	48

PARTIE 5 : BILAN.....50

1.	FRUSTRATIONS.....	51
2.	RÉUSSITES.....	51
2.1.	<i>Ce que l'organisme d'accueil m'a apporté</i>	51
2.2.	<i>Ce que j'ai apporté à l'organisme d'accueil</i>	52

CONCLUSION.....	53
BIBLIOGRAPHIE	55
TABLE DES ANNEXES	58

Toute reproduction sans accord express de l'auteur à des fins autres que strictement personnelles est prohibée.

Résumé

La mise en place d'une méthodologie de veille en vue d'optimiser la gestion de l'annuaire Sibel implique beaucoup de tâches : la maintenance du site, la collecte, le traitement et la diffusion de l'information. Les logiciels de vérification de la validité des liens sont testés pour la maintenance du site. Et, les outils de veille sont proposés aussi bien pour alimenter Sibel par des nouvelles pages web que pour signaler les changements.

Mots-clés : Sibel, veille, méthodologie de veille, vérification des liens, outil de veille, check web links, xenu's link sleuth, bullseye, copernic

Abstract

Developing on methodology of surveying permitting to optimize the management of Sibel repertory encompasses several actions : maintaining the site , collecting, processing and disseminating information.

The softwares specialized in checking the validity of the links are tested for maintaining the site. Some survey tools are proposed to provide Sibel with new web pages and warn changes

Keywords : Sibel, survey, survey methodology, checking the links, surveil tool
Check web links, xenu's link sleuth, bullseye, copernic

Remerciements

Au terme de ce stage intitulé Mise en place d'une méthodologie de veille à partir de la collection-teste de l'annuaire Sibel, nous adressons nos vifs et sincères remerciements à tous ceux qui de près ou de loin nous ont apporté leur soutien pour l'aboutissement de ce stage.

De prime abord, nous sommes redevables au maître de conférence Omar LAROUK et à la directrice du département des ressources documentaires Anne MAYER qui malgré leurs multiples occupations, ont bien voulu accepter la direction de la présente étude. Merci pour leurs observations qui nous ont aidé à mieux appréhender notre sujet sur plus d'un aspect.

Le corps professoral du DESS Réseaux d'information et document électronique dont les enseignements ont beaucoup enrichi nos réflexions mérite particulièrement nos égards.

Nous sommes également reconnaissant au personnel du département des ressources documentaires pour l'accueil et leur aide qui ont fait de ce stage une expérience professionnelle très agréable et formatrice. En effet, leur disponibilité à nos préoccupations sur le terrain a permis d'enrichir nos entretiens qui ont conduit à l'amélioration de nos vues. Qu'ils en soient remerciés.

Introduction

A l'heure où tout semble tourner désormais autour des progrès technologiques et de l'explosion de l'information, il est important de s'intéresser de plus près à la façon dont toute organisation peut surveiller son environnement pour être informée de toutes évolutions liées à son domaine de prédilection.

En effet, l'Internet offre de formidables possibilités pour détecter, atteindre ou surveiller une multitude de données, la mise en place d'un système de veille devient souvent une nécessité étant donné que sur le Net de multiples informations susceptibles d'enrichir un site ou une base de données sont disponibles.

C'est pourquoi une bonne gestion de l'annuaire Sibel (Sciences de l'information et des bibliothèques en ligne) s'impose pour pouvoir collecter, traiter et diffuser les informations et assurer la maintenance du site. Ainsi, Sibel dispose d'une collection susceptible de répondre aux questions posées sur l'avenir de la bibliothéconomie et sciences de l'information en générale. C'est ainsi qu'il est souvent visité par les spécialistes dans les domaines susmentionnés.

Dès lors, Armand affirme que le siècle que nous vivons connaît une révolution de l'information plus simple et plus radicale que celle qui marque les principales étapes de la communication. En d'autres termes, les informations suffisantes et pertinentes permettent une prise de décision rationnelle sur le développement scientifique, technologique et socio-économique d'une institution voire d'une personne.

Tout bien considéré, diffuser des informations en ligne est une bonne chose mais savoir surveiller son environnement pour pouvoir collecter, traiter et diffuser en temps réel nécessite une méthodologie de veille débouchant sur une mobilisation des ressources technologiques importantes. On peut toutefois affirmer que la veille n'est pas limitée à une classique surveillance de l'environnement, elle nécessite une volonté d'innovation, de repérage de technologies et d'informations nouvelles.

Son domaine s'étend de la recherche des informations jusqu'à leur interprétation et leur utilisation pour créer une vision de l'environnement. L'utilisation de l'Internet dans un processus de veille nécessite plusieurs phases : collecte et recueil des informations grâce aux outils de recherche et de veille, exploitation de cette masse documentaire pour en dégager un sens (validation, élaboration et synthèse), diffusion des résultats.

On le sait, s'y retrouver parmi la masse d'informations disponibles n'est pas chose aisée. Voilà pourquoi la veille est selon nous beaucoup plus large et peut s'appliquer tant aux entreprises qu'à l'étudiant qui projette sa carrière. La veille est avant tout un travail de terrain. C'est l'art de chercher les nouveautés et les changements intervenus par rapport à sa thématique en vue de s'alimenter davantage.

C'est ainsi que Sibel cherche à mettre en place un dispositif de veille pour répondre efficacement au besoin de la recherche des spécialistes de l'informations, en leur fournissant des informations pointues et adaptées qui les alertent sur tout ce qui peut intéresser leur activité. Ce dispositif de veille doit permettre aussi aux personnes qui alimentent cet annuaire de gagner du temps car leur charge de travail et aujourd'hui très lourde.

En effet, tout travail scientifique recourt à des procédures appropriées comprenant des méthodes et des techniques. On retiendra qu'au sens le plus large et le plus général, la méthode est constituée de l'ensemble des opérations intellectuelles par lesquelles une discipline cherche à atteindre des vérités qu'elle poursuit, les démontrer, le vérifier. La technique est considérée comme étant l'ensemble des instruments de travail.

Concrètement, pour réaliser et optimiser comme il se doit notre mission relative à la mise en place d'une méthodologie de veille à partir de l'annuaire Sibel, nous avons cherché d'abord à nous documenter (recherche documentaire) sur les divers aspects concernant les tâches à réaliser et ceux de la méthodologie permettant de

vérifier les liens hypertextes (internes et externes), de choisir les requêtes en vue d'interroger et d'évaluer les outils de veille à sélectionner.

Pour une étude d'analyse et d'évaluation, le recours aux méthodes quantitatives (statistiques) nous a été nécessaire. La méthode statistique avec ses techniques a permis la collecte systématique des résultats de vérification des liens et de l'interrogation des outils de veille sélectionnés et la mise en tableau.

De ce qui précède, comme on le sait, la mise en place d'une méthodologie de veille implique l'emploi d'outils de recherche dotés de fonctions de veille. Ainsi, Internet est une source d'information inévitable pour alimenter une démarche de veille tant au point de vue scientifique, commercial, concurrentiel que réglementaire. Du reste, il est difficile d'effectuer des recherches efficacement à l'intérieur des tâches quotidiennes sans être rapidement surchargé de travail. Il devient donc essentiel d'avoir une bonne connaissance des outils de veille et d'alerte pour pouvoir mettre en place un dispositif de veille permettant d'alimenter sa base. Aussi, l'utilisation des logiciels de vérification de la validité des liens est importante en vue de découvrir rapidement les liens défectueux et de les réparer pour offrir aux Internauteurs un site qui fonctionne bien.

Ce rapport de stage aborde cette question et se décompose en quatre parties.

La première partie plante le décor en présentant d'abord le département des ressources documentaires qui est notre organisme d'accueil (ses missions) pour pouvoir circonscrire par la suite les tâches à accomplir dans le cadre de notre stage.

La deuxième partie présente Sibel (Sciences de l'information et des bibliothèques en ligne) : son historique, sa structure et son mode d'emploi en vue de donner à nos lecteurs une idée sur cette base des données riche en information sur la bibliothéconomie et les sciences de l'information en générale.

La troisième partie est consacrée à la méthodologie de veille en commençant par montrer les différents aspects que revêt la veille : veille stratégique, environnementale, technologique, concurrentielle..., les outils de veille et leurs fonctionnalités.

La quatrième partie donne les détails relatifs aux tâches réalisées pendant la période du stage. Il s'agit de la vérification de la validité de liens de Sibel et de la proposition des outils de veille et les démarches appliquées dans le but d'atteindre les résultats escomptés.

En fin, une bibliographie, suivit des annexes, permet au lecteur intéressé d'approfondir le sujet

Partie 1 : Présentation du département des ressources documentaires

Le département des ressources documentaires est l'un des six départements que compte l'Enssib, Ecole nationale supérieure des sciences de l'information et des bibliothèques. Cette dernière a pour mission de former les conservateurs et les bibliothécaires de l'Etat et des collectivités territoriales, les cadres des services de documentation et d'information scientifique et technique, et de développer la recherche en sciences de l'information, bibliothéconomie et histoire du livre.

Grand établissement d'enseignement supérieur, elle a été créée par décret en 1992 et a succédé à l'ENSB (Ecole nationale supérieure de bibliothécaires), créée à Paris en 1963 et transférée à Villeurbanne en 1974. Début 1999, elle a fusionné avec l'Institut de formation des bibliothécaires, ce qui en fait le seul établissement national de formation des personnels de catégorie A des bibliothèques.

L'enssib est placé sous tutelle du Ministère de la jeunesse, de l'éducation et de la recherche. En annexe : l'organigramme de l'Enssib.

Le point ci-après présente le département des ressources documentaires.

1. Les missions du département des ressources documentaires

Les missions du département des ressources documentaires sont les suivantes :

- une mission de collecte et de diffusion des contenus dans le domaine de la bibliothéconomie et des sciences de l'information,
- une mission de partenaire de la pédagogie
- une mission de partenaire de la recherche
- une mission de pôle associé auprès de la BNF dans le domaine des sciences de l'information

En effet, les responsabilités du département s'exercent sur les deux bibliothèques, site de la Doua et site de Grand clément, le service des éditions, le site web (aspects éditoriaux)

Les clients qu'il sert sont :

- les bibliothécaires, conservateurs, professionnels de l'information en formation initiale à l'école
- les enseignants et les chercheurs en sciences de l'information et de la communication internes ou externes à l'école
- les professionnels des bibliothèques en formation continue à l'école ou qui préparent les concours internes ou encore qui souhaitent simplement se tenir informer
- les professionnels qui préparent les concours externes d'accès aux métiers des bibliothèques
- les étudiants qui à un moment de leur cursus, croisent les bibliothèques ou les SIC, étudiants en architecture, en histoire
- l'ensemble des personnels de l'école quels que soient leur statut et leur fonction

Les deux bibliothèques susmentionnées mettent à la disposition des lecteurs un fonds spécialisés en sciences de l'information gérés par un système de gestion de bibliothèque (logiciel PortFolio version 5 de la société Biblio Mondo) et un réseau de cédéroms.

En plus des ouvrages (imprimés) composés de monographies, périodiques, usuels, travaux des étudiants et de chercheurs, etc., l'Enssib, par le biais du département des ressources documentaire, possède un bon nombre des ressources électroniques :

- cédéroms (à prêter ou à consulter sur place sur les postes de la bibliothèque, ou à consulter à partir du réseau de l'enssib)
- des travaux des étudiants et des chercheurs (consultables ou non consultables, plus de 300 disponibles en version pdf en ligne)
- l'Opac de la bibliothèque (catalogue comprenant les notices des monographies, des ouvrages de références, des travaux universitaires, des CD-ROM).

En annexe les détails relatifs aux personnel, ressources matériels et financières ainsi que les activités du département des ressources documentaires

Les lignes suivantes précisent les tâches que nous avons réalisé au Département des Ressources Documentaires.

2. Définition du problème : objet du stage

Comme on le sait, l'annuaire Sibel met à disposition des spécialistes de l'information les informations les intéressants. Chaque jour sur Internet les pages web naissent, disparaissent et les modifications sont opérées.

Sibel se veut cet outil susceptible d'informer en temps réel les spécialistes de l'information des nouvelles parutions et changements intervenus dans les pages web correspondant à ses domaines de prédilection.

Ainsi pour mener à bien cette mission les responsables de Sibel ont jugé utile de mettre en place un dispositif de veille pouvant faciliter son alimentation.

De ce qui précède, l'objectif de mon stage au département des ressources documentaires s'articule sur les tâches ci-après. Il s'agit de :

- la vérification de la validité des liens hypertextes de Sibel à l'aide de logiciels de vérification des liens. Pour ce faire, il y a lieu de choisir, sur le Net, les logiciels susceptibles d'accomplir comme il se doit la mission relative à la maintenance du site ;
- proposer un ensemble d'outils de veille pouvant être utilisés en vue de mettre en place un dispositif de veille permettant d'alimenter Sibel par des nouvelles pages web et de signaler les changement dans celles (pages) déjà en ligne ;
- tester les outils de veille sélectionnés : choix des requêtes, interrogation des outils sélectionnés, évaluation des résultats ;

De plus, toutes les tâches susmentionnées ont été réalisées dans les installations du département des ressources documentaires (bibliothèque de la Doua) à partir du 1 juin 03 au 15 septembre 03.

Le point ci-après présente Sibel.

Partie 2 : présentation de Sibel

1. Aperçu historique de Sibel

Plusieurs étapes permettent de retracer l'historique de Sibel.

En effet depuis 2002, l'Enssib propose sur son site web un annuaire en ligne spécialisé en sciences de l'information et des bibliothèques dénommé Sibel (sciences de l'information et des bibliothèques en ligne).

Il sied de noter que Sibel tire son existence de Sitebib (coopération entre sites web en bibliothéconomes et sciences de l'information) qui fut un site créé en 1998 par une association d'établissement souhaitant coordonner leur travail relatif au signalement de ressources dans le domaine des bibliothèques et des sciences de l'information.

Le but de cette coopération était de proposer des liens ordonnés vers des sites indiquant eux-mêmes de liens vers des ressources de web, et de définir, en outre, le ou les sites de références pour chaque sujet. C'est ainsi que depuis 1998, Sitebib proposait près de 135 liens relatifs aux bibliothèques et aux sciences de l'information, repartis sur 11 principales rubriques thématiques. Ces liens pointaient vers les pages web des sites web des membres partenaires.

En effet, l'alimentation des rubriques fonctionnaient de la manière suivante : les établissements partenaires s'étaient répartis le sujet en fonction de leur pôles d'expérience. Ainsi, ils proposaient leurs liens dans un ou plusieurs domaines. Toutefois, si la gestion de Sitebib était partagée sur le plan du contenu, elle ne l'était pas en revanche sur le plan technique. En effet, toute proposition était dans un premier temps soumise sur une liste de discussion (sitebib@enssib.fr) commune aux partenaires, puis dans un deuxième temps véritablement intégrée, après une approbation des différents membres, par Dominique Lahary.

Il n'est pas sans intérêt de rappeler qu'avec le temps, il a été constaté que le système de gestion n'était pas le plus adapté à la situation. L'esprit de travail de Sitebib relatif à un mélange de coopération et d'autonomie ne pouvait plus être

véritablement respecté tant que les différents partenaires n'étaient pas autonome au sujet de l'alimentation des pages. De plus, la mise à jour constante n'était plus respectée. D'où la nécessité de mettre en place un système de base de données a été envisagée. Car il était indispensable d'améliorer les conditions de participation des différents partenaires et d'améliorer la fonctionnalité de Sitebib. Malgré cet effort, le projet de transformer le site web Sitebib (hébergé par l'ABF) en une base de données qui serait alimentée par les différents membres de l'équipe de Sitebib n'a malheureusement pu être mené à bien. Raison pour la quelle , le projet a été abandonné en vue d'envisager la fusion entre Sitebib et Sibel afin de créer un annuaire commun géré par le service de la bibliothèque de l'Enssib.

