

**ENSSIB**

Ecole Nationale Supérieure des  
Sciences de l'Information et des Bibliothèques

**DEA en**

**SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION**

*Option "Systèmes d'information documentaire"*

**Note de synthèse bibliographique**

**OUTILS DE VEILLE STRATEGIQUE :  
essai de typologie**

**Véronique HENRY**

---

Mars 1998

# OUTILS DE VEILLE STRATEGIQUE : essai de typologie

## **Résumé**

Dans ce travail de synthèse, nous nous sommes interrogé sur les outils utilisables pour mettre en oeuvre le processus de veille stratégique. Si nombre d'entre eux proviennent du monde de la documentation, d'autres plus spécifiques se développent depuis quelques temps. Nous en proposons une typologie par catégories.

## **Mots-cles**

veille stratégique - intelligence économique - bibliométrie - veille technologique -

**Limites** : cette synthèses de recherches bibliographiques présente des outils informatiques utilisables dans un processus de veille stratégique. Elle ne prétend pas être exhaustive, mais tente de cerner les principales catégories que l'on peut distinguer.

## **Abstract**

*We have looked for strategic monitoring tools. Some of them come from libraries informatics systems, but others more specifics are now in developpement. We propose to classify them.*

## **Keywords (choosed by author)**

*competitive intelligence - information technology strategies - technology monotoring strategic information scanning system - competitive information system*

**Limits** : *this document doesn't make an exhaustive list of business intelligence tools. It just tries to perceive the different existing categories.*

## SOMMAIRE

---

<b>INTRODUCTION.....</b>	<b>4</b>
<b>METHODOLOGIE DE RECHERCHE : LES SOURCES D'INFORMATION SUR LES OUTILS DE VEILLE.....</b>	<b>6</b>
Les enseignants :.....	6
Internet :.....	7
Recherches documentaires :.....	8
Autres contacts.....	9
Acquisition des documents :.....	9
<b>LES OUTILS DE VEILLE STRATÉGIQUE.....</b>	<b>10</b>
Avant propos :.....	10
Introduction.....	12
Vocabulaire :.....	12
1. <i>EN AMONT</i> :.....	16
2. <i>LA COLLECTE D'INFORMATIONS</i> .....	16
LOGICIELS DE COMMUNICATION.....	16
LOGICIELS et PROGICIELS DOCUMENTAIRES.....	16
LA RECHERCHE SUR INTERNET.....	17
LES MOTEURS DE RECHERCHE et FEDERATEURS DE BASES.....	18
LOGICIELS DE FORMATAGE ET MEMORISATION.....	20
3. <i>TRAITEMENT DU CORPUS RECUEILLI</i> .....	20
OUTILS BIBLIOMETRIQUES.....	21
CARTOGRAPHIE ET ARBRES DE CONNAISSANCE.....	23
ANALYSE LINGUISTIQUE.....	26
4. <i>ANALYSE</i> .....	27
5. <i>DIFFUSION ET CONSULTATION</i> .....	28
6. <i>OUTILS "MULTICATEGORIES" (SYSTEMES)</i> .....	29
<b>CONCLUSION.....</b>	<b>31</b>
<b>BIBLIOGRAPHIE ET ADRESSES WEB.....</b>	<b>33</b>

## INTRODUCTION

---

Alors que les entreprises sont constamment poussées par une compétition accrue, l'information a pris une valeur si importante que savoir l'utiliser avec pertinence est une gageure de réussite. C'est ce que l'on appelle *l'intelligence économique*, terme générique qui englobe les notions de veille stratégique, commerciale, technologique, financière, ou industrielle. Activité de l'entreprise à part entière, elle nécessite une bonne organisation afin de déboucher sur une véritable aide à la décision et participer au processus d'innovation.

Si le concept d'intelligence économique n'est pas si récent que ça, son développement concret en France et en Europe date de la fin des années 80. Les entreprises ont maintenant pris conscience de l'importance d'intégrer cette activité de veille, mais la mise en oeuvre du processus n'est pas toujours facile.

En effet, l'intelligence économique recouvre toute une série d'actions continues, permettant de recouper méthodiquement les renseignements importants, définir les informations à rechercher en priorité, les sources pertinentes à cibler et les acteurs chargés de les mobiliser [DURA93]. D'autre part, l'analyse réalisée nécessite souvent de s'intéresser à une grande quantité d'information.

*"Le système d'information qui en découle repose sur la compétence des hommes et des experts, mais aussi sur un support informatique très organisé, décentralisé, menant l'information sur le bureau des utilisateurs par un réseau adéquat." [DOU93]*

Ainsi, ce sont des personnes et des outils, essentiellement informatiques, qui vont mettre en oeuvre les différentes étapes de l'intelligence économique. Cependant, en 1990, une étude menée par le CERSI [DESJ90] informait d'un certain "manque d'outils pour traiter, synthétiser et exploiter l'information afin d'en maximiser l'utilisation et d'en faciliter la circulation" et d'un "rapport flou et des frontières indéterminées entre veille et documentation."

Récemment, en 1997, une enquête sur la veille stratégique [ROIR97] faisait ressortir du côté britannique, comme du côté français d'ailleurs, un manque d'outils et de méthodes pour le traitement et l'optimisation de l'information collectée. Certaines entreprises, surtout les grandes, ont un service de veille équipé, alors que d'autres (souvent les PME/PMI) s'adressent volontiers à des agences spécialisées.

Mais depuis quelques temps de nombreux outils sont développés (des logiciels essentiellement), par des universités ou des entreprises françaises, et couvrent une grande partie du processus de veille.

Il semble important d'expliquer rapidement quelles sont les différentes étapes de ce processus, avant de faire une liste des outils existants, car à chacune d'elles vont correspondre des matériels, des logiciels spécifiques. La plupart des spécialistes de l'intelligence économique reconnaissent au moins trois étapes dans le processus : la collecte, le traitement et la diffusion de l'information [MADA97]. Certains l'élargissent en y incluant par exemple l'expression des besoins [OBER97], la mise en forme et la mémorisation de l'information [JAKO88]. De façon plus exhaustive H. Lesca [LESC97] ajoute aux étapes intellectuelles de ce processus :

- La définition de l'orientation de la veille technologique pour l'entreprise ;
- La formation des "traqueurs" d'information ;
- La transformation des signaux faibles en information utile et signifiante ;
- L'intégration des informations ainsi recueillies dans un processus décisionnel.

Nous verrons donc, dans l'analyse des informations que nous avons collectées, quels outils il est possible d'associer à ces moments forts du cycle de veille.

Le travail qui fait l'objet de cette synthèse est en quelques sorte une "veille sur la veille", que nous avons initiée afin de donner des bases nécessaires à la mise en place ou à l'amélioration d'un processus d'intelligence économique.

Avant de présenter la synthèse de notre travail et la bibliographie qui s'y rapporte, nous décrivons les étapes successives de collecte de l'information.

- ◆ *"To provide maximum benefit to a decision-maker, intelligence information must be relevant, delivered quickly, and presented in a way that facilitates rapid comprehension."* [HOHH94]

## METHODOLOGIE DE RECHERCHE : les sources d'information sur les outils de veille.

Avant de commencer toute recherche d'information, il a été nécessaire de préciser minutieusement le sujet. En effet, il est essentiel de cerner le thème, et donc le vocabulaire qui sera employé. Que recherche-t-on ? Une liste d'outils informatiques utilisés et/ou utilisables pour la recherche, le traitement et la diffusion de l'information stratégique et technique<sup>1</sup>. Cette liste sera donc divisible en catégories d'outils, et les sources devraient nous permettre d'apprécier la valeur de chacun d'eux.