De ce fait, la base de données (Sibel) renseigne sur : les sites web d'associations, de bibliothèques, catalogues de bibliothèques francophones, des journaux électroniques, formations en Sic, listes de discussion, ministères, dossiers thématiques. Elle propose à ce jour environ 1000 références dans les domaines de la bibliothéconomie et des sciences de l'information.

Les internautes peuvent proposer des sites, qui sont validés par :

- Sabrina Guadagna, responsable de la Base Sibel – rubrique Bibliothèque, Formation, Organisations sciences de l'information
- Florence Favier, rubrique Bases de données, Livre et lecture, Ressources en ligne, Autres thèmes. De plus, il y a eu des stagiaires qui ont été chargés de pouvoir contribuer à la réussite du projet Sibel

Il y a mille et une chose à dire à propos de Sibel mais il nous faut présenter l'essentiel et en décrire son mode d'emploi.

2. Mode d'emploi de Sibel

Deux parties composent l'annuaire Sibel en vue de faciliter son fonctionnement. Il s'agit de mode administration et mode utilisateur.

2.1. Le mode administration

C'est à partir d'ici qu'on peut alimenter la base Sibel. Chaque personne qui alimente la base a un identifiant particulier. Cette mesure, outre sécuriser la base,

permet de remplir automatiquement le champ. Ainsi la validation de son mot de passe lui donnera accès à la page d'accueil de l'administrateur intitulé Administration. Cette page contient la liste des principales catégories du répertoire. L'administrateur peut ajouter des nouvelles catégories en utilisant la fenêtre qui se trouve en bas de la liste. A côté de chaque figurent les deux mots cliquables [éditer] et [supprimer] permettent à l'administrateur d'éditer les informations d'une catégorie ou bien de supprimer une catégorie qui existe Pour créer, éditer ou supprimer des sous catégories, l'administrateur peut accéder à une liste de sous-catégories et éventuellement ajouter un nouveau site soit alimenté la Base (cataloguer un nouveau site) en complétant les 6 champs correspondants.(cfr figure 1 et 2)

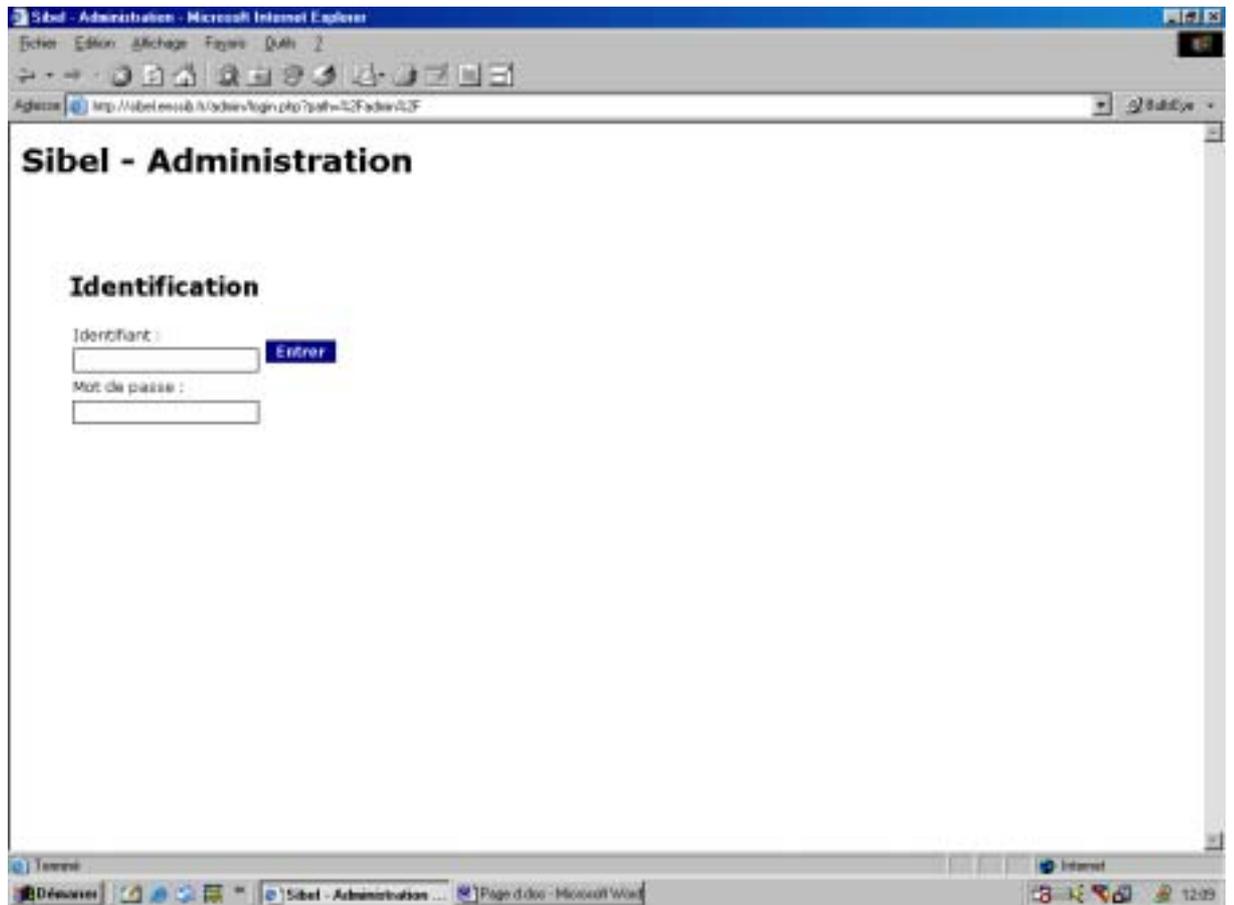


Figure 1. La partie administration

La figure ci-après donne les détails relatifs à l'alimentation (ajouter un nouveau site web) de Sibel et les 6 rubriques à remplir.



Figure 2. Partie relative au catalogage d'un site web pour Sibel.

Le point suivant parle du mode utilisateur.

2.2. Le mode utilisateur

Pour pouvoir consulter l'annuaire Sibel, l'utilisateur accède en cliquant sur l'option Sibel disponible sur la page web de la Bibliothèque de l'Enssib. C'est à ce niveau que l'utilisateur peut taper sa requête en vue d'obtenir l'information souhaitée.



Figure 3. Sibel : partie utilisateur

Le point qui suit tente de fixer les esprits sur la notion de la veille et ses composantes ainsi que la méthodologie relative à la veille.

Partie 3 : Méthodologie de la veille

1. La veille

La veille, c'est l'art de l'anticipation. C'est la capacité de suivre des centaines de sources d'information pour y repérer les nouvelles pages et sites ainsi que les modifications.

Il existe de nombreuses définitions de ce terme. L'AFNOR définit la veille comme un état de vigilance permettant la collecte, l'exploitation et la diffusion sélective et permanente d'informations sur l'environnement d'une organisation sous tous ses aspects (économiques, juridiques, culturels, sociaux, historiques, écologiques...). Activité continue et en grande partie itérative visant à une surveillance active de l'environnement technologique, commercial pour en anticiper les évolutions.

Tout compte fait, la veille qu'elle soit scientifique et technique, stratégique ou documentaire se formalise schématiquement dans la mise en place d'un processus organisé, composé de trois étapes successives qui sont la collecte, le traitement et la diffusion de l'information. Schématiquement car il ne s'agit pas d'établir une fois pour toutes un processus standardisé qui serait applicable à tout environnement. Au contraire, ce processus doit être directement adapté aux besoins et orientations spécifiques de la structure dans laquelle il prend place. Il suppose l'application d'une méthodologie, de traitement et de diffusion de l'information ainsi que l'emploi d'outils spécifiques et spécialisés de programmes d'interrogation, d'analyse et de traitement de l'information.

De ce qui précède, pour Rostaing Hervé, le mot veille est souvent employé pour désigner cette activité de surveillance de l'environnement des entreprises. On trouve ainsi les applications veille industrielle, veille stratégique, veille environnementale, veille informative, veille technologique, veille concurrentielle, veille commerciale... Toutes ces veilles ne sont pas censées couvrir les mêmes activités.

Le point suivant définit les différents types de veille sus indiqués.

2. Définition des concepts relatifs à la veille

En effet, la veille s'applique aux domaines d'activités les plus divers, elle concerne ainsi un très large éventail d'entreprises, d'organisations et de personnes.

Cela dit, la veille stratégique est un terme général qui englobe différentes autres formes de veille. De ce fait, en fonction de ce qu'on souhaite surveiller il est possible de distinguer autres expressions relatives à la veille.

2.1. Veille environnementale

Henri Samier et Victor Sandoval, affirment qu'un exemple de veille environnementale consiste à surveiller la notion de risque écologique. Cette surveillance n'est pas inhérente à un secteur ou un domaine, elle est totalement transversale. C'est dans un tel cas que la méthode fonctionnelle s'applique et permet de définir un carnet d'adresses de surveillance qui sera la base de la veille automatique.

2.2. Veille technologique

Ce type de veille constitue sans aucun doute la forme la plus connue de veille notamment en France et ailleurs.

La veille technologique est constituée par l'ensemble des techniques visant à organiser de façon systématique la collecte, l'analyse, la diffusion de l'exploitation des informations techniques utiles à la sauvegarde de la croissance des entreprises. En d'autres termes, son domaine d'observation est uniquement celui des informations scientifiques et techniques. C'est ainsi que Laurent Hermel souligne que la veille technologique qui est parfois veille scientifique et technologique s'intéresse aux acquis scientifiques et techniques, aux produits ou services, au design, aux procédés de fabrication aux système d'information

2.3. Veille concurrentielle

Comme son nom l'indique, son domaine d'observation est l'univers concurrentiel. La veille concurrentielle permet de pister les démarches actives, actions de développement, déploiement, vers d'autres secteurs et domaines d'activités, fausses pistes et leurs destinés à égarer les curieux, intrusions diverses, dépôts de brevets, travaux de recherche et ceci de la part de concurrents, directs et indirects. En effet, on peut retenir que, la veille concurrentielle s'intéresse aux concurrents actuels ou potentiels, aux nouveaux entrants sur le marché qui peuvent apparaître avec des substitutions. Cette veille est à la jonction de la veille technologique et de la veille commerciale.

2.4. La veille commerciale

La veille commerciale a pour but de collecter, traiter et diffuser des informations sur les produits et les marchés. Elle est donc intimement liée au lancement de nouveaux produits et à la détection de nouveaux marchés. Pour cette raison, une attention particulière est consacrée aux études de marché (qualitatives ou quantitatives). Cette veille concerne également l'observation des clients, des fournisseurs, des sous-traitants ou des distributeurs. Au delà des études marketing, il s'agit de s'intéresser à l'évolution des besoins des clients sur le long terme.

Il est cependant important de souligner que l'on parle de veille stratégique, de veille environnementale, de veille technologique, de veille concurrentielle ou de veille commerciale, le concept de base reste le même, c'est à dire la veille informative finalisée et décisionnelle. Il s'agit d'obtenir les informations pertinentes et utiles qui permettent une bonne réactivité de l'institution concernée face à des menaces externes ou face à des évolutions significatives de son environnement. C'est dans ce contexte que nous alignons la veille relative à cette étude.

De plus, la veille, on l'a compris, ce n'est pas seulement le recueil d'informations. L'information n'a de valeur que si elle parvient au bon moment et sous la bonne forme à la bonne personne qui en a le bon emploi. Pour cela, il faut mettre au point une méthodologie de veille.

3. Méthodologie de veille

Dans ce point, nous proposons le cadre méthodologique d'une veille automatique sur l'Internet. Cette méthodologie expose l'ensemble des processus de veille automatique utilisable aussi bien pour les entreprises, les PME que les personnes individuelles ayant la volonté de surveiller activement et légalement l'Internet avec efficacité sans perdre de temps. Pour cela, Henry Samier et Victor Sandoval préconisent cinq méthodes dissociées : la méthode des cibles, la méthode fonctionnelle, la méthode PmeFind, la méthode Director et la méthode ActivFind qui sont utilisables individuellement ou collectivement en fonction du besoin des cyberveilleurs.

La première méthode, dite des cibles, suit une logique de focalisation (ciblage) dont l'objectif est de trouver rapidement les référentiels d'une cible afin de la surveiller en continu et de surveiller les échanges de flux d'informations entre cette cible et son environnement extérieur.

La deuxième méthode, dite fonctionnelle, se fonde sur l'analyse fonctionnelle et l'analyse de la valeur et a pour but d'organiser la surveillance en fonctions autonomes.

La troisième méthode, PME, se fonde sur un besoin de méthode légère adaptée aux petites structures et organismes et s'applique donc à une petite ou moyenne entreprise.

La quatrième méthode, director, répond aux besoins des responsables de services en entreprise et qui ont des fonctions telles que recherche et développement, marketing, achat, logistique, commerciale innovation

De ce qui précède, le processus de la veille demande de la rigueur faute de quoi le travail effectué ne donnera aucun résultat, il est donc nécessaire d'adopter une organisation adaptée ainsi qu'une certaine méthodologie. Pour cela, il faut mettre au point une méthodologie de veille. Comme on le sait, la veille s'effectue en quatre étapes à savoir :

- orientation de la veille (spécification ou identification des besoins d'information ; adéquation besoin/réponse
- recherche et collecte de l'information (désignation/choix des outils)
- exploitation et traitement des informations (vérification et analyse)

- diffusion et appropriation des informations (désignation et mémorisation)

De plus, plusieurs méthodes de veille existent. Nous pouvons, en outre, souligner de types qui dépendent de l'outil à utiliser à savoir

- le push : vous paramétrez les informations que vous voulez recevoir et vous recevez (par courriel par exemple) des alertes vous informant des nouvelles informations ;
- le pull : vous paramétrez votre stratégie de recherche (votre requête ; périodicité de la veille) et vous modifiez cette stratégie d'après les résultats que vous obtenez

Il est vrai que chaque méthode répond à une veille particulière. Nous pouvons en distinguer trois types de veille :

- être tenu au courant d'un certain type d'information : vous indiquez quels sont les canaux d'information qui vous intéressent ;
- veille sur une page ou un site : vous avez identifié des sources (pages ou site) et vous voulez être tenu au courant des actualisations ou nouvelles informations qui apparaissent sur le site. Selon la veille, cela va de la simple alerte jusqu'au téléchargement d'un site (ou de la partie qui vous intéresse)
- veille sur une information : vous voulez savoir si de nouveaux sites traitent d'une information. Des agents, appelés agents autonomes automatisent votre veille en prenant dans certains cas en compte les résultats déjà trouvés.

En effet, la mise en place d'une veille sur un type d'information à partir des outils de veille et d'alerte ne repose pas sur l'emploi d'une méthode unique et infaillible applicable en toutes circonstances. Les outils sont multiples, de qualité inégale, et les solutions nombreuses.

La veille requiert donc l'emploi complémentaire :

- d'outils automatiques dotés de fonctionnalités de veille
- des fonctions des outils bureautiques (navigateur, courrier électronique, etc.)
- des outils d'aide à la recherche d'information incarnés par les répertoires, moteurs, sites thématiques ou spécialisés.

Pour les outils automatiques de surveillance de sites et des pages, on préférera ceux qui permettent un paramétrage assez précis pour éviter de recevoir des alertes inutiles : mots – clés, filtrage des changements hors contenus, etc.

L'utilisation des outils bureautiques doit fournir une assistance aux opérations de veille :

- les navigateurs permettent l'organisation de dossiers et de sous-dossiers de veille
- les logiciels de messagerie électronique peuvent jouer un rôle utile (ex : eudora : créer de boîte aux lettres, un filtre)

En effet, actuellement les outils de veille sont nombreux comme on l'a souligné ci - dessus. Ainsi, il est tout de même important pour nous de parler des agents pull et push.

4. Les outils de veille

4.1. Typologie d'outils de veille

Deux grandes familles des outils de veille qu'on appelle parfois « agents ». Il s'agit :

- des outils de veille (agents de type pull) : l'utilisateur est actif
- des outils de veille (agents de type push) : l'information est poussée jusqu'à l'utilisateur.

4.2. Les outils de veille on line : technologie pull

La finalité des agents est d'aider à surveiller un ou plusieurs sites web et, éventuellement, les archives des forums de discussion ou des listes de diffusion. On distingue deux possibilité :

Les agents d'alerte par e-mail (MyUpdate, The informant...) vous envoient un courrier électronique chaque fois que tout ou partie de la page web que vous avez sélectionné est modifier (MyUpate), ou bien quand une recherche d'informations produit de nouveaux résultats au sein d'un ou plusieurs moteurs (The informant). Dans certains cas, (The Informant), il est également possible de filtrer les changements à surveiller en choisissant des mots – clés appropriés.

Les agent d'alerte « off-line » sont des logiciels clients qui s'installent sur votre ordinateur et que vous pouvez finement personnaliser afin de réaliser des activités de veille. Ces agents réalisent les mêmes tâches que les agents on line mais

permettent également de copier sur votre ordinateur des sites entiers. Il suffit de leur indiquer l'adresse de site et l'heure à laquelle vous souhaitez qu'ils s'activent. Ils vont se connecter de façon automatique au réseau afin de rapatrier les données qui vous intéressent. Vous pouvez ensuite mener vos activités de recherche et de veille à partir des données enregistrées localement avec un affichage immédiat. L'inconvénient majeur de ces agents est qu'ils doivent (pour des raisons d'efficacité) se baser sur les dates de dernières modifications qui sont souvent peu pertinentes. Il suffit qu'une virgule ou qu'une image soit déplacée pour que l'agent interprète qu'il y ait eu changement.

4.3. Les outils de veille : technologie push

Les agents « push » sont très efficaces pour suivre l'actualité et pour diffuser des informations à des cibles très larges. Là aussi, on distingue deux cas :

Les agents « push on-line » permettent d'obtenir des nouvelles personnalisées sous forme de rubriques thématiques (actualité sportive, actualité politique, actualité financière, actualité météo,...etc) ainsi que beaucoup de services annexes (réservations, achats, loisirs, communautés électroniques...). Tout fonctionne en ligne et vous n'avez pas besoin d'installer de logiciel client. Les informations sont diffusées au sein d'une page de sommaire personnelle qui s'active dès lors que vous avez défini votre profil d'utilisateur. Ce genre de service est simple et assez efficace mais très limité pour des activités de veille sophistiquées. En effet, à part les contraintes liées à la connexion Internet, vous disposez souvent d'un choix très restreint d'options pour personnaliser votre page.

Les agents « push off-line » reposent sur le même principe mais nécessitent d'installer un logiciel. Ainsi, une fois que vous avez défini votre profil au sein d'une liste d'options de plus en plus riches, votre logiciel télécharge automatiquement, à partir d'un serveur distant, le contenu des chaînes auxquelles vous avez souscrit.

Dans les lignes qui suivent, nous allons tenter de définir les fonctionnalités des outils de veille

5. Quelques fonctionnalités des outils de veille

Les fonctionnalités des outils de veille évoluent régulièrement. Ces outils sont dotés des fonctions permettant de mettre en place un dispositif de veille et ce, pour n'importe quelle organisation. Il s'agit des fonctions suivantes :

5.1. Fonction d'alerte

L'alerte met en œuvre la fonction push et fonctionne souvent sur la base d'un profil établi par l'utilisateur. Elle consiste en l'envoi d'un signal, consécutivement à l'apparition d'une information couvrant les termes de sa demande.

Il existe des alertes dont on peut paramétrer la fréquence afin de regrouper les messages. Cela permet de solutionner le problème en partie mais seulement pour des informations dont on sait qu'elles peuvent tolérer un délai de traitement. Cette forme d'alerte est proposée par exemple sur les sites d'éditeurs scientifiques comme Blackwell publisher ou Elsevier qui procèdent à l'envoi de sommaires de revues et/ou de tables des matières des dernières publications. On peut distinguer deux usages d'alerte. La première concerne des personnes qui utilisent à titre individuel : dans ce cas, ils auront intérêt à utiliser la fonction de façon raisonnée, en privilégiant si possible les outils qui permettent de définir des profils assez fins et retournent des résultats couvrant bien les requêtes. C'est le cas des chercheurs, qui vont avoir à gérer une importante quantité de courriers électroniques, sans en savoir nécessairement le temps. La deuxième utilisation de l'alerte peut s'inscrire dans un processus de veille orientée et organisée qui donne lieu à l'établissement d'un programme de gestion des tâches et du temps

5.2. Fonction d'analyse

Elle permet le classement d'un ensemble de résultats en catégories sémantiques. Certains moteurs de recherche disposent de cette fonction. De façon générale, elle consiste à extraire des documents collectés les termes et syntagmes nominaux connexes et sémantiquement.

L'analyse, dans la mesure où les algorithmes appliqués à la méthode d'extraction sont capables de produire une catégorisation acceptable, constitue une aide très utile et intéressante pour l'utilisateur. L'outil est supposé se substituer à l'intervention de ce dernier à un moment de l'étape de sélection des documents

adaptés. En effet, c'est la consultation des résultats par balayage séquentiel d'une grande quantité d'information qui est évitée à l'utilisateur grâce à l'analyse, et celle-ci fournit automatiquement des éléments (les termes extraits) de représentation du corpus. Outre le gain de temps, cette opération présente l'avantage de pallier les problèmes de classements de résultats des outils automatiques de recherche. Ce n'est plus la pertinence attribuée automatiquement aux documents qui est prise en compte par l'analyse, mais les contenus de l'ensemble des documents. Le deuxième avantage qui en découle est que les résultats se situant au-delà des deux ou trois premières pages consultées habituellement par la majorité des utilisateurs seraient pris en compte.

5.3. Fonction filtrage

Cette tâche doit résoudre le problème de la surcharge d'information en limitant le nombre de réponses qui arrivent à l'utilisateur. Plusieurs outils disposent de cette fonction dont l'activation est à la charge de l'utilisateur. Parmi les outils testés, ce sont principalement les métamoteurs Copernic, BullsEye et aussi Strategic finder qui ont permis d'observer cette fonction.

Le filtrage des résultats est parfois très sélectif et peut donner lieu à la suppression d'une grande partie des documents rapportés. Dans ce cas, il est une indication de non-adéquation de la requête avec le mode d'interrogation choisi et doit amener l'utilisateur à réduire le nombre de termes de sa requête ou bien à modifier la sélection des outils. En effet, si les opérateurs booléens choisis par l'utilisateur ne sont pas supportés par les différents outils interrogés, c'est l'opérateur par défaut (généralement « ou ») qui prévaut. Cela explique que très peu de documents rapportés correspondent au final à la requête initiale si elle est composée de plusieurs termes.

5.4. Conservation des requêtes

Certaines requêtes doivent être régulièrement réitérées, et il peut être nécessaire de les appliquer régulièrement à plusieurs outils de recherche. Cela suppose de passer du temps à saisir successivement des formules d'interrogations dans les cases à renseigner de chaque outil. Ce problème peut se solutionner de façon basique en utilisant les ressources bureautique telles que les favoris d'Internet

explorer par exemple. Il suffit de conserver les premières pages de résultats d'une interrogation pour chaque moteur sollicité, et de les relancer périodiquement (c'est bien la page de résultats qu'il faut conserver, les termes saisis par l'utilisateur sur la page d'accueil du site n'étant pas enregistrés). Mais la conservation des requêtes est une option que possèdent plusieurs outils de veille et d'alerte, dont ceux qui autorisent d'un profil personnalisé.

5.5. Surveillance des modifications de contenus

Cette fonction est intégrée par les outils de surveillance de pages web, et c'est dans une moindre mesure l'une des fonctions des aspirateurs de site (les aspirateurs de sites vont, eux, signaler les pages qui ont subi une modification suite à l'activation de la mise à jour par l'utilisateur). Elle est destinée à éviter à l'utilisateur de répéter une tâche récurrente et laborieuse lorsqu'il s'intéresse à un grand nombre de pages.

Par défaut, la surveillance s'exerce le plus souvent quotidiennement. Des alertes sont adressées à l'utilisateur dès qu'un changement est observé. Elles peuvent prendre la forme d'un signal apparaissant sur le bureau de l'utilisateur, d'un courrier électronique en html qui permet d'accéder directement à la page concernée, d'un message inscrit sur le site de l'outil grâce à la personnalisation de l'interface, ou d'un message sur le téléphone portable, ordinateur de poche.

Cette fonction d'autonomie d'un agent est appelée à une grande popularité, notamment dans le cas de veille sur un sujet donné.

5.6. Ciblage des sources interrogées

Une fonction typique des outils de veille et d'alerte est (de rassembler un grand nombre de sources à interroger en fonction d'une thématique ou d'un type d'information. c'est par exemple le cas des annuaires et répertoires thématiques, par exemple celui de l'UNESCO qui tend à un inventaire exhaustif des différents types de sources en sciences humaines et sociales (centre de documentation, organisme et centres de recherche, établissements et institutions, banques de données...). Les outils de recherche d'articles de presse intègrent cette fonctionnalité en recensant pour l'utilisateur une très grande quantité de titres de journaux, revues, agences de presse, etc...qu'ils organisent en fonction d'un

classement thématique. L'interrogation de ces sources se fait de manière globale ou sélective. L'intérêt de ces outils est donc d'épargner à l'utilisateur l'interrogation successive d'un grand nombre de sources distinctes grâce à un formulaire d'interrogation unique. Certains outils comportent une option de personnalisation de l'interface, qui contribue à simplifier l'utilisation en réduisant le nombre d'étapes préalables à l'interrogation.

La partie suivante permet de décrire et d'évaluer les tâches, relatives à notre stage

Partie 4 : Optimisation de la gestion de l'annuaire Sibel

Cette quatrième partie tente de décrire la démarche relative à l'évaluation de l'annuaire Sibel en vue d'optimiser sa gestion, à savoir la vérification des liens hypertexte, le choix de requêtes, l'interrogation et évaluation des outils de veille utilisés.

1. Des outils de vérification des liens hypertextes

Pour vérifier la validité des liens internes et externes de Sibel, nous nous sommes aperçu qu'il était assez fastidieux de voir les pages une à une pour détecter les liens défectueux. On le sait, Sibel compte près de mille liens qui orientent vers les pages web disponibles et d'autres sites sur les sciences de l'informations et des bibliothèques. C'est pourquoi il était intéressant pour nous de pouvoir utiliser les logiciels conçus pour ce genre d'activités.

L'utilisation de ces logiciels (gratuits sur l'Internet) permet de connaître l'état des liens hypertextes d'un site web et d'identifier les liens défectueux en vue d'envisager leur réparation et ce, pour assurer la maintenance du site.

A entendre les internautes, peu nombreux sont ceux qui apprécient les erreurs sur l'Internet notamment l'erreur 404 (les pages web supprimées ou démenagées).

Le webmaster se doit donc de veiller à ce que tous les liens soient fonctionnels entre les différentes pages de son site.

Pour nous, cette vérification devrait s'étendre à l'ensemble des liens disponibles de Sibel, qu'il s'agisse des liens vers ses propres pages et les autres sites web.

La vérification des liens se faisait régulièrement, soit au début et à la fin de chaque semaine. Pour y parvenir, deux logiciels ont été sélectionnés. Il s'agit de logiciel Check Web Links et Xenu's Link Sleuth que nous présentons dans les points ci-après.

1.1. Les logiciels de vérification des liens

1.1.1 Choix des logiciels

La création d'un site web oblige en général à mettre en ligne un grand nombre de pages. Ces pages ont souvent de nombreuses caractéristiques communes (position de certains calques, police, titres et couleurs pour n'en citer que quelques unes). La maintenance d'un site web exige aussi l'utilisation des logiciels de vérification des liens.

Ainsi, le choix d'un logiciel de vérification des liens n'est pas difficile, que dans la mesure où le Web nous en offre en abondance. A moins d'avoir à gérer un répertoire de signets aux dimensions herculéennes, et encore, on peut facilement obtenir satisfaction avec l'un ou l'autre des logiciels téléchargeables gratuits.

Pour choisir et sélectionner les logiciels susceptibles de vérifier les liens de Sibel, la meilleure façon de le faire pour nous était d'essayer puis de comparer. C'est ainsi que nous avons retenu Check Web Links. De plus, il était possible de consulter en ligne les résultats d'une enquête effectuée en janvier 2002 auprès des membres de l'Association des professionnels de l'information et de la documentation, des résultats qui s'avèrent flatteurs, sommes toute, envers les fonctionnalités du logiciel Xenu's Link Sleuth. C'est grâce à cette enquête que nous avons aussi opté pour Xenu's Link Sleuth. En ce qui nous concerne, nous avons remarqué, après essaie et comparaison, que Check Web Links et Xenu's Link Sleuth sont faciles à utiliser et donnent de bon résultats.

Pour ce qui est de l'utilisation de ce types de logiciels, Catherine Benier affirme que la plupart des logiciels de vérification des liens hypertextes, qu'ils soient gratuits ou payants, s'utilisent aisément ; le téléchargement est rapide.

En effet, les seules tâches manuelles requises que nous avons constaté, en utilisant les logiciels de vérification des liens susmentionnés sont de copier l'adresse du site où se trouve la liste de liens à vérifier (<http://sibel.enssib.fr>), la coller à l'endroit désigner dans l'interface du logiciel utilisé puis lancer le processus automatique de vérification tel que nous allons le voir dans les copies d'écrans après la présentation de chaque logiciel de vérification de liens sélectionnés.

1.1.2 Check Web Links

Il s'agit comme, on l'a dit, d'indiquer à Check web links l'adresse de la page d'accueil de votre site (<http://sibel.enssib.fr>) et la vérification peut commencer. (figure 4)

Toutefois, Check Web Links n'est pas avare en paramétrage, bien au contraire. Ainsi il est possible d'optimiser la vérification des liens. Il est également possible de préciser la nature de la vérification des liens. Vous pouvez vérifier les liens externes qui renvoient vers d'autres sites web.

Contrairement à la majorité de ses concurrents, Check Web Links permet également de vérifier la validité des liens de votre site web en local, c'est à dire sur votre disque dur lorsque vous êtes déconnecté.