Dans un premier temps, nous avons répertoriés les sources qui, intuitivement, semblaient utiles pour trouver de l'information sur les outils de veille :

- les périodiques, surtout ceux spécialisés sur les nouvelles technologies ;
- les magazines d'informatique et les publicités qu'on y trouve ;
- les répertoires des associations, des éditeurs ;
- les documents glanés sur les salons, les congrès... ;
- les démonstrations de produits, ou des contacts avec des centres de recherche, des fabricants, des revendeurs ;
- les comptes rendus de missions des collègues de travail<sup>2</sup> ;
- les enseignants (universités et écoles), en France et à l'étranger ;
- les responsables de veille d'autres entreprises ou organismes ;
- les banques de données spécialisées ;
- sur Internet : serveurs commerciaux et spécialisés, les labos de recherche publics ou privés ; les news et autres services de diffusion de l'information ;
- les professionnels de la veille (cabinets, agences privées ou publiques...) ;
- les listes de diffusion (notamment celles provenant de l'étranger, par exemple des ambassades).

Ces différentes sources ont toutes été utilisées, mais certaines se sont révélées beaucoup plus fructueuses que d'autres. La recherche s'est faite par étapes en fonction des résultats déjà trouvés, afin d'affiner ou d'ouvrir un domaine particulier. Ainsi, des sources ont été sollicitées plusieurs fois (ex : internet).

### ✓ Les enseignants :

Ce sont des enseignants chercheurs de l'université de Lyon I et de Grenoble II, ainsi que des intervenants extérieurs, qui nous ont donné les premières informations et les

---

<sup>1</sup> Nous verrons que des outils correspondants à d'autres étapes de la veille ont également été repérés.

<sup>2</sup> En effet cette recherche a été effectuée dans le cadre d'un stage au sein d'un service de veille technologique au CNET de Grenoble (Centre National d'Etudes des Télécommunications).

contacts à prendre : des cours sur les systèmes d'information, la bibliométrie ou la recherche sur internet ; des noms d'associations, d'organismes ou de spécialistes effectuant des recherches dans le domaine ; des sites web à consulter ; des articles à lire ; des démonstrations de logiciels ; etc. En effet, dans le domaine des outils de veille, un nombre assez restreint d'acteurs jouent un rôle important. Ce sont en particulier des organismes d'état et des laboratoires universitaires, qui ont initié ensemble le développement de ce secteur en France. Donc, suite à ces rencontres avec les enseignants, nous avons effectué des recherches précises pour connaître les développements actuellement en cours et les synthèses écrites.

### ✓ Internet :

C'est sur le web que nous avons trouvé les informations sur les principaux organismes français et étrangers conseillés par les enseignants, entre autre :

- le CRRM, Centre de Recherche Rétrospective de Marseille ([crrm.univ-mrs.fr](http://crrm.univ-mrs.fr)) ;
- l'IRIT, Institut de Recherche d'Information Technologique ([atlas.irit.fr](http://atlas.irit.fr)) ;
- la SFBA, Société Française de Bibliométrie Appliquée ([crrm.univ-mrs.fr/sfba/sfba.html](http://crrm.univ-mrs.fr/sfba/sfba.html)) ;
- la SCIP, Society of Competitive Intelligence Professionals ([www.scip.org](http://www.scip.org)) ;
- l'ECAM, Centre Européen de Mathématiques Appliquées d'IBM ([www.fr.ibm.com/france/ecam/ecam](http://www.fr.ibm.com/france/ecam/ecam)) ;
- l'IEP de Lyon (<http://iep.univ-lyon2.fr/SE71/Veille/agents.htm>) ;
- l'INIST, Institut National de l'Information Scientifique et Technique ([www.inist.fr](http://www.inist.fr)).

Ces sites ont été soigneusement analysés, dans le but de repérer des informations sur des outils de veille.

Le site du CRRM présente, outre les activités générales de recherche sur la bibliométrie, un outil développé dans ce laboratoire. Chaque partie du logiciel est présentée en détail. La SFBA est rattachée au CRRM, et les présentations de colloques figurent dans le même site. Après lecture des textes du dernier colloque, il s'est avéré qu'aucun outil nouveau à notre connaissance n'était présenté.

Le site de l'IRIT propose un outil assez différent de celui du CRRM, et une série d'autres outils couvrant d'autres étapes de la veille.

Autre source intéressante : le site du CEVEIL (Centre d'Expertise et de Veille Inforoutes et Langues, au Québec). Il présente en particulier les travaux d'étudiants de l'Ecole de Bibliothéconomie et des Sciences de l'Information de l'Université de Montréal, où l'on retrouve des synthèses sur le thème des technologies de l'information [CEIV96].

Une recherche a également été effectuée avec les moteurs de recherche par mots-clés et les métamoteurs, surtout à partir des noms des outils que nous voulions analyser plus en détail. Ces recherches ont été la plupart du temps fructueuses et

nous ont permis d'atteindre les sites des concepteurs ou des distributeurs de ces produits.

Par contre la recherche "exploratoire" avec des termes comme "veille technologique" ou "intelligence économique" n'a pas vraiment permis de trouver de nouveaux outils par rapport à ceux déjà en notre connaissance.

### **✓ Recherches documentaires :**

A la Bibliothèque de l'ENSSIB, l'interrogation du catalogue en ligne a permis de trouver 36 documents pertinents sur la veille technologique (avec également des informations plus spécifiques à l'informatique documentaire et internet).

Nous avons interrogé les CD-ROMs suivants :

- Docthèse : 19 réponses, dont 3 retenues.
- Lisa : 37 réponses, dont 4 retenues.

Nous avons consulté les périodiques suivants :

Sur Internet :

- Internet Professionnel
- Internet Research
- Planète Internet

Sur l'informatique :

- 01 Informatique
- Informatique professionnelle
- Monde informatique

Sur les technologies :

- Technologies internationales
- Vigie (Multimédia et inforoutes, Informatique appliquée)

La lecture d'articles sur le thème de la veille, ou des rubriques présentant de nouveaux produits a enrichi nos connaissances sur le domaine.

A la Bibliothèque de l'Université Lyon1 nous avons consulté des périodiques et une thèse (micro-fiche).

A la Bibliothèque du CNET de Grenoble, nous avons également procédé à une interrogation du catalogue, et avons pu ainsi consulter, outre les périodiques spécialisés, quelques ouvrages généraux sur la veille.

Nous avons fait une recherche sur Téléthèse (interface Minitel), mais sans trouver d'information nouvelle. Par contre, l'interrogation de Pascal a été fructueuse : plus de 1000 articles répondaient à l'interrogation par les termes "veille stratégique", "veille technologique", "bibliométrie" et "intelligence économique". Les réponses se

recoupant parfois, et étant par ailleurs déjà connues souvent, nous n'avons conservé qu'une dizaine de références.

Par ailleurs, la consultation des catalogues accessibles par le Web ont permis de repérer entre autre une thèse pertinente.

### **✓ Autres contacts**

Enfin des contact de professionnels nous ont beaucoup apporté dans cette recherche :

- des laboratoires de recherche (pour des explications sur les outils développés),
- des distributeurs d'outils de veille (entre autre pour des démonstrations),
- des responsables de services de veille dans d'autres entreprises (pour les systèmes mis en place et les outils utilisés).

### **✓ Acquisition des documents :**

Nous avons pu consulter une bonne partie des documents cités dans les bibliothèques (ENSSIB, Lyon1 et CNET), mais également en les imprimant directement à partir du web. Pour les autres articles ou thèses non disponibles sur place, nous avons utilisé le prêt inter-bibliothèques.

- ◆ Pour cette recherche, nous pouvons conclure que les informations principales nous ont été apportées par des spécialistes ou des enseignants, et que les documents récupérés par une recherche documentaire plus classique n'ont fait que compléter et confirmer ces derniers.

## **Les OUTILS de veille stratégique**

---

### **Avant propos :**

- Afin de couvrir les différentes étapes du processus de veille, de nombreux outils informatiques peuvent être utilisés. Il ne faut cependant pas les confondre avec

les méthodes à appliquer, car ils ne constituent qu'un apport, un gain de temps et d'efficacité<sup>3</sup>.