La fiche technique du logiciel Check Web se présente comme suit :

- Editeur : Alterego
- Taille du fichier 70 Ko
- Version 1.2
- Prix gratuit
- Langage : Anglais/Français
- Système : windows 95/98/Me/NT/2000/XP

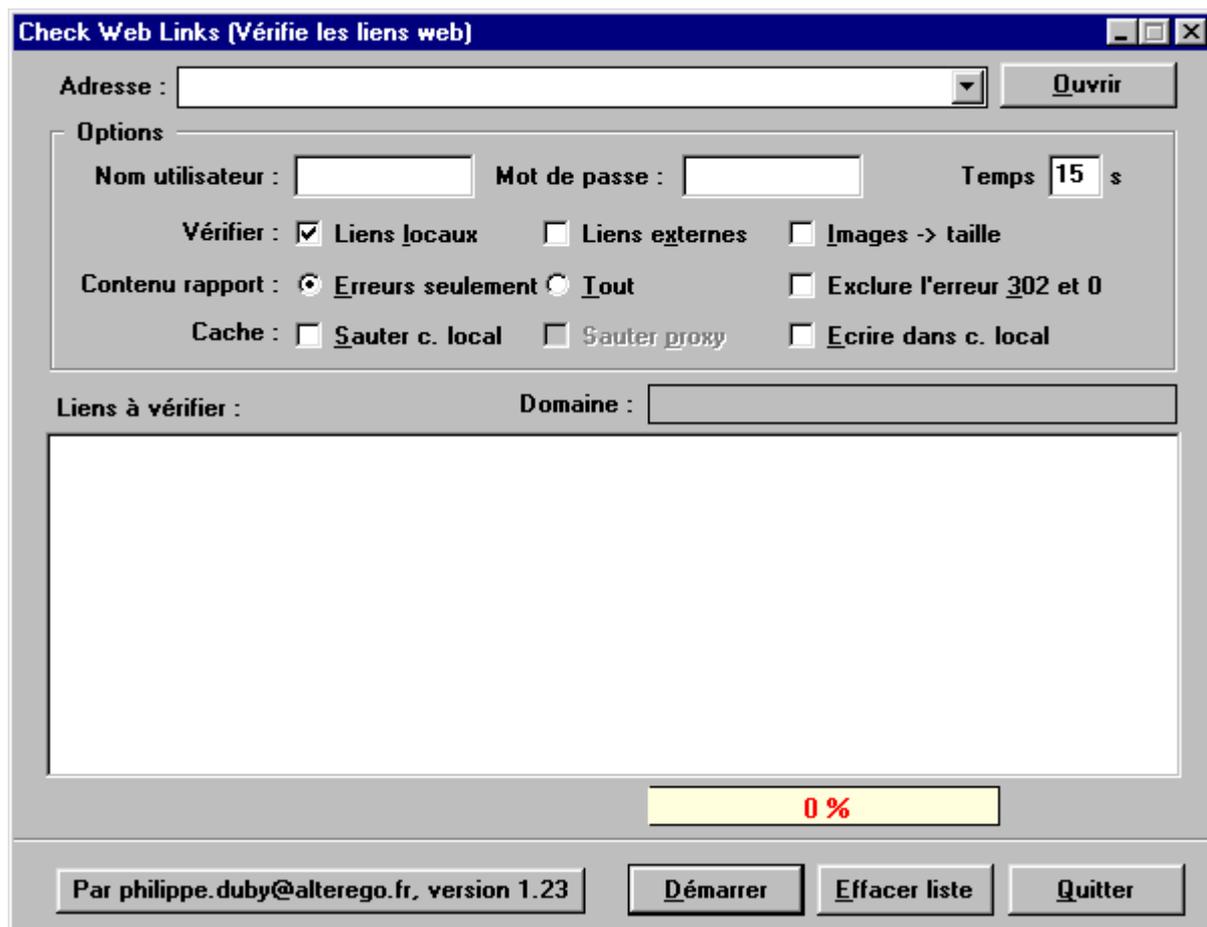


figure 4: logiciel Check Web Links

Le logiciel gratuit Check web links permet d'analyser les pages d'un site web et de repérer d'éventuels liens (internes et externes) qui ne fonctionnent pas. Il s'agit d'un programme 32 bit, réalisé en C++ par l'informaticien Philippe Duby, dont le source est libre. Il n'est plus mis à jour depuis 1999 mais fonctionne très bien et son emploi est simple.

1.1.3 Xenu's Link Sleuth

Xenu's Link Sleuth (<http://homme.snafu.de/tilman/xenulinkhtml>) reste certainement le plus connu et le plus utilisé des freeware vérificateurs de liens. Xenu's link Sleuth vérifie les liens internes et externes de tout site web, y compris les liens des images, des cadres, des plug-in, des scripts, etc... Il donne les codes d'erreur des liens brisés, et produit un plan du site analysé. Il permet d'identifier les erreurs dans les liens internes ou externes (une faute de frappe est arrivée), de

repérer les fichiers manquants (oubli de chargement, modification du nom...) et de s'assurer que les liens externes existent toujours.

Xenu's link sleuth peut être utilisé non seulement pour vérifier ses propres pages, mais aussi pour analyser et mieux comprendre la structure d'un site extérieur. Xenu's Link Sleuth est en anglais, mais facile à utiliser. Sa fiche technique se présente comme suit :

Taille : 223 kb

Licence : Gratuitiel

Système : Windows 95/98/Nt

Langue : Anglais

Auteur : Tilman Hausher

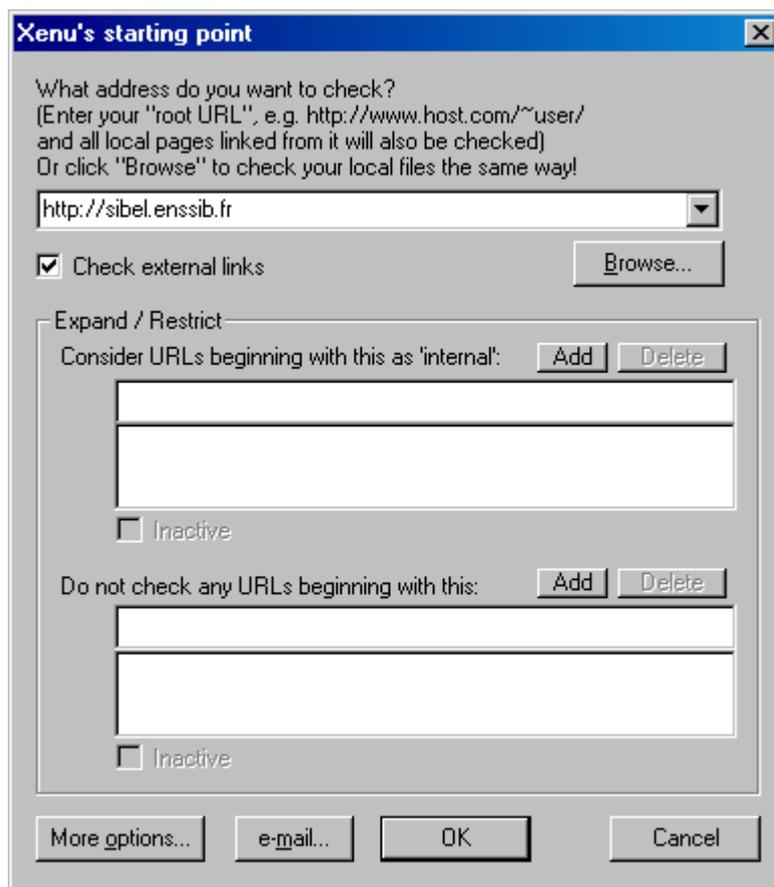


Figure 5 : Logiciel Xenu's link sleuth

Xenu's Link Sleuth s'assure que tous les liens sont corrects, aussi bien les liens internes (HTML, images ou autres) que les liens externes. Il peut vérifier une copie locale de votre site ou bien directement sur le Web.

Xenu's Link tout comme Check Web Links peuvent générer un rapport HTML qui contient la liste des liens brisés, la liste des liens externes et un plan de votre site. Ce rapport est utile dans la mesure où il fournit les éléments susceptibles de résoudre le problème lié à la maintenance du site.

1.1.4 Résultat de vérification des liens

Pour mieux interpréter les résultats de vérification des liens, il serait intéressant, d'avoir une idée sur les codes d'erreur Internet les plus populaires qu'on peut trouver sur le net. La liste des codes susmentionnés nous la donnons en annexe de la présente étude.

Les deux copies d'écran ci-après (check web lenks et xenu's link sleuth) renseignent sur les résultats obtenus lors des vérifications de la validité des liens hypertextes de la collection Sibel (liens internes et externes)

1.1.5 Check Web Links

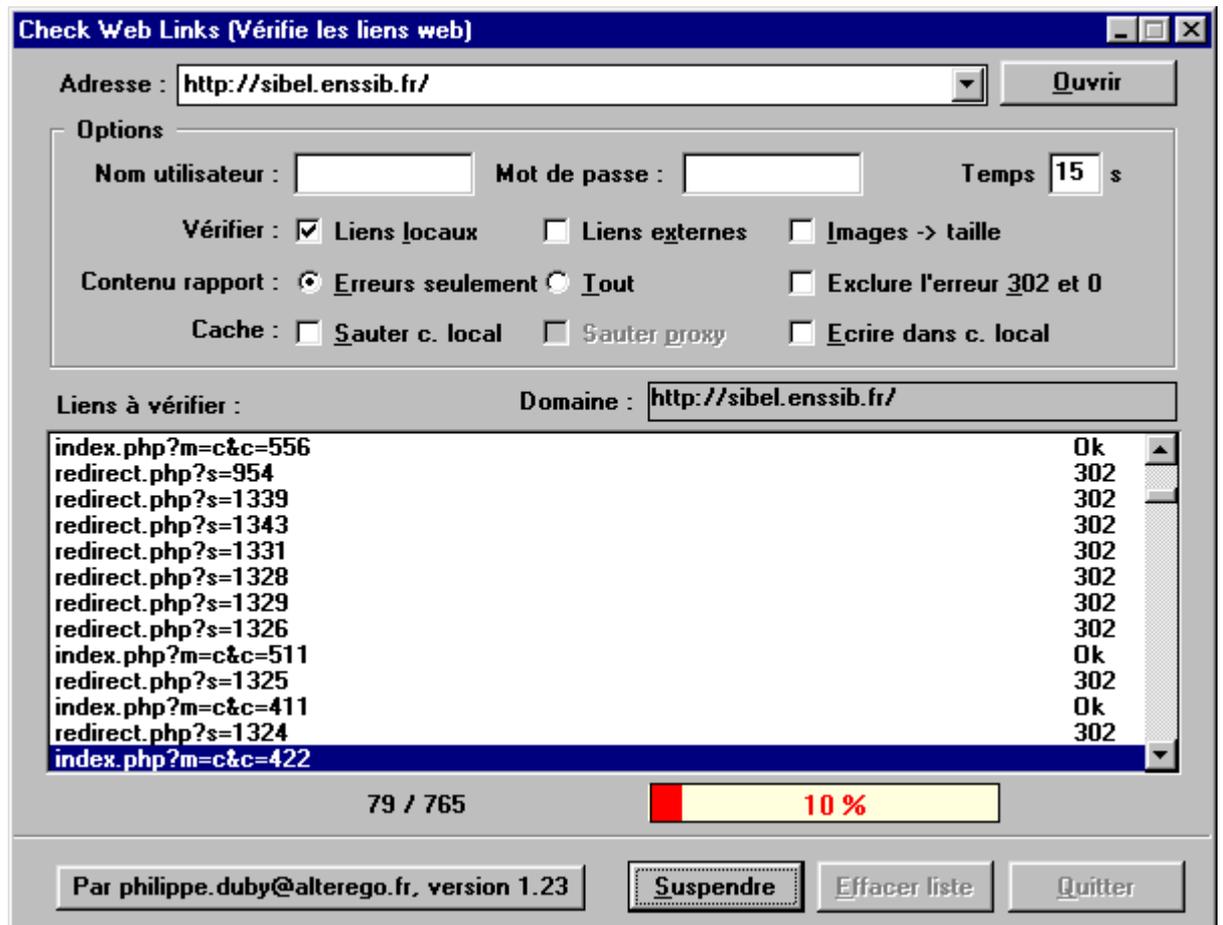


Figure 5 : Check Web Links : Vérification des liens

L'écran ci dessus montre, en guise d'exemple, l'un des codes erreurs Internet enregistrés lors de vérification des liens de Sibel par le logiciel Check Web links

L'écran ci-après donne le rapport de vérification des liens par le logiciel xenu's link sleuth

1.1.7 Tableau des codes erreurs Internet enregistrés dans le site Sibel

301	Document déplacé de façon permanente.
302	Document déplacé de façon temporaire
403	Pas d'autorisation d'accès au serveur
404	La page demandée n'existe pas. C'est l'erreur la plus courante.
500	Erreur interne du serveur
502	Mauvaise passerelle d'accès

Comme le montre le tableau ci dessus, les codes erreurs Internet enregistrés sont les codes : 301, 302, 403, 404, 500, 502. Il y a lieu de rappeler que c'est au début et à la fin de chaque semaine qu'on procède à la vérification des liens. Et, les codes erreurs sus indiqués étaient assez souvent enregistrés lors du lancement des logiciels pour pouvoir vérifier les liens. On souligne souvent que plus le site web est importante, plus il y a le risque de pouvoir rencontrer ne serait-ce l'un des codes erreurs disponibles sur le net. C'est pour dire que Sibel est une base de données importable (1000 liens) c'est ainsi que nous avons enregistré des liens brisés soit momentanément (code 302), soit définitivement (code 404 : la page demandée n'existe plus), soit encore le document souhaité est déplacé de façon permanente (code 301) etc. On le sait, c'est chaque jour sur le Net que les sites naissent et disparaissent. C'est pourquoi, on ne peut en aucun cas être surpris dès lors, pendant la navigation, qu'on tombe sur l'un des codes erreurs Internet sus indiqués. De plus, le fait de vérifier assez souvent la validité des liens atteste que les responsables du Site tiennent à son entretien et par conséquent, cherchent à satisfaire comme il se doit leurs visiteurs.

Il est tout de même important de noter que de tous les codes erreurs le plus courant reste celui de la page introuvable, donc le célèbre code 404. Ce message survient lorsque la page a été déplacée ou retirée d'un serveur après le passage d'un robot. Ce code est un indice du rafraîchissement de la banque de données de l'automate de recherche.

Comme nous venons de le dire, la présence des codes erreur sur une base de données décourage les internautes et diminue la qualité du site web et la considération qu'on l'attribue.

De ce qui précède, il n'est pas sans intérêt de dire un mot sur la démarche à respecter pour pouvoir réparer les erreurs. C'est une tâche qui doit toujours être envisagée après avoir effectué la vérification des liens et que les erreurs sont constatées.

La procédure ci-après permet de réparer les erreurs sans risque de commettre d'autres.

En effet, pour réparer une erreur, la première étape est de cliquer sur le lien pour s'assurer qu'il est bien « mort ».

- il arrive souvent qu'un lien ne soit rompu que de façon temporaire et qu'il fonctionne à nouveau
- Parfois, vous serez redirigé vers une autre URL et vous devrez mettre à jour la nouvelle adresse du site.
- Vérifiez qu'il n'y a pas de fautes de frappe dans l'URL
- Si le serveur lui-même fonctionne, il pourrait s'agir d'un problème temporaire.
- Si l'adresse racine du site (root URL) ne fonctionne pas, vous pouvez envisager la suppression du lien car le site pourrait effectivement avoir définitivement disparu.

De ce fait, la maintenance du site s'impose c'est ainsi qu'il convient donc d'effectuer régulièrement une vérification de la validité des liens externes vers d'autres sites, des liens internes. La plupart des outils recherche et signale également les liens redirigés qu'il convient de mettre à jour.

Après la vérification de la validité des liens hypertextes et, pour tester les outils de veille sélectionnés, il était indispensable, pour nous, de procéder d'abord au choix de requêtes et ce, en respectant une démarche bien déterminée.

2. Des outils de veille

2.1. Choix des requêtes

Le choix des requêtes, on le sait, est une étude indispensable dans la mesure où, il permet d'interroger les outils de recherche en vue d'obtenir les informations souhaitées.

En effet, la requête est un mot - clé ou une association des mots - clés. Elle peut être primaire ou secondaire.

La requête primaire est un mot – clé générique, souvent seul, générant beaucoup de trafic. Ex : bibliothèque, documentation etc.

La requête secondaire : requête primaire associée avec un ou plusieurs autres mots – clés. Ex : bibliothèque nationale,

Beaucoup de recherches se font avec, non seulement, un mot générique, sinon avec des combinaisons de deux, trois ou même quatre mots. La moyenne, selon nos observations, est de deux à quatre mots par recherche.

En effet, la recherche permet de définir une image de la réponse que l'on veut obtenir. D'où la nécessité de choisir les mots-clés permettant de renvoyer les réponses souhaitées. Il est donc extrêmement important de passer du temps à choisir soigneusement ces mots – clés. C'est ce que nous avons fait pour cette étude. Ainsi, nous avons mis du temps pour choisir soigneusement les mots – clés en parcourant presque la quasi totalité des pages web Sibel.

Pour nous, l'objectif de choix de mots – clés est d'obtenir un taux de réussite élevé lors des recherches. Pour cela, nous nous sommes mis à la place des internautes et examiner chaque possibilité d'expressions simples et complexes pouvant mettre en évidence les références de Sibel. Car la qualité et la valeur des résultats d'une recherche par mots – clés dépendent directement des mots – clés choisis.