- Nous ne nous attacherons pas à lister le matériel informatique de base nécessaire à l'acquisition et l'installation de ces logiciels. Cependant il faut noter que les connexions aux réseaux, l'accès à des bases de données, le support de logiciels assez volumineux et la quantité d'informations traitées demande une capacité mémoire importante (il est évidemment nécessaire d'avoir une bonne capacité de stockage de l'information recueillies).
- Nous n'avons conservé ici que les outils utilisables dans le cadre d'une activité de veille stratégique. Les outils spécifiques de "datawarehouse", liés à la notion d'informatique décisionnelle, ne sont pas cités, bien qu'ils concernent la gestion et la diffusion de l'information dans l'entreprise. Seuls certains de ces logiciels trouvent une application dans un processus de veille stratégique.
- Les outils présentés sont essentiellement français, pour deux raisons. D'une part, les laboratoires français ont fait preuve de dynamisme dans le développement d'outils dans des domaines comme la bibliométrie par exemple, et sont à l'origine de la plupart des logiciels existants. D'autre part, la spécificité du domaine (traitement d'un corpus de données qui sont le plus souvent textuelles) nous amène à penser que les outils francophones sont plus adaptés à notre langue car ils en respectent la logique. Bien sûr, la plupart permettent de travailler sur un corpus de textes en anglais, la langue la plus courante dans les différents domaines de recherche. Seuls quelques "incontournables" développés à l'étranger (comme les moteurs de recherche) sont donc ajoutés à la liste. Mais il faut savoir que de nombreux pays ont entamé des recherches sur le domaine (datamining et bibliométrie).
- Il est évident que nous n'avons pas pu tester tous les outils que nous décrivons. Les informations recueillies proviennent souvent des sociétés qui les ont produites, et de ce fait sont toujours très "optimistes" quant à l'efficacité de leur produit. Cependant, de nombreux articles de presse nous ont permis d'apporter un minimum de critique à certains de ces outils.
- Enfin le classement que nous proposons est assez subjectif, dans la mesure où la plupart de ces outils n'ont pas qu'une seule et unique fonctionnalité. Nous avons choisi de faire ressortir celle qui les caractérise le mieux.

---

<sup>3</sup> L'utilisation de moyens informatiques pour rechercher l'information stratégique n'est pas non plus incompatible avec la collecte de documents papier ou de données informelles, recueillis lors de salons, de visites, etc.

## **Introduction**

Lors du colloque IDT93, B. Ragot affirmait que, malgré les systèmes complexes et automatique d'analyse existant, *"les machines, même les plus sophistiquées", étaient "encore et pour longtemps dénuées d'imagination"* [RAGO93]. Quelques années plus tard, la liste des outils informatiques de recherche, de traitement, et d'analyse de l'information n'a cessé de croître. S'il est évident que la machine ne remplacera pas l'homme dans le jugement final de pertinence accordé à une information, les logiciels disponibles effectuent un tri considérable sur la masse croissante de documents, et constituent une aide non négligeable pour à la fois gagner du temps et faire ressortir l'essentiel.

Alors que jusqu'à présent on utilisait essentiellement une panoplie d'outils venus tout droit du monde de la documentation, on trouve maintenant des logiciels ou progiciels véritablement dédiés à l'intelligence économique. Ces outils intègrent plusieurs étapes du processus de veille stratégique, généralement le traitement et l'analyse de l'information (aspects linguistiques et bibliométriques). Cependant, certains auteurs sont tentés de réduire les outils de veille stratégique à ces logiciels infométriques. Or, si l'on reprend les étapes du processus de veille, tel que présenté par H. Lesca (voir p. 5 introduction), et qu'on essaie d'y rattacher les outils actuellement disponibles sur le marché, nous nous rendons compte que certaines des étapes ne sont pas encore vraiment "outillées" (par exemple l'identification des signaux faibles et l'intégration des informations traitées dans un processus de décision) [LESC97]. Mais il semble difficilement envisageable de tout vouloir automatiser, certaines opérations intellectuelles ne pouvant se passer de l'intervention d'experts.

Lorsqu'on parle d'analyse automatisée de l'information recueillie, il faut distinguer deux sources qui induisent deux types de traitements : les banques de données bibliographiques fournissent des documents structurés en champs, alors que d'autres sources apportent des documents plein texte. Dans le premier cas les outils bibliométriques s'appliquent parfaitement au tri, à la comparaison et au regroupement des données, alors que dans le second cas il faudra utiliser des techniques statistico-linguistiques pour analyser le texte dans son ensemble.

De même, en fonction des besoins d'information (d'exploration, de prospective, de structuration...) on n'appliquera pas tout à fait les mêmes traitements [ROUS97].

## **✓ Vocabulaire :**

Un vocabulaire particulier aux notions de veille stratégique et d'intelligence économique s'est développé parallèlement à ces concepts. Souvent issus de l'anglais, ces termes recouvrent certaines parties du processus, et donc les outils qui y correspondent. Attention cependant aux traductions littérales, il existe des nuances dont il faut tenir compte.

- **Moteur / métamoteur**

Les moteurs de recherche sont bien connus des utilisateurs du Web. Il faut distinguer ceux qui fonctionnent par mots-clés et ceux qui fonctionnent par thèmes (qui se développent actuellement d'ailleurs). Les métamoteurs interrogent simultanément plusieurs moteurs dans leur syntaxe, et dédoublonnent les résultats finaux.

- **Agent intelligent**

En plein essor pour la recherche de données sur internet, les agents intelligents sont actifs en permanence, en traitement de fond, et rapatrient régulièrement les résultats de leur travail. Ils "voyagent" dans le réseau de façon autonome, toujours actifs même quand on éteint son ordinateur, et rapatrient le résultat de leur recherche lorsqu'on le rallume. Leur but est d'extraire en temps réel l'information utile au sein d'une masse de données réparties sur des milliers de serveurs, en la filtrant de façon automatique, suivant des critères choisis par l'utilisateur [GAZE98]. Comme l'analyse linguistique reste délicate, rares sont les moteurs qui soient immédiatement opérationnels. Ils sont donc dotés d'une faculté d'apprentissage : par rapport à ce que l'on décide de garder sur une précédente recherche, en observant les sites favoris des caches du navigateur, etc. Ils ont également une fonction de dédoublonnage et de vérification permanente de la validité des liens.

Les grands moteurs de recherche utilisent des agents intelligents (Topic par exemple), mais il existe aussi des outils moins puissants téléchargeables sur le net à moindre coût.

- **Endogène**

L'information qui n'apparaît que grâce à la comparaison ou à l'association de données contenues dans des champs ou des documents différents est dite endogène. A opposer à l'information exogène.

- **Infométrie**

Adopté en 1987 par l'IFD (International Federation of Documentation), ce terme désigne l'ensemble des activités métriques relatives à l'information, couvrant aussi bien la bibliométrie que la scientométrie [NOYE95].

- **Bibliométrie** ("bibliographie / mesure")

Notion apparue dans les années 70, la bibliométrie est une application des mathématiques au domaine de l'information et de la documentation<sup>4</sup>. De plus en plus appliquée au domaine de la veille stratégique, la bibliométrie est un *« système d'analyse dynamique permettant de positionner les sujets et les techniques »*

---

<sup>4</sup> Auparavant, à partir des années 30, on parlait surtout de scientométrie (étude quantitative de la science et de la technologie). On peut maintenant la définir comme étant la bibliométrie spécialisée au domaine de l'IST.

[DOU90]. On utilise entre autres les statistiques, les probabilités, les systèmes experts pour recenser, trier et analyser automatiquement un corpus de données pouvant être très volumineux. Ces méthodes sont en particulier efficaces sur des références bibliographiques, qui sont des documents bien structurés. Les outils bibliométriques vont pouvoir mettre en évidence les mots-clés et les liens qu'ils entretiennent entre eux, mais aussi les filiations entre auteurs (en mesurant les citations), les thèmes et l'intensité des collaborations entre les laboratoires de recherche.

Ces outils utilisent des techniques d'indexation automatique pour lister les mots-clés du corpus de documents. Puis seuls les termes porteurs de sens sont conservés grâce à l'analyse linguistique et morpho-syntaxique, la lemmatisation (méthode qui consiste à simplifier à leur forme canonique les verbes, adjectifs et autres substantifs). L'étape suivante consiste alors à regrouper les mots-clés pour constituer des agrégats (clusters) avec des liens plus ou moins forts. D'autres techniques travaillent sur les nuages de données (méthode multidimensionnelle) pour permettre au veilleur d'analyser les rapports entre elles et suivre leurs évolutions. On peut alors, par recoupement, *"mettre en évidence des informations stratégiques comme les sujets de recherche émergeant sur un thème donné, la composition des équipes d'experts qui collaborent sur un sujet particulier, les experts les plus importants, les organismes phares, etc. (...)* La qualité des résultats obtenus dans ce type de travail dépend d'une complémentarité forte entre le spécialiste des logiciels utilisés et les experts du domaine étudié, capables d'interpréter par exemple les signaux perçus sur des cartes" [THIL97a].