Sachant que les outils de recherche tout comme de veille, par le biais de moteurs de recherche cherchent les mots – clés dans trois endroits (dans le titres de pages web, dans les listes de mots – clés non structurées proposés par le chercheur, dans le résumé du sujet traité dans le site ou page), nous avons respecté les procédures relatives au choix de mots – clés à savoir :

- lister les mots – clés du secteur d’activité de Sibel et de choisir les associations de mots pouvant réellement taper les internautes ;
- les mots – clés choisis devraient caractériser des pages web du site Sibel ;
- faire sortir les mots – clés les plus importants de manière à mettre les pages web de Sibel en valeur dans les résultats de recherche des outils de veille sélectionnés
- saisir correctement les requêtes et lancer la recherche
- tenter de choisir les mots – clés (association) qui sont assez long sans l’être trop. Les mots clés courts peuvent produire une abondance de résultats erronés et les mots long peuvent être trop précis
- placer entre guillemets les requêtes afin que les mots – clés formant la requête se trouvent cote à cote dans les résultats obtenus.

De plus, nous avons compris qu’effectivement, plus les choix de mots – clés (requêtes) est pensé, meilleurs sont les retours. (En annexe, la liste des requêtes choisies).

Le point ci-après tente de donner quelques notions relatives à la pertinence.

2.2. Quelques notions de pertinence de résultats par rapport à Sibel

Sylvie Dalbin définit la pertinence comme étant le degré de corrélation entre une question et la réponse apportée, la pertinence est un concept clé de la recherche d’information depuis des années 50 .

A l’heure actuelle, la notion de pertinence préoccupe tout chercheur dans la recherche d’informations. D’une façon générale, la notion de pertinence vise le document et son adéquation à une requête généralement définie par le croisement de deux notions transversales de la recherche d’information et de la RDI (recherche documentaire informatisée) ; le bruit et le silence documentaire :

- bruit : documents retrouvés non pertinents
- silence : documents pertinents non retrouvés

On le sait, le bruit et le silence sont mesurés par le taux de rappel et de précision :

- taux de rappel : proportion de documents pertinents retrouvés par rapport à l’ensemble des documents pertinents de la base documentaire

- taux de précision : proportion des documents pertinents par rapport à l'ensemble des documents récupérés.

De plus, comme l'indique Sylvie Dalvin « la pertinence concerne plusieurs aspects de recherche d'information : au moins six aspects de la pertinence :

- la pertinence de la représentation du document par rapport à la requête
- la pertinence de la représentation du document par rapport au thème
- la pertinence du document lui – même par rapport à la requête
- la pertinence du document par rapport au thème
- la pertinence de la requête par rapport au système documentaire
- la pertinence de l'information trouvée par rapport au besoin initial »¹

De ce qui précède, avant toute démarche d'évaluation, il serait intéressant de bien caractériser la pertinence de ce qu'on veut mesurer et évaluer. En ce qui nous concerne, pour mieux évaluer et comparer les outils de veille sélectionnés, la notion de pertinence par rapport à Sibel sera caractérisée par le positionnement des documents Sibel (sites ou pages web Sibel) par rapport à l'ensemble des documents pertinents trouvés, et ce par outil de veille interrogé.

Ainsi donc, pour mieux illustrer et appliquer cette notion de pertinence, nous avons construit une grille d'évaluation des résultats de recherche par rapport à Sibel.

2.3. Grille d'évaluation des résultats d'interrogation des outils de veille

Critères	Description	Pointage
Pertinence de réponses	Le site ou pages web Sibel est dans les 5 premières références	20 points
	Les sites ou pages web Sibel est entre la 6 ^{ème} et la 10 ^{ème} références	15 points
	Le site ou pages web Sibel est entre la 11 ^{ème} et la 15 ^{ème} références	10 points
	Le site ou pages web Sibel est entre la 16 ^{ème} et 20 ^{ème} références	5 points
	Le site ou page web Sibel est entre la 21 ^{ème} et plus	0 point

Tableau construit par nous même

La grille d'évaluation ci-dessus permet d'évaluer la pertinence de réponses relatives aux outils de veille sélectionnés et interrogés.

5 niveaux de pointage forment cette grille et facilitent notre démarche d'évaluation des outils de veille relatifs à la pertinence des résultats par rapport à Sibel.

Ces cinq niveaux de pointage peuvent être interprétés comme suit :

- Dans les 5 premières références du résultat de recherche, le (la) ou les sites et pages web Sibel sont considérés pertinents et l'outil de veille obtient 20 points multipliés par le nombre des pages ou sites web Sibel qui y sont alignés.
- Entre la 6^{ème} et la 10^{ème} références, les sites ou pages web sont pertinents. 15 points multipliés par le nombres des sites ou pages web Sibel sont accordés à outil de veille
- Entre la 11^{ème} et la 15^{ème} références, les sites ou les pages web Sibel sont aussi pertinents pour 10 points accordés à l'outil de veille multipliés par le nombre des sites et pages web Sibel
- Entre la 16^{ème} et la 20^{ème} références, les sites ou les pages web sont pertinents. 5 points seulement seront accordés à l'outil de veille et multipliés par le nombre des pages et sites web Sibel disponibles
- Entre la 21^{ème} et plus, les sites et les pages web Sibel ne sont pas pertinents. Par conséquent l'outil de veille aura zéro point.

L'objectif poursuivit dans cette analyse est de mettre en exergue l'outil de veille qui alignera dans le top 20 de ses résultats de recherche plus de pages et sites disponibles dans la base données Sibel.

En effet, la citation de Dalbin Sylvie et cette grille d'évaluation, nous les empruntons comme technique permettant d'évaluer les outils de veille sélectionnés.

Le tableau ci-après donne une idée globale sur les résultats de recherche de chaque outil de veille par rapport au positionnement des sites et pages web Sibel. Les tableaux complets des résultats de recherche sont donnés en annexe.

2.4. Résultat de recherche des outils de veille

Outils de veille	Nbre/requêtes	Nbre/réponses	Nbre/repenses sibel/outil de veille	Positionnement				
				Entre 5 ^{ème} réf.	Entre 6 ^{ème} et 10 ^{ème}	Entre 11 ^{ème} et 15 ^{ème} réf.	Entre 16 ^{ème} et 20 ^{ème} réf.	Entre 21 ^{ème} et plus
Copernic	16	594	53	28	1	8	7	9
Bullseye	16	354	27	22	2	0	2	1

Pour mieux apprécier les outils de veille utilisés, la notion de pertinence et la grille ci-dessus nous les considérons comme étant des instrument importants permettant d'évaluer la performance des outils de veille bien qu'il existe aussi d'autres moyens pour le faire.

Le tableau ci-dessus permet de dire que sur 16 requêtes choisis, les outils de veille sélectionnés, classés par ordre d'importance, ont donné les résultats suivant :

2.4.1 Copernic

Au sujet de Copernic, le tableau renseigne que les 16 requêtes ont fournies 53 résultats dont 28 sont alignés dans les 5 premières références, un seul entre la 6^{ème} et la 10^{ème} références, 8 résultats placés dans l'espace situé entre la 11^{ème} et la 15^{ème} références, 7 résultats sont entre la 16^{ème} et la 20^{ème} réf. et en fin 9 résultats considérés non pertinents sont enregistrés à partir de la 21^{ème} référence. C'est pour dire que sur les 53 résultats, 44 sont pertinents contre 9 qui ne les sont pas du fait qu'ils sont hors le Top 20 de Copernic. On la dit, la notion de la pertinence est généralement définie par le croisement de deux notions relatives au bruit (documents retrouvés non pertinents) et le silence (document pertinents non retrouvés) nous pouvons dire que sur 594 résultats obtenus grâce aux 16 requêtes utilisés, le taux de bruit est considérable soit $594 - 44 = 550$ documents non pertinents (bruit) ont été enregistrés dans Copernic. On le sait, le bruit et le silence sont mesurés par le taux de rappel et de précision. Dans ce cas, sur 221 documents pertinents disponibles dans la base Sibel correspondant aux 16 requêtes utilisées,

Copernic n'a trouvés que 44 documents pertinents. Ainsi, le taux de rappel est de 22% et le silence de 78%.

Donc Copernic n'a pas aligné même la moitié des documents pertinents de Sibel dans son top 20 et ce, en rapport avec les requêtes utilisées.

S'il faut appliquer la notion de la grille d'évaluation on retiendra le calcul suivant pour Copernic.

Nbre résultats obtenus	Positionnement	Nombre des points à multiplier	Total
28	5 premières références	20	560
1	Entre la 6 ^{ème} et la 10 ^{ème}	15	15
8	Entre la 11 ^{ème} et la 15 ^{ème} références	10	80
7	Entre la 16 ^{ème} et la 20 ^{ème} références	5	35

Ce tableau donne le total des points obtenus par Copernic. Il ressort que dans les 5^{ème} références Copernic a aligné 28 documents contre 1 document enregistré (entre la 6^{ème} et la 10^{ème}), 8 documents entre la 11^{ème} et la 15^{ème} références et en fin 7 documents à partir de la 16^{ème} et 20^{ème} référence. Comme vous pouvez le constater, dans ce tableau, nous n'avons pas fait allusion aux documents enregistrés dans l'espace situé entre la 21^{ème} et plus étant donné que ces documents sont d'office considérés non pertinents (cfr tableau en annexe)

2.4.2 BullsEye

A propos de Bullseye, Les 16 requêtes ont donné lieu à 27 résultats dont 22 dans les 5 premières références contre 2 entre la 6^{ème} et la 10^{ème} réf., aucun résultat (0) entre la 11^{ème} et la 15^{ème} réf., 2 entre la 16^{ème} et la 20^{ème} réf. et un seul au delà de la 21^{ème} réf.

De ce qui précède, sur 221 documents pertinents de la base Sibel répondant aux 16 requêtes construites, le taux de rappel est de plus ou moins 16% et, le silence de 84%.

En ce qui concerne la notion de la grille d'évaluation, pour BullsEye on peut retenir ce qui suit :

Nbre résultats obtenus	Positionnement	Nombre des points à multiplier	Total
22	5 premières références	20	440
2	Entre la 6 ^{ème} et la 10 ^{ème}	15	30
0	Entre la 11 ^{ème} et la 15 ^{ème} références	10	0
2	Entre la 16 ^{ème} et la 20 ^{ème} références	5	10

En effet, les deux tableaux d'évaluation susmentionnés renseignent que Copernic a renvoyé plus de résultats pertinents (44 résultats pour 22% de taux de rappel) par rapport à BullsEye (26 résultats pour 16% de taux de rappel). De ce fait, nous pouvons dire également que Copernic dans son top 20 aligne plus de résultats pertinents de la base de données Sibel que BullsEye. L'écart de documents pertinents de Sibel entre Copernic et BullsEye est de 17 documents en faveur de Copernic soit $44 \text{ (Copernic)} - 27 \text{ (BullsEye)} = 17$. Cet écart est quand même considérable.

La question qui reste à poser est celle de savoir si Copernic interroge simultanément les meilleurs moteurs de recherche que BullsEye. Le point relatif à l'étude comparative nous en dira plus.

Le point ci-après tente de donner une brève présentation des outils de veille interrogés.

2.5. Les outils de veille sélectionnés

2.5.1 Copernic

Il y a 3 versions de Copernic 2000 (Copernic 2000, Copernic 2000 plus et Copernic 2000 Pro). La version de base gratuite de Copernic est disponible sur le marché sur site www.copernic.com.

Copernic est un outil permettant d'effectuer une veille sur un thème donné. Les recherches sur ce thème sont lancées simultanément sur plusieurs moteurs de recherche et annuaires. Les informations récoltées sont classées par ordre de pertinence, les doublons et liens périmés sont automatiquement éliminés. Pour une consultation hors line, il est possible de télécharger une partie ou l'intégralité des documents trouvés. Il est également possible de relancer périodiquement la recherche et de n'afficher que les nouvelles données. La figure ci-dessous présente l'interface de recherche de Copernic

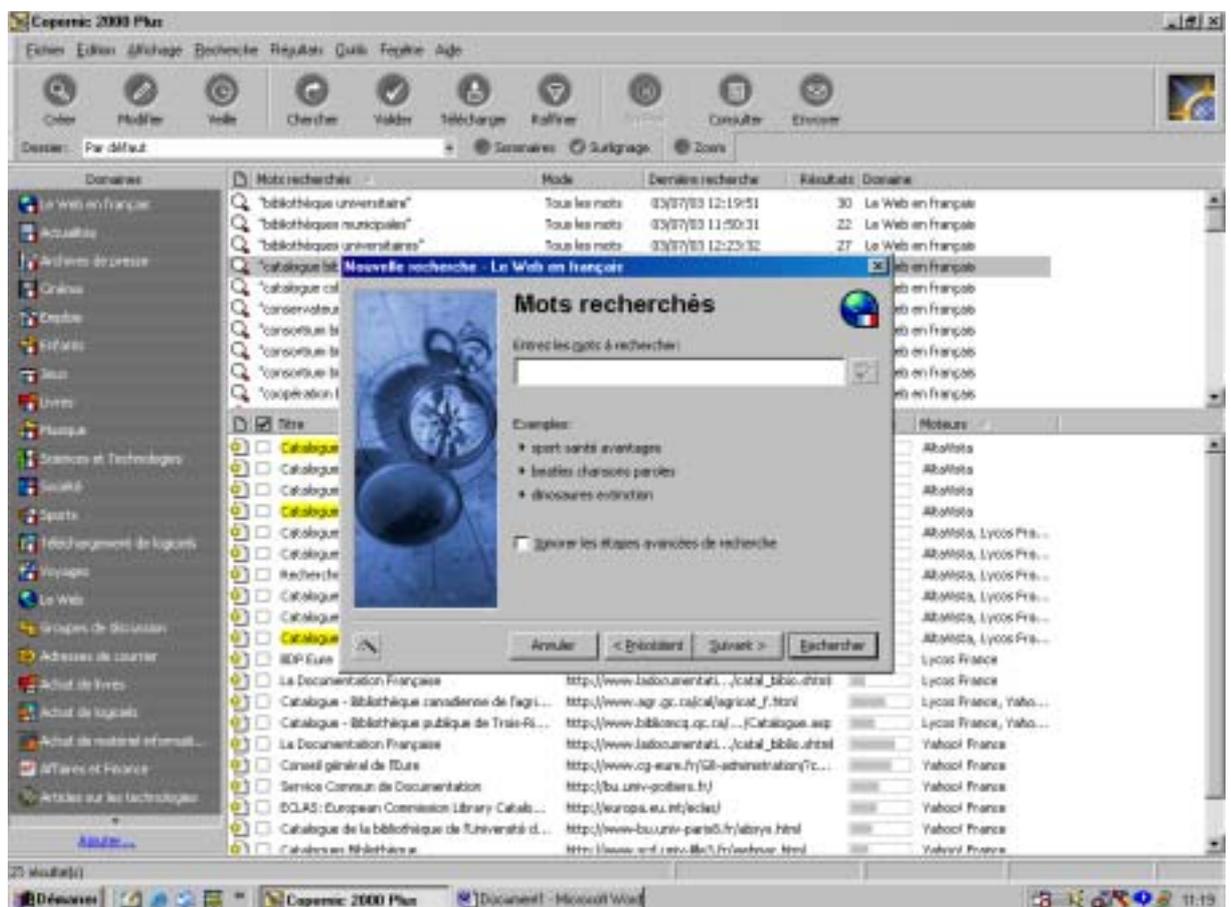


figure 4 : formulaire de recherche et résultats

2.5.2 BullsEye

BullsEye est, selon les spécialistes, le meilleur logiciel du marché. Une version basique gratuite (BullsEye 2) est disponible sur le site : www.intelliseek.com

Ce logiciel permet d'interroger un grand nombre de moteurs de recherche. Il permet également d'interroger des bases de données. Dans la version BullsEye Pro 2, les recherches peuvent faire l'objet d'une veille systématique grâce à des mises à jour automatisées.

La figure ci-après permet de découvrir le formulaire d'interrogation de BullsEye.