- **cartographie**

A partir d'un corpus donné, les logiciels bibliométriques produisent une représentation graphique mettant en évidence les relations hiérarchiques entre les termes (classification), les positionnant les uns par rapport aux autres (calculs de densité, de centralité). La visualisation graphique facilite l'interprétation et l'analyse des données. Peu de logiciels ne font que de la cartographie, en général cette application se retrouve dans les outils bibliométriques et des systèmes plus complets.

- **datawarehouse**

Littéralement "entrepôt de données", le datawarehouse est en fait le stockage et la mise à disposition des données de l'entreprise. Des outils permettent d'interroger, de visualiser et d'analyser ces bases ainsi constituées.

- **datamining**

Outil complémentaire au datawarehouse, le datamining est un procédé interactif, couplé au système d'information décisionnel de l'organisation, permettant d'expliquer des résultats, de corrélérer des événements, de prévoir des comportements ou de

confirmer des hypothèses. Les outils de datamining utilisent entre autres les techniques des arbres de décision et les réseaux de neurones [GAZE98].

## **1. EN AMONT :**

Avant de se lancer dans une activité de veille, il est important de définir les besoins de l'entreprise, les bases éventuellement existantes, et de comprendre ce que sont des informations à caractère stratégique. Le CERAG (laboratoire de recherche de l'Ecole Supérieure des Affaires à Grenoble) a développé ces dernières années un ensemble d'outils informatiques permettant de couvrir ces étapes de la veille.

On notera en particulier les prototypes suivants [LESC97] :

- Pour mesurer l'intérêt de la mise en place d'une veille stratégique et identifier les besoins de l'entreprise : PERTINENCE ;
- Pour identifier les acteurs concernés par la veille : CIBLE, système d'aide au raisonnement interactif ;
- Pour mesurer l'existant et la capacité d'une entreprise à développer l'activité de veille : FENNEC. Composé d'une base de connaissance, d'une base de faits, et d'un moteur d'inférence, le logiciel permet de faire un diagnostic précis avant d'initier ou d'améliorer une activité de veille ;
- Pour sensibiliser les acteurs (traqueurs ou pas) : SENTINEL et OASIS, didacticiels qui proposent des "cours" théoriques (les références principales de la veille) et des exercices pratiques de traque de l'information.

## **2. LA COLLECTE D'INFORMATIONS**

L'étape de collecte d'information implique le plus souvent une recherche sur des sources électroniques (banques de données, internet...)<sup>5</sup>. Les différents moteurs et logiciels de recherche permettent de collecter des informations structurées ou plein texte, et il faut ensuite compiler ces documents dans une base de données pour passer à l'étape du traitement.

### **✓ LOGICIELS DE COMMUNICATION**

Ces logiciels permettent l'extraction des informations dans des bases. On citera par exemple XTALK (Crosstalk Communication), INFOLOG (I + K), COM 2400 (La Commande Electronique), KXCOM (Kortex) ou DIALOGLINK (Dialog) [CAST94].

### **✓ LOGICIELS et PROGICIELS DOCUMENTAIRES**

Pour gérer des données textuelles, les logiciels et progiciels documentaires permettent de créer des interfaces d'interrogation spécifiques, et de diffuser les

---

<sup>5</sup> En effet, plus de 90 % de l'information est aujourd'hui disponible sous forme électronique.

documents suivant différents modes (de plus en plus souvent par le biais d'une interface web). Ils sont très adaptés aux bases internes de l'entreprise.

### ✓ LA RECHERCHE SUR INTERNET

On ne compte plus aujourd'hui les outils proposés gratuitement sur internet.

**Sur le Web** plusieurs catégories peuvent cependant être définies<sup>6</sup> :

- les moteurs de recherche par mots-clé
- les moteurs thématiques
- les annuaires
- les métamoteurs (intégrateurs complexes : lancent plusieurs moteurs simultanément)
- les agents intelligents (voir p. 12) : à distinguer des robots des moteurs de recherche, les agents intelligents permettent de faire un premier tri. On notera par exemple :
  - AUTONOMY (requête en langage naturel ; tient compte des résultats sélectionnés par l'utilisateur ; capable de communiquer avec d'autres agents ayant les mêmes centre d'intérêt).
  - FIREFLY (intégré à My Yahoo -recherche avec profils utilisateurs ; met en relation des utilisateurs partageant les mêmes centre d'intérêt).
  - JANGO (spécialisé dans la recherche de prix proposés sur les sites Web ; affiche les résultats au fur et à mesure).

D'autres logiciels sont utiles pour effectuer la recherche et le tri de l'information sur internet. DIGOUT4U par exemple ([www.arisem.com](http://www.arisem.com)) reformule la requête posée en langage naturel et utilise les moteurs de recherche sur plusieurs niveaux de profondeur des liens. Il procède ensuite à une analyse sémantique des pages récupérées pour hiérarchiser les documents par degré de pertinence. A noter : Digout4U est capable d'effectuer la recherche indifféremment en français et en anglais.

**Sur les groupes de news et les listes de diffusion**, l'analyse des messages qui circulent se révèle souvent fructueuse et permet de faire apparaître des domaines d'intérêt émergents. L'un des outils les plus connus est DEJA NEWS. Proposant de nombreuses options de recherche, cet outil permet de visualiser une liste sous plusieurs angles, de retrouver des informations sur les auteurs des messages (à quelles listes ils appartiennent, le nombre de messages qu'ils

---

<sup>6</sup> Pour une liste complète et commentée d'outils de recherche sur Internet : <http://www.adbs.fr/adbs/viepro/sinfoint/lardy/chap4.htm> [LARD97]

envoient...), et donc de trouver à la fois des réponses et des contacts utiles [OBER97].

Toujours dans cette catégorie d'outils, on notera MECIA (société Inforama International), qui est un logiciel permettant d'observer en continu un flux de messages par rapport à des profils (par mots-clés) ; quant à L4U (développement de Taiga), il filtre les informations dans plusieurs langues [LESC97].

**Rapatriement de sites** : certains logiciels permettent de rapatrier tout ou partie d'un site pour en exploiter le contenu hors ligne <sup>7</sup>. Cette fonction peut aussi être une des options des moteurs de recherche.

#### **Outils de repérage de modification d'un site :**

Par exemple Net AttachePro ([www.tympani.com](http://www.tympani.com)) ou Tierra Highlight ([www.tierra.com](http://www.tierra.com)). Cette fonction fait aussi souvent partie d'outils plus complets de recherche et de traitement de l'information.

### **✓ LES MOTEURS DE RECHERCHE et FEDERATEURS DE BASES**

Les grands fabricants de ces moteurs utilisés par de nombreux logiciels sont Verity et Fulcrum. Mais d'autres spécialistes de la documentation développent également des outils puissants.

VERITY développe une gamme de produits destinée aux applications de recherche d'informations dans l'entreprise, sur Internet et autres sources (bases de données, CD-ROM).

- TOPIC est un système intelligent de recherche documentaire en texte intégral permettant de constituer des bases de connaissance. L'interrogation peut se faire soit en texte libre, soit au moyen de concepts existant dans la base, et la diffusion par messagerie offre un rapport personnalisé en fonction du profil d'intérêt de l'utilisateur. Très utile pour la recherche d'informations ciblées, le moteur Topic est utilisé par de nombreux fournisseurs de GED (Adobe, Lotus, Siemens, etc). Topic offre toute une famille de produits de recherche, de filtrage, d'extraction et d'analyse d'information. Pour l'activité de veille, on peut citer le WebSearcher, un serveur qui comprend un module de recherche intégré, et les Topic Agents.
- INTELLISERV, autre outil de Verity, permet d'effectuer des recherches dans des sources d'information hétérogènes contenues dans un serveur interne, et de surveiller automatiquement certains sujets par une technologie de push. Les

---

<sup>7</sup> On en trouvera également une liste à l'adresse suivante : <http://www.adbs.fr/adbs/viepro/sinfoint/lardy/aspirat.htm>.

documents sont transmis par ordre de pertinence, et un résumé automatique est fourni pour chacun. La diffusion aux utilisateurs peut se faire par le web, par messagerie ou par télécopie [GAZE98].