Figure 5 : formulaire de recherche

Après cette brève présentation, en annexe, nous comptons donner tous les détails relatifs aux deux outils de veille sélectionnés (portrait de Copernic et de BullsEye).

Il est vrai que sur le Net, les métamoteurs sont en nombre croissant. Parmi les métamoteurs sélectionnés au départ, dans le cadre de la mise en place d'une

méthodologie de veille pour Sibel, nous avons constaté que certains ne sont que des métarecherches (vivisimo et ixquick) donc des métamoteurs qui ne sont pas dotés des fonctions de veille tel que développées plus haut, et que d'autres le sont mais exigent un éditeur (navigateur) pour pouvoir consulter les résultats de recherche. (c'est le cas de strategic finder 2.0). De plus d'autres encore ne rendent pas disponible sur le Net la version de démonstration qu'on pourra télécharger et essayer pendant ne serait-ce que 15 jours. C'est ainsi que notre choix s'est porté sur Copernic et BullsEye étant donné qu'ils mettent à disposition du grand public la version démonstration téléchargeable pour essai.

En effet, nous avons choisi Copernic parce que la version démonstration est disponible sur le Net. Et, présente une facilité d'utilisation, malgré sa puissance, et une interface intuitive. En outre, selon nos investigations, Copernic reste un logiciel stable et le plus utilisé sur le marché (6,7 millions d'utilisateurs dans le monde) et donne de bons résultats.

BullsEye, a été choisi, non seulement parce qu'il est facile à utiliser, mais c'est par le fait que le service informatique de l'enssib dispose déjà de ce logiciel qu'il avait installé sur l'ordinateur que nous avons utilisé durant la période de stage. Aussi, s'il faut en croire les spécialistes, BullsEye est parmi les meilleurs outils de veille qu'on rencontre sur le marché. Est-ce l'étude comparative le confirmera ?

Dans le point qui suit, nous tenterons de faire une étude comparative de Copernic et de BullsEye.

2.6. Etude comparative des outils de veille sélectionnés

La mission de base des outils de veille est de fournir l'information aux utilisateurs au moyen des trois opérations qui sont : la collecte, le traitement et la diffusion. BullsEye et Copernic sont deux outils de veille qui remplissent cette mission.

Il est important de rappeler ici que les outils de veille (métamoteurs) interrogent en parallèle plusieurs moteurs et outils de recherche. C'est pour cette raison que BullsEye et Copernic sont aussi des outils intermédiaires, traduisant les questions des utilisateurs dans la syntaxe propre à chaque moteur. Ainsi, pour les deux outils sus indiqués, l'utilisation courante ci-dessous a été enregistrée :

- choix des moteurs de recherche à interroger (1000 pour les deux)
- choix du nombre de réponses par moteur
- saisie et envoi de la requête

De ce fait, les intérêts des outils de veille (métamoteurs) sont les suivants :

- une seule requête pour plusieurs outils
- utilisation enrichie des outils de recherche
- permettent un défrichage, un panorama des ressources sur un thème.

Eu égard de ce qui précède, BullsEye est l'un des principaux concurrents de Copernic. Car ces deux outils sont dotés de fonctions de veille et presque des mêmes caractéristiques qui donnent satisfaction à leurs utilisateurs (cfr tableau comparatif en annexe). De plus, ils (Copernic et BullsEye) explorent les moteurs de recherche et répertoires les plus prestigieux.

De ces deux versions (BullsEye Pro et Copernic Pro), les évolutions techniques ci-après, permettent de tirer des leçons suivantes :

BullsEye Pro est accompagné d'un module de «Tracking» grâce auquel il est possible de surveiller des pages web en étant alerté par mail de chaque nouveau changement. Et, Copernic pro de son côté permet de résumer instantanément tout document trouvé et alerte également par mail chaque nouveau changement et ce, grâce à son module « summarize ».

Selon les spécialistes ; les fonctionnalités de BullsEye en terme de surveillance étaient plutôt supérieures à celles de son concurrent Copernic jusqu'à la sortie de Copernic Agent Pro au mois d'octobre 2002. Apparemment, le perfectionnement et la toute nouvelle interface de Copernic Agent a été fatale à BullsEye.

De plus, étant donné que Intelliseek a décidé de discontinuer la vente de son logiciel BullsEye, Copernic offre une mise à niveau concurrentielle. Cet événement a renforcé le positionnement de Copernic comme chef de file dans ce marché. C'est ainsi que le produit a été complètement redessiné et lancé sous le nouveau nom de Copernic Agent. L'ajout de nouvelles fonctionnalités et de technologies centrées sur les besoins des professionnels et des entreprises visait à en faire un réel assistant de recherche qui commence là où les moteurs de

recherche traditionnels s'arrêtent. De ce fait, le succès obtenu par Copernic Agent n'est sûrement pas étranger au destin de BullsEye dont la dernière version datait de presque 2ans. Intelliseek , non seulement a décidé de discontinuer son produit phare BullsEye mais est allé plus loin en mettant fin au soutien technique de celui-ci.

Une telle décision donne plus d'avantage à Copernic en ce qui concerne le choix des outils de veille permettant de mettre en place un dispositif de veille pour ce qui concerne Sibel.

De ce qui précède, il y a lieu de proposer aux responsables de l'annuaire Sibel, le logiciel Xenu's Link Sleuth pour la vérification de la validité des liens en vue de la maintenance (entretien) du site et Copernic pour pouvoir faciliter la gestion de l'annuaire Sibel et ce, en mettant en place un dispositif de veille (collecte, traitement et diffusion de l'information).

Cette dernière partie donne le bilan du travail réalisé au département des ressources documentaires.

Partie 5 : Bilan

Le stage au sein du département des ressources documentaires a de façon générale été pour moi une enrichissante expérience professionnelle. Etant donné que ce stage m'a aidé d'acquérir de nombreuses connaissances et compétences relatives à la veille et d'appliquer quelques notions apprises en DESS- réseaux d'information et de document électronique.

De ce qui précède, je veux devoir tenter ici de dresser un bilan relatif aux nombreux points positifs et les très peu nombreuses frustrations constatées lors de ce stage.

1. Frustrations

Nous n'avons pas pu tester BullsEye Pro et Copernic Pro étant donné que l'organisme d'accueil ne dispose que de la version nom professionnelle pour BullsEye et pour Copernic, nous avons téléchargé la version de démonstration.

2. Réussites

Dans l'ensemble, nous avons pu accomplir comme il se doit notre mission susceptible de mettre en place une méthodologie de veille pour le compte de la base Sibel de la manière suivante :

- 1- Nous avons réussi à vérifier la validité de liens hypertextes (internes et externes) de la base Sibel en utilisant les logiciels de vérification des liens sélectionnés à savoir Check web et Xenu's link sleuth ;
- 2- Nous avons pu proposer un ensemble d'outils de veille (BullsEye et Copernic) permettant d'alimenter la base Sibel par de nouveaux sites répondant à sa thématique et de signaler les modifications intervenues dans le site web disponibles dans ladite base.
- 3- Nous avons apporté des réflexions relatives à la correction des liens défectueux enregistrés et à la mise en place d'une méthodologie de veille

Par ailleurs, il est tout de même important, dans les lignes qui suivent, de dire un mot aussi bien sur ce que l'organisme d'accueil nous a apporté que sur ce que nous avons apporté à l'organisme d'accueil.

2.1. Ce que l'organisme d'accueil m'a apporté

- 1- Tout au long de ce stage, nous avons été confronté à de nombreuses compétences et des connaissances. De prime abord, il s'agit de compétence des différentes personnes de l'équipe du département des ressources documentaires. Ceci se réfère à la disponibilité de l'équipe de la bibliothèque pour nous expliquer le fonctionnement de Sibel et d'autres aspects relatifs au fonctionnement des outils de veille.

2- Lors du travail sur la vérification de la validité des liens, nous avons pu appréhender les logiciels de vérification de liens Check Web Links et Xenu's Link sleuth, les utiliser en vue de signaler les codes erreurs Internet enregistrés. Cette tâche, grâce à ce stage, nous a donné une connaissance sur les logiciels de vérification de liens et les codes erreurs Internet. Aussi, l'utilisation des outils de veille en effectuant des recherches, en sélectionnant les réponses pertinentes par rapport à Sibel.

2.2. Ce que j'ai apporté à l'organisme d'accueil

- 1- Suite aux résultats de vérification de la validité des liens hypertextes de Sibel enregistrés, les responsables de cette base sont informés qu'il y a des liens brisés et, il peut en avoir davantage. C'est ainsi que nous croyons avoir apporté un esprit nouveau relatif à la vérification, de façon régulière, de la validité des liens internes et externes de cette base.
- 2- Pour avoir proposé les deux outils de veille (Bullseye et Copernic) et une fois mise en place, on pourra dire désormais la base Sibel sera alimentée automatiquement par des nouveaux sites et les modifications qui y seront constatées dans les sites existants seront signalées. Ainsi les spécialistes de l'information seront toujours à la pointe des informations les intéressant.
- 3- Une fois que les propositions sont observées, nous croyons, premièrement, que ce dispositif permettra aux responsables de l'annuaire Sibel de gagner du temps étant donné que leur charge de travail est lourde. Et, deuxièmement, l'utilisation de logiciel de vérification des liens donnera lieu à un entretien efficace de Sibel.

La conclusion nous en dira davantage sur le déroulement de notre stage dans le département des ressources documentaires dont la mission, comme on le sait, consistait, en résumé, à vérifier la validité des liens et de proposer un ensemble d'outils de veille pour le compte de Sibel.

Conclusion

Au terme de ce stage, il est tout de même important de rappeler que deux missions nous a été confiées notamment :

- la vérification de la validité des liens de Sibel en vue de découvrir les liens défectueux bref connaître l'état actuel de la base en faisant l'étude de l'existant relative aux liens hypertextes internes et externes et d'en faire de propositions relatives à la réparation des liens défectueux enregistrés.
- la proposition d'un ensemble d'outils de veille permettant d'alimenter Sibel par des nouveaux sites web correspondant à sa thématique et de signaler (par e-mail « technique d'alerte ») les modifications enregistrées dans les sites déjà en ligne dans cet annuaire.

De ce fait, les 4 mois de stage (juin à septembre) à l'Enssib plus précisément au département des ressources documentaires (bibliothèque de l'enssib) a été riche en apprentissage a telle enseigne qu'il était indispensable pour nous de mettre à profit les connaissances apprises au cours de notre formation en DESS - Réseaux d'information et de document électronique en vue de remplir comme il se doit notre mission.

Tout compte fait, lorsqu'une base de donnée prend de l'ampleur le risque de découvrir soit l'un des codes erreurs n'est plus à démontrer. Assez souvent ces types des codes erreurs, non seulement, décourage les internautes, mais peuvent diminuer à juste titre la considération et la qualité du site. Voilà pourquoi, grâce à ce stage, nous avons pu acquérir une connaissance assez approfondie relative à l'utilisation des logiciels de vérification de liens disponibles gratuitement sur Internet. De plus, à partir des résultats de vérification des liens obtenus, nous étions amené à maîtriser les codes erreurs et leurs signification. La connaissance de ces codes erreurs et leurs signification étaient d'une nécessité obligeante en vue de mieux analyser les résultats de vérification des liens et d'en faire de proposition.

Comme on le sait, dans le monde de l'information, la veille devient de plus en plus une activité qui ne laisse plus indifférent les spécialistes de l'information notamment les bibliothécaires, documentalistes et les veilleurs eux même. En effet, l'idée de surveiller l'environnement pour des raisons diverses a fait que les outils de veille, à l'heure actuelle, soient en nombre croissant sur le net.

Dans notre démarche de proposer les outils de veille, il fallait tester beaucoup des métamoteurs c'est ainsi que nous avons également approfondie notre connaissance en ce qui concerne l'utilisation de ces types d'outils de recherche dont la particularité est d'interroger simultanément d'autres sources en nombre considérable (moteurs de recherche, annuaire, répertoire, web invisible etc., forum).

De ce qui précède, pour interroger les outils de veille sélectionnés, le choix de requêtes a semblé être prioritaire d'où la nécessité de respecter une certaine démarche relative au choix de mots – clés caractérisant Sibel. Cette tâche n'était pas facile étant donné qu'il fallait explorer ce site en vue de choisir les mots-clés pouvant mettre en valeur cette base de données. Ainsi, 16 requêtes ont été sélectionnés pour interroger et évaluer les deux outils de veille sélectionnés en mettant en avant la notion de la pertinence telle que nous l'avons voulue pour Sibel. Et une grille d'évaluation a été construite par nous même pour pouvoir noter les outils de veille sélectionnés.

Après avoir testé et évalué les logiciels de vérification des liens et les outils de veille, nous sommes arrivés à la conclusion selon laquelle, les logiciels ci-après sont proposés aux responsables de Sibel :

- Xenu's Link Sleuth pour la vérification de la validité des liens)
- Copernic pour alimenter l'annuaire Sibel par des nouvelles pages web et de signaler des changements).

Les deux logiciels sus indiqués pourront contribuer efficacement à la gestion de l'annuaire Sibel en vue d'offrir aux internautes un outil de travail susceptible de les renseigner sur la bibliothéconomie et les sciences de l'information.

Bibliographie

1. Ouvrages

ABOU Olivier, ANDRIEU Olivier. Imaginer son site web : toutes les questions à se poser avant de mettre en place un site. Les Ulis, Microsoft Press, 2001, 384 p.

HERRMEL Laurent. Maîtriser et pratiquer la veille stratégique. Paris, AFNOR, 2001, 95p.

JAKOBIAK F. Pratique de la veille technologique, Paris, Les Editions d'organisation, 1991

LESCA Humbert. Veille stratégique : l'intelligence de l'entreprise. Lyon, Astier, 1994, 146 p.

MARTINET B., REBAULT J.M. La veille technologique, concurrentielle et commerciale, 2^{ème} éd., Paris, Les Editions d'organisation, Paris, 1992.

REVELLI Carlo. Intelligence stratégique sur Internet : comment développer efficacement des activités de veille et de recherche sur les réseaux. Moteurs de recherche, réseaux d'experts, agents intelligents. Paris, Dunod, 1998, 212 p.

SAMIER Henry, SANDOVAL Victor. La veille stratégique sur l'internet. paris, Lavoisier, 2002, 191p.

2. Rapports et mémoires

ARAYA-YAHANNES BEKELE. Sibel, sciences de l'information et des bibliothèques en ligne. Rapport de stage. Lyon, Enssib, 2002, 54 p.

BAYON Nicolas. Mise en place d'un portail de veille sur internet : rapport de stage. Lyon, Univ. Lyon 1, 2000, 64 p.

DELISLE Cynthia. Les outils de recherche sur internet : typologie et principales caractéristiques. Lyon, Enssib, 2000, 54 p.

ENSSIB. Département des ressources documentaires : rapport. Lyon, Enssib, 2001

GERARD Anouk. Bibliothèque de l'école nationale supérieure des sciences de l'information et des bibliothèques (Enssib) : rapport de stage. Lyon, 2003, 34 p.

HENRY Véronique. Outils de veille stratégique : essai de typologie. Lyon, Enssib, 1998, 35 p.