FULCRUM : la société canadienne Fulcrum Technologies est un des principaux éditeurs de logiciels de recherche d'informations pour l'entreprise.

On retiendra en particulier Knowledge Network, une gamme modulaire de logiciels centrée sur une architecture web. Elle comprend un moteur de recherche documentaire capable d'intégrer des informations issues de plusieurs types de sources. L'information est présentée sous la forme d'une arborescence personnalisée regroupant les documents en fonction de leur contenu. Knowledge Network utilise la recherche en langage naturel et l'analyse morpho-syntaxique [GAZE98].

KNOWLEDGE BROKERS<sup>8</sup> : Développé par le Xerox Research Center Europe et en utilisation pilote sur des sites européens, il s'agit d'une suite de logiciels qui composent un outils de recherche d'information sur Internet et sur des bases de données, une "plate-forme de médiation" entre l'utilisateur final et des sources diverses. Les Knowledge Brokers permettent d'interroger simultanément des sources hétérogènes, avec une même équation ; les résultats composeront un documents unique, dont le format sera utilisable quelles que soient les sources qui auront servi à sa création.

---

<sup>8</sup> Littéralement "Courtiers de connaissance".

- ◆ La plupart de ces outils reposent désormais sur une adéquation entre les algorithmes statistiques et la linguistique [CAHI97]. Par exemple la société française ERLI, spécialisée en informatique linguistique, vient d'adapter à Knowledge Network des technologies de traitement de l'information : Information Retriever (requête en langage naturel) et Terminology Manager (ajoute des connaissances à la base de connaissances linguistique personnalisable). Des accords de coopération ont également été passés avec Verity pour compléter le traitement sémantique de Topic.

### **✓ LOGICIELS DE FORMATAGE ET MEMORISATION**

Si les données collectées doivent faire l'objet d'un reformatage pour permettre le traitement, on pourra utiliser des logiciels spécifiques (si cette fonction n'est pas déjà incluse dans les outils de recherche). Exemple : INFOTRANS (I + K), F2F (Frankia), FILTER (Information Automation LTD) ou REFWRITER (JSA LTD) [CAST94].

De la même façon, s'il faut mémoriser les données dans des fichiers internes, on pourra choisir entre les différents types de systèmes de gestion des bases de données (SGBD) et les gestionnaires électroniques de documents (GED).

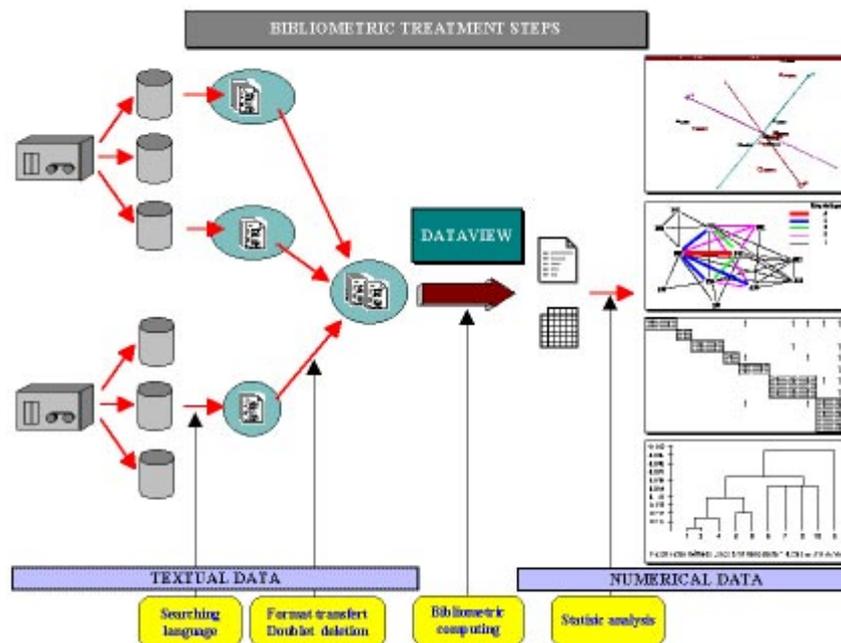
### ***3. TRAITEMENT DU CORPUS RECUEILLI***

Une fois la base de connaissance constituée, il est nécessaire d'appliquer un certain nombre de traitements pour faire un tri, classer et représenter de façon simple le corpus d'informations (en particulier les données en provenance d'internet, souvent volumineuses et de formats divers). C'est ici qu'interviennent les outils d'analyse statistique, bibliométriques et cartographiques. Notre but n'est pas d'entrer dans les détails de fonctionnement de ces logiciels, mais de présenter les finalités des plus connus.

## ✓ OUTILS BIBLIOMETRIQUES

DATAVIEW, développé par le CRRM (Centre de Recherche Rétrospective de Marseille ; <http://crrm.univ-mrs.fr/commercial/software/dataview/dataview.html>)

Les concepteurs de Dataview sont partis du constat suivant : dans le traitement bibliométrique, peu de processus sont entièrement automatisés, sauf deux étapes : la collecte (logiciels de communication) et l'analyse des données (logiciels statistiques, avec une présentation graphique des données numériques). Le problème est de pouvoir traiter non seulement les données numériques, mais aussi les données textuelles, avec un même outil. Quelques recherches sont en cours <sup>9</sup>, mais souvent ces outils sont limités dans un traitement spécifique. Par contre, Dataview a été développé dans l'optique de pouvoir traiter automatiquement les données textuelles comme les données numériques, et d'offrir plusieurs types d'analyses. Le schéma suivant<sup>10</sup> explique la démarche de Dataview : un langage de recherche sur des sources variées et hétérogènes, le reformatage des données avec élimination des doublons, le traitement bibliométrique, puis l'analyse statistique (avec plusieurs types de représentations graphiques).



<sup>9</sup> Dont : Brookes, PatStat + , Patent trend analysis, Leximappe

<sup>10</sup> <http://crrm.univ-mrs.fr/commercial/software/dataview/fig1.gif>

- ◆ L'avantage de Dataview est de pouvoir apporter, dans un processus de veille stratégique, des éléments déterminants de type : suivi des technologies de concurrence, évolution des tendances, état de l'environnement technologique, mise en place de systèmes d'alerte, etc. Cependant, d'autres méthodes complémentaires (outils statistiques infographiques) doivent être utilisées si l'on veut se servir efficacement de Dataview [OBER97].

TETRALOGIE développé par l'IRIT (Institut Régional d'Information Technologique, à Toulouse). Il s'agit en fait d'un ensemble d'outils intégrés qui recherche des sites web pertinents, retrouve les noms des chercheurs les plus actifs dans un domaine et met ainsi en évidence des collègues invisibles.

Tétralogie fait partie du système ATLAS, développé sous l'impulsion d'organismes d'état (CEDOCAR<sup>11</sup> en particulier) et la SFBA<sup>12</sup> dans le but de favoriser le développement de la veille technologique. C'est donc le maillon d'une chaîne de logiciels qui comportent :

- des outils réalisés par le CRRM et le CEDOCAR : techniques d'acquisitions de données et procédés statistiques préliminaires.
- à l'INIST : élaboration de programmes qui incorporent les méthodes d'extraction d'informations des textes libres et les moyens de formatage des données (Neurodoc).

Le logiciel est alimenté par des banques de données bibliographiques, textuelles ou factuelles. Il utilise les différentes méthodes d'analyses des données (ACP, AFC, CAH etc.) et des techniques spécifiques de visualisation (2D à 4D, animation interactive ou automatique)<sup>13</sup>.