3. **Articles et revues**

BERNIER C. A la recherche du lien perdu. Revue Québécoise des professionnels de l'information documentaire, vol 31 n°3, 2002

CONTAL Philippe. Fiabiliser les sources du web. Archimag n° 164, mai 2003

STILLER Henri. Veille sur Internet : analyse et synthèse de l'information. Archimag n° 165, Juin 2003

4. **Ressources Internet**

Agentland. BullsEye Pro, atteignez votre cible, 2003. Disponible sur
<<http://www.agentland.fr/pages/leam/articles/bullseye.html>>
(consulté le 20.07.03)

BOUTHOUMIEN L., CONDOL D., MALARD S. La veille, 1997. Disponible sur
<http://www.urfist.cict.fr/veille.html> (consulté le 15.08.03)

Copernic technologies inc. Famille de produits Copernic, 2003. Disponible sur
<<http://www.copernic.com/fr/produits/agent/index.html> > (consulté le 18.07.03)

ENSSIB. <http://www.enssib.fr>. (consulté régulièrement)

Frameset. Xenu's Link Sleuth, 2001. Disponible sur
<<http://www.agentland.fr>> (consulté le 12.07.03)

PC Astuces. Check Web Links, 2001. Disponible sur
<<http://www.pcestuces.com/logitheque/checkweblinks.htm> >(consulté le 13.7.03)

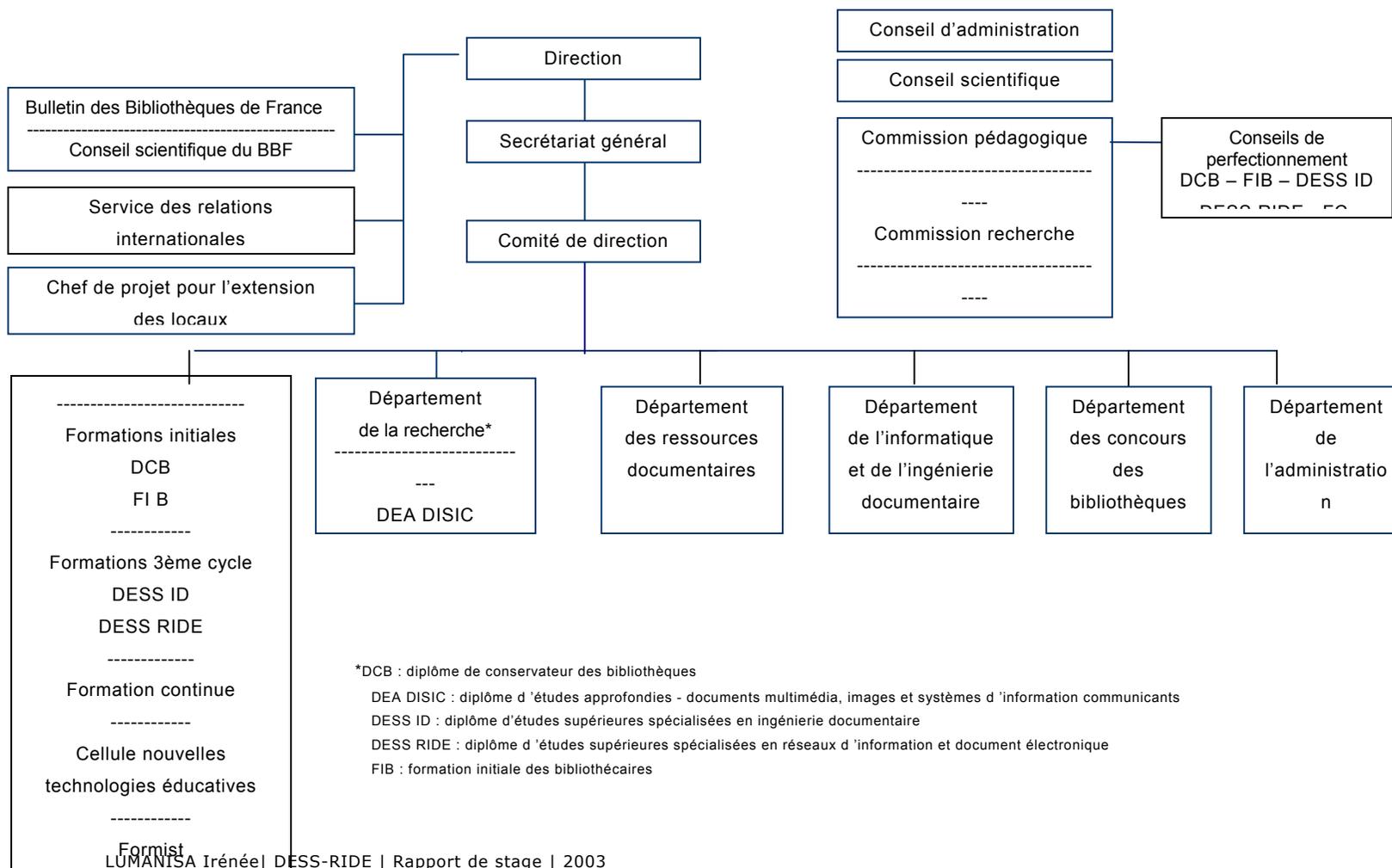
SIBEL. <http://sibel.enssib.fr>. (consulté régulièrement)

Veille.com. Concepts de veille, 2002. Disponible sur
<http://www.veille.com/> (consulté le 12.07.03)

Table des annexes

1. LE PERSONNEL, RESSOURCES MATÉRIELLES ET FINANCIÈRES DU DÉPARTEMENT DES RESSOURCES DOCUMENTAIRES	III
1.1. LE PERSONNEL , RESSOURCES MATÉRIELLES ET FINANCIÈRES DU DÉPARTEMENT DES RESSOURCES DOCUMENTAIRES	III
2. RESSOURCES MATÉRIELLES, DOCUMENTAIRES ET FINANCIÈRES	III
2.1. SALLES DE LECTURE, ÉQUIPEMENT ACTUEL	III
2.2. COLLECTIONS : TYPOLOGIE DES RESSOURCES DOCUMENTAIRES	IV
2.3. RESSOURCES FINANCIÈRES	V
3. RÉPARTITION DES REQUÊTES PAR NOMBRE DES RÉPONSES PERTINENTES SIBEL	VIII
4. TABLEAU DES PRINCIPAUX CODES ERREURS INTERNET	IX
5. PORTRAIT BULLSEYE ET COPERNIC.....	X
5.1. BULLSEYE	X
5.2. MODE D'EMPLOI.....	XI
5.3. FONCTIONNALITÉS	XII
5.4. TABLEAU DE RÉSULTATS D'INTERROGATION DE BULLSEYE	XIII
6. PORTRAIT COPERNIC.....	XIV
6.1. MODE D'EMPLOI.....	XV
6.2. PRINCIPALES CARACTÉRISTIQUES DE FONCTIONNEMENT	XVI
6.3. RÉSULTATS D'INTERROGATION DE COPERNIC	XVII
LE TABLEAU CI-APRÈS TENTE UNE ÉTUDE COMPARATIVE DES DEUX OUTILS DE VEILLE SÉLECTIONNÉS (BULLSEYE ET COPERNIC)	XVIII
7. TABLEAU COMPARATIVE DES OUTILS DE VEILLE.....	XIX

Organigramme de l'enssib



1. Le personnel, ressources matérielles et financières du département des ressources documentaires

1.1. Le personnel, ressources matérielles et financières du département des ressources documentaires

Actuellement, 18 personnes travaillent dans le centre département des Ressources documentaires.

Un conservateur	
Bibliothèque de La Doua : <ul style="list-style-type: none"> - deux bibliothécaires - deux bibliothécaires adjoints spécialisés - un assistant de bibliothèque - une assistante de communication - 3 magasiniers (3 mi-temps) 	Bibliothèque de Grandclément : <ul style="list-style-type: none"> - une bibliothécaire - un bibliothécaire adjoint spécialisé - une documentaliste - un magasinier
Service de l'édition : cinq personnes (4 temps plein et 1 mi-temps)	

2. Ressources matérielles, documentaires et financières

2.1. Salles de lecture, équipement actuel ²

² Source : Rapport du Département des Ressources documentaires de l'ENSSIB, 2000-2001.

Site de La Doua	Site de Grandclément
Une salle de lecture de 50 places et une salle de références de 15 places assises.	Une salle de lecture de 40 places assises
4 PC pour la consultation du catalogue Portfolio (v. 5), 2 PC multifonctions en libre accès, réseau de cédéroms ARCHIMED commun aux deux sites (installation septembre 2000), poste multimédia, Internet, lecteurs de microfiches, Minitel, photocopieuses	2 PC permettant la consultation du catalogue Portfolio (v. 5), réseau de cédéroms ARCHIMED commun aux deux sites (installation septembre 2000), poste multimédia, accès Internet, lecteur de microfiches, Minitel, photocopieuse

2.2. Collections : typologie des ressources documentaires

Site de La Doua	Site de Grandclément
La plupart des ouvrages et les périodiques de l'année en cours sont en libre accès. --25 000 livres --286 titres de périodiques vivants --482 titres morts --cédéroms, rapports, mémoires, thèses, dossiers documentaires.	Toutes les collections sont en libre accès. --5000 livres --130 titres de périodiques vivants --cédéroms, travaux des élèves-bibliothécaires, dossiers documentaires.
Collections spécialisées en bibliothéconomie, informatique documentaire, sciences de l'information et histoire du livre.	Collections particulièrement développées en bibliothéconomie, bibliographie générale, usages des publics, formation, administration.
Classification DEWEY	Classification DEWEY
Les notices sont informatisées à 90% et accessibles en ligne.	Les notices sont informatisées à 90%, la base n'est consultable que sur place.
Acquisitions 2001 : --792 livres (568 français, 224 étrangers) --31 nouveaux périodiques (9 français, 22 étrangers)	

2.3. Les ressources électroniques

En plus des ouvrages (imprimés) composés de monographies, périodiques, usuels, travaux d'étudiants et de chercheurs, etc., l'ENSSIB possède un bon nombre de ressources électroniques :

- Cédéroms (à prêter, ou à consulter sur place sur les postes de la bibliothèque, ou à consulter à partir des postes du réseau de l'ENSSIB)
- Des travaux des étudiants et des chercheurs (consultables ou non consultables, plus de 300 disponibles en version pdf)

- L'OPAC de la bibliothèque (catalogue comprenant les notices des monographies, des ouvrages de références, des travaux universitaires, des CD ROM).
- Le dossier de presse de l'ENSSIB (PERIS' COP)
- Service des sommaires en SIC (86 titres de revues francophones et étrangères en SIC figurant au catalogue.) [abandonné depuis juin 2002 pour être peu à peu remplacé par le dossier de presse]
- Dossiers thématiques
 - Censure et bibliothèque
 - Économie du livre ancien
 - Métiers des bibliothèques (mémento des textes paru au Journal Officiel)
 - Normes et formats
- Les signets de la bibliothèque
 - Bases de données en ligne (une quarantaine)
 - Revues francophones (en SIC) (par ordre alphabétique)
 - Journaux électroniques (30 en anglais, 1 en Espagnol)
 - Les synthèses de biblio-fr et adbs-info (accès par thèmes des synthèses des listes de discussion; liens vers les archives) (34)
 - Un annuaire de ressources en ligne

2.4. Ressources financières

- Une bonne partie du budget du département des ressources documentaires de l'ENSSIB est utilisée pour le projet SIGB.

Affectation	Montant
Investissement SIGB	600 KF
Fonctionnement	
Prestations SIGB (reprise des données)	400 KF
Acquisitions La Doua :	
Monographies et cédéroms	320 KF
Périodiques	320 KF
Acquisitions Grandclément :	

Monographies et cédéroms	90 KF
Périodiques	90 KF
Salaires	300 KF
Divers (frais de déplacements, fournitures diverses, maintenance du SIGB, bases de données...)	180 KF
TOTAL	2 300 KF (350 632 .73 €)

Budget

BUDGET 2002 - DEPARTEMENT DES RESSOURCES DOCUMENTAIRES - BIBLIOTHEQUES ET WEB

Compte	Destination	Objet	Détail ou commentaire	FRANCS	EUROS	Qui gère
60	FOURN	Matériel divers	Tout, Microfilm, Fimulac, BMF, Protocol, CompuType... petit matériel et traitement des collections - Tous commandes de papeterie hors matériel professionnel doit passer par Raymond	10 000,00 F	1 524,45 €	Ghislaine
61	DOLIV	Ouvrages français ou étrangers	ouvr. suites (WEKA environ 20 000 F)	177 000,00 F	26 983,48 €	Danielle
	DOPER	Abonnements français ou étrangers	ouvr. étrangers (environ 180 000 F)	320 000,00 F	48 783,56 €	Gabrielle
	DONLI	Ouvrages électroniques français ou étrangers	Ensemblement collections. Acquisitions centralisées à La Doua. Cette destination est fusionnée avec BODON	0,00 €	0,00 €	
	GCLIV	Ouvrages français ou étrangers	WEKA 7	60 046,00 F	9 153,95 €	Gabrielle
	GOPER	Abonnements français ou étrangers		60 000,00 F	9 146,94 €	Gabrielle
	AGENE	Congrès		0,00 F	0,00 €	
62	COTIS	Cotisations diverses	Inscriptions à faire prendre par Valérie Aymard			
	BODON	Abonnement à des bases de données	ABF, GUTTENBERG, CERL, MEDIAT, LORIS, LIBER, AZER... : nous proposons de garder le CERL, que nous finançons depuis plusieurs années déjà. Nous avons de moins en moins de collections (à terme, ils vont disparaître) et de plus en plus d'abonnements à des bases de données en ligne.	70 000,00 F	10 671,43 €	Danielle
	PEB	PEB	Toutes lectures de PCB	4 951,00 F	754,78 €	Pierre
	AGENE	Voyages	Transport + hôtel	20 000,00 F	3 048,98 €	Janine
	AGENE	Frais de réception	Repas, transport pour un invité	0,00 €	0,00 €	
	RELIU	Prestations de reliure	Reliure des ouvrages abimés pour conservation	10 000,00 F	1 524,49 €	Janine
	WEB	Prestations graphiques	Mise à la charte graphique du site sur 2002, planifiée pour juin 2002, nouveau site avec nouveau visuel pour le prochain Doc Forum de décembre 2002, version française et version anglaise	30 000,00 F	4 573,47 €	Anne
	SITT	Système de gestion de bibliothèque, partie prestations et gestion des données	Selon devis Jouve sur projet SIGS 2001 : 492 752 F. Poursuite du projet sur 2002 avec prestations liées à l'acquisition d'un système de GEIDE : 198 000 F	520 125,00 F	79 292,55 €	Anne
64	SALAI	Contractuels	Emploi Jouve + Florence Favier + contractuel de janvier à juillet en remplacement de Pierre Laroque (Laura Ricci)	350 000,00 F	53 357,16 €	Anne
205	SIOLOG	Système de gestion de bibliothèque, partie logicielle	Selon devis InterMondo sur projet SIGS 2001 : 1 112 236 F Poursuite du projet sur 2002 avec acquisition d'un système de GEIDE : 295 000 F	1 000 000,00 F	152 449,02 €	Anne
TOTAL				2 732 122,00 F	416 509,31 €	

Note : le projet système d'information (destination SITT et SIOG) représente 2 000 000 F au total sur 2001 et 2002

3. Répartition des requêtes par nombre des réponses pertinentes Sibel

N°	Requêtes	Nbre rép. Sibel
1	« association archivistes »	3
2	« association bibliothécaires »	9
3	« association conservateur »	1
4	« bibliothèque interuniversitaire »	9
5	« bibliothèque municipale »	30
6	« bibliothèque nationale »	47
7	« bibliothèque numérique »	11
8	« bibliothèque publique »	24
9	« bibliothèque universitaire »	31
10	« « catalogue collectif »	11
11	« consortium bibliothèque »	7
12	« coopération bibliothèque »	11
13	« école science information »	14
14	« réseau national bibliothèque »	7
15	« service commun documentation lyon	6
16	« professionnel information »	20
Total	16 références	221 rép. Sibel

Les tableaux de l'interrogation des outils de veille sont présentés à la fin du portrait relatif à chacun des outils.