*Voici par exemple le résultat d'une recherche croisée (noms d'experts, sujets de recherches, travaux...) sur des bases bibliographiques. La carte est obtenue en appliquant une AFC (analyse factorielle de correspondance) sur une matrice récapitulant les données [<http://atlas.irit.fr/petitexemple.html>].*

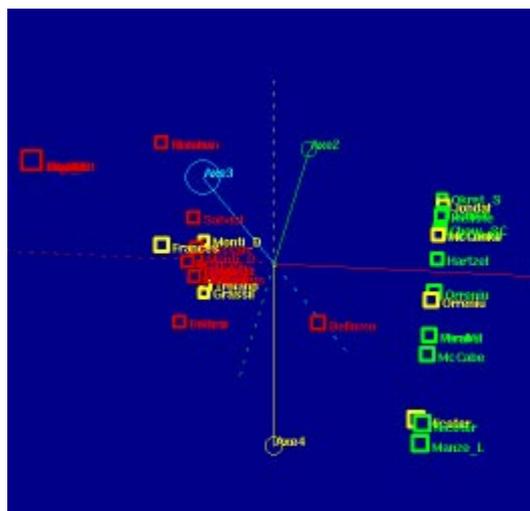
*On voit clairement apparaître des groupes qui travaillent et collaborent sur un même sujet de recherche. Cette carte met aussi en évidence les noeuds, c'est-à-dire les experts qui sont au centre d'un réseau.*

---

<sup>11</sup> Centre de Documentation de l'Armement

<sup>12</sup> Société Française de Bibliométrie Appliquée

<sup>13</sup> Voir pour plus de détails la page <http://atlas.irit.fr/analyse.html>



Ce type de représentation, minutieusement manipulé et analysé, pourrait mettre en évidence des sujets de recherche émergents (information endogène, c'est-à-dire difficilement décelable)<sup>14</sup>.

- ◆ L'ADIT<sup>15</sup> utilise Tétralogie dans son processus de recherche et d'analyse de l'information. Le logiciel permet d'identifier dans un premier temps les réseaux de collaboration et leur organisation, détermine leurs travaux, leur évolution dans le temps, etc. Après que des experts aient examiné ces résultats, des thèmes très précis de surveillance sont définis, et une recherche automatique est lancée avec TOPIC afin de programmer une veille en continu [THIL97a].

### ✓ CARTOGRAPHIE ET ARBRES DE CONNAISSANCE

Des logiciels de bibliométrie ou de traitement intègrent souvent la cartographie à leur outil (Périclès, Tétralogie, Dataview...). Mais quelques logiciels sont dédiés à cette fonction, dans le but de faire émerger du sens à une grande masse de données<sup>16</sup> :

NEURONAV + : classe par thème un ensemble de documents indexés détenus au sein de l'entreprise et synthétise ces informations sous forme de cartes. Il offre de plus un environnement d'exploration des documents par navigation entre les mots-clés, les documents et les thèmes. La visualisation graphique met en évidence l'importance de chaque thème dans le corpus. Il est possible d'effectuer une comparaison entre plusieurs cartes, de modifier l'indexation et de réduire la masse des documents analysés.

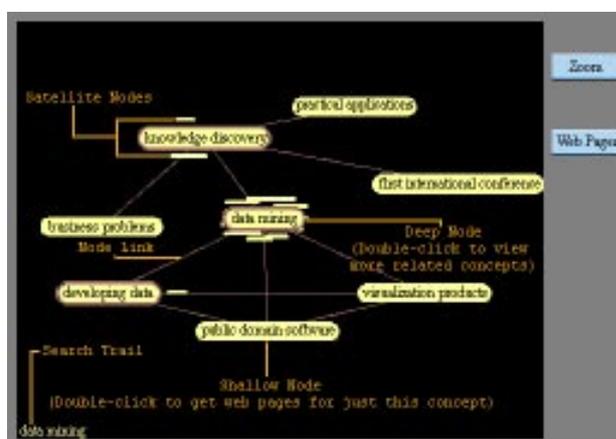
<sup>14</sup> Cependant, certains spécialistes pensent que Tétralogie ne peut être utilisé pour traquer des signaux faibles [LESC97].

<sup>15</sup> Agence pour la Diffusion de l'Information Technologique

<sup>16</sup> On retrouve dans cette catégorie les logiciels d'analyse de données qualitative et quantitative, qui offrent une présentation graphique des traitements [COUE96].

- ◆ Ce logiciel est par exemple très utile pour se familiariser avec un fonds documentaire [GAZE98].

SEMIOMAP (Laboratoire de sémiotique informatique du Pôle universitaire Léonard de Vinci - [www.semio.com](http://www.semio.com)) : ce logiciel indexe l'ensemble des pages web existantes, et sur cette base fournit une sorte de carte sémantique sous la forme de diagramme présentant des agrégats statistiques [NAJM97]. Le traitement du volume de données est très rapide. Les concepts dépendants de l'équation de recherche sont représentés comme des noeuds qui sont liés entre eux pour former des clusters (réseaux de concepts). Les noeuds les plus "profonds" sont encadrés, et un système Search Trail permet de garder le cheminement de la recherche. La fonction Zoom permet quant à elle de refaire une recherche plus pointue sur des noeuds sélectionnés. Enfin le bouton Web Pages donne un accès direct aux pages concernant le concept.



Le module Semioscan identifie sur la carte les modifications à date donnée.

- ◆ Cet outil s'appuie essentiellement sur un traitement statistique de la proximité entre les mots du texte [GAZE98].

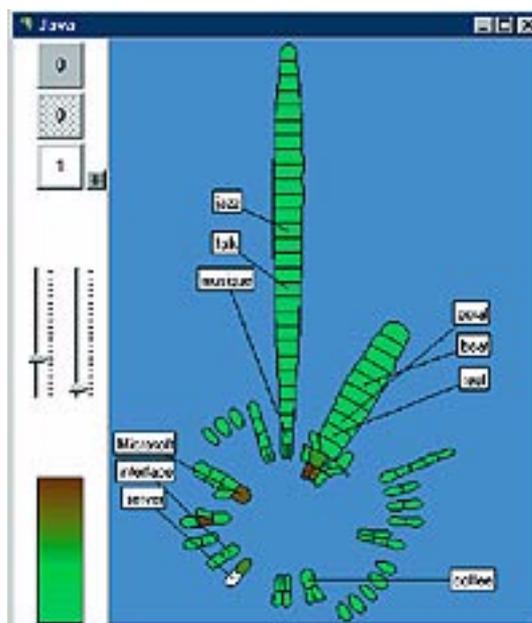
SAMPLER (Société Cisi) : ce logiciel analyse un corpus de texte en procédant à une extraction automatique des termes et en faisant une représentation graphique du lexique ainsi formé. Des clusters mettent en évidence des liens plus ou moins forts entre les mots (des co-occurrences) et font ressortir des domaines émergents. Ce logiciel peut aussi être utilisé comme une aide à la reformulation documentaire [GAZE98].

- ◆ Ce logiciel est utilisé par l'ADIT, en combinaison avec Tétralogie.

U-MAP (Société TRIVIUM - [www.umap.com](http://www.umap.com)) : le logiciel permet de paramétrer une requête (français / anglais), de sélectionner des sources, et de rapatrier les pages web sélectionnées. L'interface conviviale permet alors de "travailler" hors ligne sur la masse des documents sans les ouvrir. Cette phase préalable permet d'éliminer l'inutile, en changeant souvent de point de vue pour s'intéresser tantôt à l'ensemble, tantôt aux détails. Umap dresse ensuite une cartographie du corpus, en mettant en

évidence les liens entre les mots, et en offrant des outils de simulation d'hypothèses, d'exploration des singularités, etc [CAHI97].

*Exemple de carte produite pour le mot "Java".*



Autre outil développé par Trivium : le logiciel GINGO construit, à partir d'une masse assez importante d'informations, des "arbres de technologie", c'est-à-dire permettant de suivre l'évolution technique d'un secteur.

INFOSCAN (<http://www.MachinaSapiens.qc.ca/machina/>) :

C'est un logiciel de filtrage et de visualisation de l'information issue du courrier électronique et des News. InfoScan travaille à partir d'un corpus constitué de documents en format texte [LARD97].

## ✓ ANALYSE LINGUISTIQUE

Les documents collectés peuvent subir une indexation permettant un traitement ultérieur. A cette fin, des outils d'analyse syntaxique, linguistique et sémantique sont intégrés à des logiciels puissants de recherche documentaire.

SPIRIT : logiciel développé par le Commissariat à l'Energie Atomique, Spirit fait une indexation automatique à la fois statistique et linguistique, à l'aide d'un dictionnaire, d'un analyseur morpho-syntaxique, et d'un module de reformulation des requêtes (multilingue ; traite le langage naturel) ; il repère dans les bases de données les variations de vocabulaire (nouveaux termes ou disparitions). Les résultats sont présentés par classe de pertinence en indiquant les unités lexicales sélectionnées. Spirit-Sense est le module sémantique qui intervient entre la phase d'analyse syntaxique et la reformulation. Les idées contenues dans les documents sont extraites automatiquement et associées à des concepts.

- ◆ Spirit est par exemple intégré à Périclès (p. 28).

TAIGA (Traitement Automatisé de l'Information Géopolitique d'Actualité). Outil de recherche sémantique des informations. Comme son nom l'indique, TAIGA est dédié au traitement et à l'analyse des informations géopolitiques et d'actualité. L'indexation se fait par thèmes, et non par mots-clés. Comme Topic, Taiga permet d'effectuer les interrogations en texte libre ou suivant des thèmes classés de façon arborescente. Taiga fonctionne dans plusieurs langues, et est expert en sémantique et en linguistique. Il transforme les textes depuis n'importe quelle langue dans un langage pivot qui regroupe les terminologies autour de champs sémantiques.

- ◆ Complexe d'utilisation, mais très rapide (il traite 1 milliard de caractères par seconde), TAIGA est utile pour la collecte des informations de potentiel disponibles dans des sources formelles [LESC97].

TEXT MINING : Développé par l'ECAM (Centre Européen de Mathématiques Appliquées d'IBM), le Text Mining est un ensemble de méthodes mathématiques et linguistiques permettant de traiter les données pour les structurer et les analyser. Deux produits d'IBM intéressent particulièrement la veille stratégique :

- TKS (Text Knowledge Server), solution d'analyse de bases documentaires de sources diverses. TKS fonctionne avec des modules de préparation des données (formatage, filtrage), de traitements linguistiques, classificatoires, statistiques, d'exportation, de rapport et de visualisation des résultats.
- TECHNOLOGY WATCH, version allégée de TKS, travaille uniquement sur des textes indexés et des données codées, mais n'inclue pas l'analyse linguistique et les modules de recherche sur Internet. Technology Watch permet de faire une analyse automatique d'information bibliographique.

- ◆ Ce logiciel est particulièrement adapté au traitement de la littérature scientifique et technique en provenance des bases de données externes mais aussi internes. [IBM98]

- Dans la gamme des outils développés par IBM, on peut également citer CRI, qui est un outil dédié à l'analyse de corpus de presse, de discours, de campagnes publicitaires, de résultats d'enquêtes d'opinion, etc [GAZE98].

LEXIMAPPE : programme débuté au début des années 80 par le Centre de Sociologie de l'Ecole des Mines de Paris et le SERPIA (INIST) et classé dans les outils de clustérisation. Leximappe fonctionne avec la méthode de mots associés. Il indexe automatiquement et repère le flux des mots (nouveaux, qui disparaissent ou qui évoluent), en gageant que "*l'association entre les termes d'indexation rende visible la trame d'un domaine de recherche*" [POLA93]. Mais des tests montrent que si l'on peut effectivement identifier des éléments intéressants avec Leximappe, les signaux faibles restent difficilement identifiables [MAUR96].

SPHINX LEXICA (société Sphinx Développement) : C'est un logiciel de traitement du langage naturel. A partir de grammaires, de dictionnaires et par l'analyse syntaxique, Sphinx Lexica procède à une analyse morphologique (lemmatisation) puis à une analyse syntaxique. Il est possible de sortir une topographie du corpus et une carte qui met en évidence les "grappes" de concepts accrochés à un mot stratégique.

#### **4. ANALYSE**

Une fois le corpus traité, l'analyse fine des documents importants permettra de mettre en évidence les signaux forts et surtout les signaux faibles, et d'adopter des méthodes permettant de les recouper afin de leur donner du sens. **C'est une étape stratégiquement importante, où l'intervention humaine est absolument nécessaire.** Quelques outils peuvent cependant servir de support à ce travail.

Par exemple PUZZLE : ce prototype développé par l'ESA [GUIL94] permet de traiter les signaux faibles (regroupement des informations éparses, fragmentaires) dans le but de leur donner du sens. Le logiciel comporte plusieurs modules :

- saisie d'informations (à l'aide d'une fiche signalétique ; attribution d'un degré de fiabilité et d'importance) et mémorisation dans une base ACCESS ;
- construction de la requête (sélection des termes dans une liste de critères prédéfinie) et interrogation de la base ;
- structuration des informations sous forme de puzzle (graphe formé de noeuds qui sont les informations issues de la sélection) ;
- gestion des informations et des fichiers. [LESC97]

## 5. DIFFUSION et CONSULTATION

L'étape de la diffusion est essentielle car l'information doit alors pouvoir devenir décisionnelle.

A part la diffusion orale ou sur support papier, on peut choisir entre plusieurs modes de mise à disposition de l'information stratégique. Il faudra auparavant cibler les différents publics concernés, et synthétiser l'information (par exemple sous forme de lettre, de note, de page Web, etc. -il existe des outils pour cette phase, par exemple Isiman, de la société Keyword).

- Le courrier électronique est bien sûr l'outil le plus rapide pour envoyer cette information, mais on peut aussi opter pour les groupes de communication ou les forums sur Internet. Mais attention à la confidentialité. A ce niveau, il peut être utile d'avoir recours aux logiciels de cryptographie.
- Il est également possible d'utiliser, pour un certain type de public, les logiciels de push pour la diffusion d'informations déjà synthétisées et stockées dans une ou plusieurs bases internes.

On peut les diviser en trois catégories [RINA97] :

- push générique : chaque utilisateur sélectionne les sites pour lesquels il veut obtenir des mises à jour automatiques. *Exemple : Pointcast.*

- push avec filtrage : les profils des utilisateurs sont stockés sur le serveur, qui collecte les informations internes et externes, et les distribue en fonction de la demande. *Exemples : Castanet (Marimba), Backweb, Intermind.*

- push avec publication et abonnement : les administrateurs créent des profils utilisateurs ; les mises à jour sont transmises automatiquement. *Exemple : Tibco.*

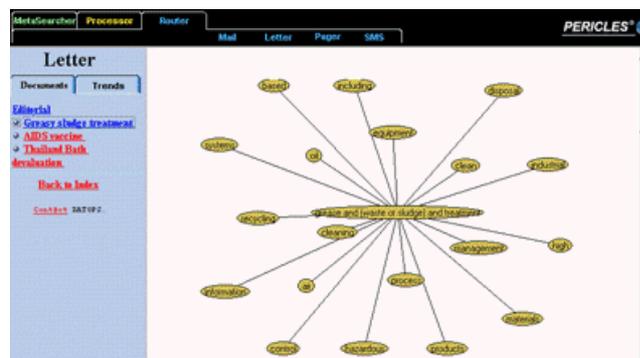
- ◆ Le choix de ces différentes techniques repose en fait sur les méthodes de diffusion que l'on souhaite mettre en place dans l'entreprise. Certains abonnements permettent de recevoir l'information par courrier électronique (par exemple une revue de presse), la présentation la plus courante étant une fenêtre qui se substitue à l'économiseur d'écran et affiche les dépêches [ANDR98].