Le tableau suivant donne les codes erreurs Internet qu'on peut rencontrer lors de la navigation

4. Tableau des principaux codes erreurs Internet

Codes erreurs	Description
301	Document déplacé de façon permanente. <i>Rapprochez-vous de la racine du site et effectuez une recherche</i>
302	Document déplacé de façon temporaire. <i>Réessayez ultérieurement</i>
400	Erreur de syntaxe dans l'adresse du document. Un message « impossible d'afficher la page » <i>Revoyez l'adresse</i>
401	Pas d'autorisation d'accès au document. <i>Demandez au webmaster du site le mot de passe</i>
402	Accès au document soumis au paiement. <i>Il va certainement vous falloir sortir votre carte bleue</i>
403	Pas d'autorisation d'accès au serveur. <i>Demandez à l'administrateur les autorisations nécessaires</i>
404	La page demandée n'existe pas. C'est l'erreur la plus courante. <i>Rapprochez vous de la racine du site et effectuez une recherche de la page. Vous pouvez aussi essayez de remplacer les majuscules par des minuscules dans le nom de la page. Pour en savoir plus</i>
405	Méthode de requête du formulaire non autorisée. <i>Il s'agit probablement d'une erreur de l'auteur du site. Il se peut aussi que votre navigateur ne convienne pas (trop ancien ou paramétré de façon incompatible)</i>
406	Requête non acceptée par le serveur
407	Autorisation du proxy nécessaire
408	Temps d'accès à la page demandée expiré <i>Le serveur est indisponible ou encombré. Réessayez plus tard ou réactualisez la demande</i>
500	Erreur interne du serveur <i>Mauvaise paramétrage de votre navigateur ou erreur de conception du site visité</i>
501	Requête faite au serveur non supprimée
502	Mauvaise passerelle d'accès
503	Service non disponible
504	Temps d'accès à la passerelle expiré <i>Mauvaise connexion à Internet réessayez plus tard après vous être déconnectés</i>

Nous avons alignés dans ce tableau les principaux codes erreurs Internet auxquelles les internautes sont confrontés. Des indications aiderons à contourner le problème. Cette liste n'est pas exhaustive.

Les points suivants donnent les portraits des outils de veille sélectionnés. Il s'agit de BullsEye et Copernic.

5. Portrait Bullseye et Copernic

5.1. BullsEye

BullsEye Pro est un agent de recherche développé par la société IntelliSeek (<http://www.intellseek.com>) BullsEye permet de consulter en parallèle des milliers de sources (moteurs de recherche, web invisible, bases de données forum). Dans sa version 2, BullsEye existe sous deux formes : la version plus et la version pro, cette dernière possédant toutes les fonctionnalités offertes par l'éditeur est sans conteste la plus complète

Tout d'abord BullsEye permet d'interroger plus de 1000 outils de recherche, repartis par grands thèmes et catégories :

Au cœur du fonctionnement de BullsEye réside le principe conjoint à tous les métamoteurs : interrogation simultanée de plusieurs outils de recherche (annuaire, moteurs et bases de données). A cette base commune de nombreuses fonctionnalités ont été ajoutées pour faire de cet outil l'un des plus puissants de sa catégorie.

En effet, le tracker de BullsEye utilise aussi le carnet d'adresses, bookmarks de Netscape ou les favoris du navigateur Internet Explorer pour suivre les changements périodiques de pages (tracker).

BullsEye génère des rapports sur les recherches du tracker ou sur les recherches de sources sur le web. Le rapport du tracker est un bulletin d'alerte indiquant seulement le changement et la modification de pages

Ces rapports sont envoyés automatiquement à un groupe d'utilisateurs par e-mail.

Avec BullsEye, les résultats de recherche sont proposés classiquement par score ou pertinence relative à chaque source, donc peu fiables. Mais BullsEye propose une organisation par non de domaine de sites (com, edu, fr)

Cet outils fortement anglophone et très complet se rapproche des outils tes que Copernic, StrategicFinder et DogOut4 mais avec moins de sources francophones que Copernic et StrategicFinder et aussi « ouvert » à l'ajout de nouvelles sources que Digout4U et StrategicFinder.

BullsEye dispose des options plus sophistiquées ainsi sont associées : formulaire de recherche avancée (combinaison des opérateurs booléens). Dictionnaire orthographique, thésaurus, et index de termes dérivés de la racine par l'utilisateur sont proposés comme assistance à la formulation de sa requête.

De plus la possibilité de classement des résultats en catégories, signalement des pages modifiées.

5.2. Mode d'emploi

Les 1000 outils de recherche qu'interrogé BullsEye sont répartis par grand thèmes et catégories :

- web (ensemble du web)
- news (presse)
- shopping (sites d'e-commerce)
- multimédia (images, audio vidéo)
- discussions (newsgroups)
- jobs (emploi)
- books (livres)
- software (logiciels) etc...

En effet, pour lancer une recherche il suffit de sélectionner la thématique voulue, de spécifier une éventuelle sous-catégorie puis de choisir ses mots-clés. D'autres options peuvent être paramétrées dans le mode avancé comme choix des outils de recherche et de type d'analyse (sans analyse, pour une vue d'ensemble rapide – suppression des liens morts, téléchargement et analyse). Ces paramétrage permettent de configurer sa requête pour être au plus proche des résultats

consultables directement au sein de l'application, BullsEye ayant un navigateur intégré.

5.3. Fonctionnalités

Les fonctionnalités intéressantes du logiciel,

- les mots-clés sont automatiquement mis en valeur dans le document retrouvé assurant une lecture plus efficace ;
- les résultats peuvent être triés, filtrés, enregistrés et exportés.
- L'interface de BullsEye est très intuitive et conviviale. L'élaboration des requêtes peut être assistée par plusieurs outils intéressants notamment :
 - un correcteur d'orthographe (anglais)
 - un index des termes phonétiquement proches
 - une interface pour la saisie de requêtes booléennes simple :
 - il est possible de filtrer les pages en fonction d'une URL ou d'un mot dans le titre
 - il est possible d'effectuer un second filtrage sur les résultats.

Pour ce faire BullsEye affiche la liste des mots indexés. De plus d'autres fonctions à considérer sont :

- Recherches spécialisées
- Recherche sur les sites de différents pays (USA, Canada, Royaume-Uni, Inde, Australie, Mexico etc.)
- Ajout de moteurs de recherche
- Navigateur amélioré pour consulter les résultats
- Recherche en langage naturel

Avec BullsEye vous pouvez consulter et naviguer les résultats simultanément

La version Pro permet d'effectuer de la veille sur des résultats, des sites, une liste des favoris, etc. De plus, le logiciel est couplé à un système d'alerte.

Il vous permet de sauvegarder les pages web directement dans des répertoires. Il est alors possible d'effectuer ses recherches en local (off line)

BullsEye tracker pour surveiller les changements d'un site

On le voit, les fonctions de BullsEye sont pléthoriques et sont fort heureusement offertes dans une interface des plus conviviales et ergonomiques. L'utilisation et le paramétrage du logiciel sont de fait facilités, la navigation y étant très intuitive.

Le tableau suivant renseigne sur les résultats de l'interrogation de cet logiciel.

5.4. Tableau de résultats d'interrogation de BullsEye

Outil de veille	Date	requêtes	Nbre De réponse	Sibel/ réponse pertinence	Positionnement				
					5° réf.	Entre 6 - 10	Entre 11-15	Entre 16-20	Entre 21 et +
Bullseye	2-07-03	« association archivistes »	23	2	2	-	-	-	-
	2-07-03	« association bibliothécaires »	24	1	1	-	-	-	-
	2-07-03	« association conservateurs »	24	1	1	-	-	-	-
	2-07-03	« association interuniversitaire »	20	3	2	1	-	-	-
	2-07-03	« bibliothèques municipales »	23	1	1	-	-	-	-
	2-07-03	« bibliothèque nationale »	25	3	3	-	-	-	-
	2-07-03	« bibliothèque numérique »	26	1	1	-	-	-	-
	2-07-03	« bibliothèque publique »	25	1	1	-	-	-	-
	2-07-03	« bibliothèque universitaire »	19	2	1	-	-	-	-
	2-07-03	« catalogue collectif »	19	3	2	1	-	-	-
	3-07-03	« consortium bibliothèque »	19	1	1	-	-	-	-
	3-07-03	« coopération bibliothèque »	19	0	-	-	-	-	-

	3-07-03	« école sciences information »	24	1	1	-	-	-	-
	3-07-03	« réseau national bibliothèques »	20	3	3	-	-	-	-
	3-07-03	« service commun documentation lyon »	25	3	1	-	-	1	1
	3-07-03	« professionnels information »	24	1	1	-	-	-	-

Le point ci après tente de dresser le portrait de Copernic.

6. Portrait Copernic

Adresse : <http://www.copernic.com> (pour téléchargement ou achat)

Sous la responsabilité de la société Agents technologies Corp.

Copernic est issue d'une entreprise québécoise, c'est un agent intelligent de recherche sur l'Internet qui interroge simultanément des centaines des sources (moteurs de recherche, sites thématiques, groupes de discussion, adresses e-mail) organisés en domaines de recherche.

Pour chaque domaine, Copernic donne accès à une sélection de site qu'il questionnera simultanément. Les recherches s'effectuent à partir de mots-clés ou expressions (bouts de phrase) que Copernic recherchera sur les sources que l'on peut inclure ou exclure de la recherche par simple sélection

A l'issu des recherches, Copernic présente les documents trouvés par ordre de pertinence des sources après avoir éliminé automatiquement les documents en double et les liens morts. De plus il met en sur brillance dans les documents les mots – clés de la recherche.

Il convient ensuite de sélectionner les documents pertinents afin qu'il les télécharge en local avec ou sans les images et qu'il élimine les pages n'existant plus.

Un rapport contenant le titre, l'adresse et le résumé des pages peuvent être générés au format HTML afin de faciliter l'analyse. La présentation du rapport est modifiable en fonction des filtres tels que le score, le titre, l'adresse ou la date du document.

La fonction raffiner permet de rajouter des mots-clés à la recherche et raffiner la recherche uniquement sur les documents téléchargés précédemment. Pour chaque

recherche, Copernic crée une ligne dans son interface ce qui donne souvent pour la veille une cinquantaine de recherches (lignes) pour un seul thème. Pour chacune d'entre elle, la mise à jour se fera automatiquement grâce à un agenda paramétrable accessible par l'onglet veille et activé par la sélection de la mise à jour automatique. Il suffira à l'ordinateur d'être allumé pour que le logiciel se lance automatiquement et réactualise la recherche.

Les nouveaux documents apparaîtront avec une icône sous forme d'étoile alors que les documents déjà téléchargés ont une icône sous forme de disquette.

L'avantage principal de cet outil est qu'il offre des sources multilingues et permet ainsi de faire des recherches en 5 langues

6.1. Mode d'emploi

Choisir dans l'assistant l'un des 7 domaines accessibles de la version freeware :

(si l'assistant n'est pas activé il faut cliquer sur la baguette magique située en bas à gauche). Les 7 domaines sont :

- le web en français
- le web
- groupes de discussion
- adresses de courrier
- achat de livres
- achat de logiciels
- achat de matériel

(les deux premiers domaines sont les plus utiles)

De ce fait, pour faire la recherche, il faut taper les mots-clés recherchés en les séparant par un espace. Par défaut Copernic recherche les pages qui contiennent tous les mots. Si la case ignorer les étapes avancées de recherche est cochée il suffit de cliquer sur rechercher pour lancer la recherche. L'interrogation des différents moteurs peut être précédée d'une phase de mise à jour automatique du logiciel. La recherche peut prendre un certain temps, il est possible de l'arrêter à tout moment par le bouton Arrête.

La recherche est modifiée et relancée, la différence par rapport à la création se situe au niveau des sauvegardes faites par Copernic. Pour chaque recherche Copernic sauvegarde les résultats sous forme compressée lisible par lui seul. En modifiant la recherche il n'y a pas création de nouveaux fichiers internes.

6.2. Principales caractéristiques de fonctionnement

- Interroge de manière simultanée jusqu'à 1000 sources pour la version Pro
- Permet l'emploi d'opérateurs booléens lors de la formulation de la requête (ET, OU, locutions).
- Permet quatre types de requête : rapide, normal, détaillé, personnalisé. Ces types se distinguent entre eux par le mode maximum de résultats permis par le moteur et par requête (fixé par défaut respectivement à 300 et à 1000).
- Détruit les doublons et liens désuets.
- Affiche, pour chaque document rapporté :
 - a) un score de pertinence
 - b) le titre
 - c) une description
 - d) l'URL
 - e) la date
 - f) l'outil de repérage initial
 - g) l'état (accessible, inaccessible, nouveau, ignoré, visité, téléchargé, raffiné)

De ce qui précède, les résultats peuvent être :

- consultés hors ligne raffinés à l'aide d'opérateur booléens (ET, OU, SAUF, locutions)
- sauvegardés en différents formats (textes, HTML, XML, dBAsE, etc.)
- possibilité de réactiver une requête pour en mettre à jour les résultats (les nouveaux documents repérés sont alors mis en évidence).
- utilisé pour « aspirer » des sites afin de permettre une consultation hors ligne

- mis à jour hebdomadaire automatique du logiciel et des différents moteurs.

Avec Copernic il y a possibilité de programmer une veille, en indiquant à Copernic quand est-ce que vous désirez qu'il relance la requête automatiquement. Pour cela, aller dans recherche/veille (cette option est bridée dans la version de démonstration).

Cette veille peut se faire manuellement en choisissant Recherche/Mettre à jour la recherche.

Pour modifier la requête (Recherche/Modifier)

Pour poser la même requête mais dans un autre domaine, on peut dupliquer la recherche (Recherche/Dupliquer).

En effet Copernic Agent professionnel – la version la plus évoluée de la gamme de produits Copernic Agent – marque une évolution certaine par rapport à la recherche traditionnelle sur le web, en intégrant des outils puissants et innovateurs pour trouver, gérer, analyser et surveiller de l'information sur l'internet.

Le tableau suivant donne les résultats de l'interrogation de Copernic.

6.3. Résultats d'interrogation de copernic

Outil de veille	Date	requêtes	Nbre De réponse	Sibel/ réponse pertine nte	Positionnement				
					5 ^e réf.	Entre 6 - 10	Entre 11-15	Entre 16-20	Entre 21 et +
Copernic	4-07-03	« association archivistes »	9	2	2	-	-	-	-
	4-07-03	« association bibliothécaires »	11	1	1	-	-	-	-
	4-07-03	« association conservateurs »	48	1	1	-	-	-	-
	4-07-03	« association interuniversitaire »	20	4	3	-	1	-	-
	4-07-03	« bibliothèques municipales »	42	5	1	-	-	1	3

	4-07-03	« bibliothèque nationale »	50	9	4	1	1	1	2
	6-07-03	« bibliothèque numérique »	25	3	1	-	1	-	1
	6-07-03	« bibliothèque publique »	38	1	1	-	-	-	
Copernic	6-07-03	« bibliothèque universitaire »	27	5	1	-	2	2	-
	6-07-03	« catalogue collectif »	40	4	2	-	1	-	1
	6-07-03	«consortium bibliothèque »	39	3	1	-	1	1	-
	6-07-03	« coopération bibliothèque »	48	4	2	-	1	1	-
	6-07-03	« école sciences information »	58	6	3	-	-	1	2
	6-07-03	« réseau national bibliothèques »	44	2	2	-	-	-	-
	6-07-03	« service commun documentation lyon »	43	3	3	-	-	-	-
	6-07-03	« professionnels information »	72	1	1	-	-	-	-

Le tableau ci-après renseigne sur l'étude comparative des deux outils de veille sélectionnés (BullsEye et Copernic)

7. Tableau comparative des outils de veille

Outils de veille/critères	Bullseye	Copernic
Recherche avec mots-clés	Oui	Oui
Recherche avec opérateurs booléens	Oui	Oui
Recherche avec opérateurs de proximité	Oui	Oui
Multilingue	Anglais	Oui
Installation et compréhension de l'outil	Facile d'installation/ 8287 Ko/windows 95/98/NTInter net explorer	Facile d'installation Ko/windows 95/98, Microsoft windows200/ NT et XP
Support	CD-ROM	CD-ROM
Outil fini, commercialisé	Oui	Oui
Critère de classement des résultats	Par pertinence	Par pertinence
Elimination des doublons	Oui	Oui
Présentation des résultats	Textuelle	Textuelle
Fonction de veille	Oui	Oui
Evolution technologique	Non	Oui
Sites	www.intellisek.com	www.copernic.com
Prix	169\$	79,95 EUR/USD