## 6. OUTILS "MULTICATEGORIES" (SYSTEMES)

L'évolution de l'intelligence économique se dirige vers le développement d'outils intégrant plusieurs fonction, couvrant plusieurs étapes du processus de veille : collecte, traitement et diffusion de l'information. En voici quelques exemples :

PERICLES (www.datops.com) En fonction d'un profil (dans le sens de "thème"), le logiciel lance plusieurs agents simples d'une durée de vie prédéfinie (à l'aide de l'outil de recherche MetaSearch). Les sources peuvent être de tout type (internet, bases de données internes, texte, etc.). Datops a créé la notion d'InfoWarning, qui détecte les termes rattachés à ceux de la requête. Les résultats sont analysés statistiquement afin de produire des cartes, des graphiques, et mettre ainsi en avant les thèmes dominant et les évolutions. Il est alors possible, à partir de la carte de connaissance ("carte sémantique"), d'affiner la recherche en ne sélectionnant qu'une partie des termes.

*Exemple de carte de connaissance produite par Périclès, où l'on voit bien les termes qui "rayonnent" autour d'un thème central de recherche.*



- ◆ Ce type de logiciel, s'il permet de visualiser sans problème les thèmes dominants (l'information "triviale"), ne fait pas ressortir assez clairement les pôles "marginaux", qui sont souvent des termes émergents dans un domaine, donc des "signaux faibles" en veille technologique. Cependant, l'avantage de Périclès est de pouvoir suivre l'évolution d'une recherche dans le temps.

IM4U (Intranet Miner For You, société Arisem - [www.arisem.com](http://www.arisem.com)) : il s'agit d'un "système de veille automatisée" combinant une triple dimension (push, recherche documentaire et analyse sémantique). IM4U recherche l'information dans le fonds documentaire interne et sur internet. L'utilisateur peut recevoir par courrier électronique les résultats de son profil de veille qui seront stockés sur l'intranet et deviendront disponible pour les autres utilisateurs.

APERTO LIBRO ([www.inforama.fr/al1.htm](http://www.inforama.fr/al1.htm)): développé par Inforama International, société de conseil en veille et intelligence économique, c'est un ensemble complet d'outils et de services permettant de collecter, indexer, rechercher, organiser et trier l'information utile. Ce système comprend entre autre un moteur de recherche sémantique (NOEMIC98, nouvelle version de Taiga) et un logiciel gestionnaire de connaissances (Wincite).

La base de donnée constituée peut être vue sous plusieurs angles ; le croisement des critères permet d'obtenir des fiches de renseignements, et les fonctions statistiques présentent des vues graphiques de la recherche. Les données brutes ainsi collectées sont filtrées en fonction des paramètres entrés par Wincite.

## CONCLUSION

---

Les différentes catégories d'outils utilisables dans le domaine de l'intelligence économique mettent en évidence le long processus permettant à l'information d'obtenir une véritable valeur ajoutée et ainsi devenir stratégique pour l'entreprise. Il est donc très important de bien choisir l'enchaînement de ces divers outils pour optimiser cette activité. Cependant, il faut se souvenir qu'en aucun cas le traitement automatique ne remplace l'intervention humaine. En effet, les experts *"sont seuls juges en matière d'analyse de contenu et de synthèse : les logiciels n'auront servi qu'à leur faciliter le travail. Seule l'expertise humaine permet d'établir véritablement la nature et les retombées stratégiques de l'information mise en évidence par l'ensemble des traitements"* [THIL98].

Pour certaines entreprises, la recherche d'informations stratégiques se limite à la constitution régulière de dossiers couvrant un thème donné. Pour d'autres, la simple analyse de ce qui a été fait ne suffit pas.

Lors de la conception d'un système de veille, il faut tenir compte des publics concernés, et des différents types de surveillance à mettre en oeuvre [CAST94]. Par exemple, on peut définir deux types de traitements :

- un premier assez simple, qui recouvre la collecte (téléchargement de documents par exemple), le reformatage, la mémorisation et la diffusion des informations.
- un second plus complexe, qui inclue un traitement linguistique et statistique (outils bibliométriques), pour des volumes importants.

De même, les opérations de surveillance peuvent être de trois ordres :

- régulière, c'est à dire de fond (ex.: recherche sur les brevets, sur des bases de données spécifiques...)
- besoin immédiat : dossier sur un thème précis
- renseignement informel : fiches techniques des produits de la concurrence, études de marché, etc.

A chacune de ces étapes correspond un ou des traitements particuliers.

C'est pour ces raisons que le système mis en place doit être souple et offrir la possibilité d'être utilisé globalement ou par modules. Certains outils pourront être manipulés directement par les utilisateurs (les traqueurs par exemple), d'autres le seront par des spécialistes (ex. les outils bibliométriques).

On peut aussi se poser le problème des différents corps de métiers intervenant dans le processus de veille : documentalistes, informaticiens, spécialistes de bibliométrie, experts, etc. Or *"faire collaborer tous les métiers n'est pas facile"*, d'où la nécessité de mettre en place un système bien *"orchestré"* par un médiateur, qui donne accès à tout ou partie des outils et des méthodes disponibles [THIL97b].

D'autre part, lors de la mise en place d'un système de veille stratégique, il est important d'évaluer au préalable l'usage et la valeur de l'information pour les utilisateur [WERN98]. A quel(s) type(s) de public s'adresse-t-on ? Quels sont leurs besoins respectifs ? L'acquisition et le développement de technologies adaptées sera fonction de la réponse à ces questions.

*"La veille technologique est un processus cyclique"* [DESV92]. Pour respecter ce processus, l'idéal est que les outils utilisés soient intégrés à un système complémentaire, modulaire et souple. Bien articulées avec le travail des experts, ces technologies peuvent contribuer au bon développement de l'intelligence économique de l'entreprise.

## BIBLIOGRAPHIE et ADRESSES WEB

---

[ANDR98]

*"Recherche d'information : le push fait l'actualité"*

ANDRIEU O.

Technologies Internationales n°40 décembre-janvier 1998, p. 3-6

[ARON97]

*"L'info, nerf de la guerre" - Dossier Intelligence Economique*

ARON Patrice et PETIT Catherine

Le Monde Informatique n°731 29 août 1997

[BONN97]

*"Comment faire de la veille online"*

BONNET C. et MACARY J.F.

Internet Professionnel n°11 juillet 1997, p. 72-75

[BOUQ95]

*"Système de veille stratégique au service de la recherche et de l'innovation de l'entreprise :*

*principes-outils-application"*

BOUQUET Valérie

Thèse de doctorat - Université Aix Marseille 3, 1995

[CAHI97]

*"Choisir un outil de recherche textuelle"*

CAHIER J.P.

Le Monde Informatique n°737 10 octobre 1997

[CAST94]

*"Conception et installation d'un système de veille technologique : application au domaine pétrolier"*

CASTANO Eric

Thèse de doctorat - Université Aix Marseille, 1994

[CEIV98]

*"Centre d'Expertise et de Veille Inforoutes et Langues"*

<http://www.ceveil.qc.ca>, 1998

[CHAN98]

*"Les agents intelligents aident l'internaute à faire le tri des informations"*

CHANIA Emmanuel

01 Informatique n°1487 27 février 1998, p. 19

[COUE96]

*"Les logiciels d'analyse de contenu"*

COUEPEL Yann

<http://www.ceveil.qc.ca/traitem/couepel/intro.htm>, 1996

[CRON94]

*"The internet and competitive intelligence : a survey of current practice"*

CRONIN B, OVERFELT K. FOURCHEREAUX K et al.

International Journal of Information management vol. 14 n°3 1994, p. 204-222

[CVTC98]

*Club Veille Technologique et Concurrentielle*

<http://www.cpod.com/cvtc>, 1998

[DESJ90]

*"Veille technologique : revue de littérature et étude de terrain"*

DESJARDIN-BOURSIER R., MAYERE A., MUET F. et SALAUN J.M

CERCI, ENSB, Villeurbanne, 1990

[DESV92]

*"La veille technologique : l'information scientifique, technique et industrielle"*

DESVALS H. et DOU H.

Dunod, Paris, 1992

[DOU93]  
*"La veille technologique"*  
DOU Henri, HASSANALY Parina, QUONIAM Luc et LA TELA Albert  
Les cahiers du Leran - Sciences de la Société n°29 mai 1993, p. 213-216

[DURA93]  
*"Pour une approche intelligente de l'intelligence stratégique"*  
DURAND Thomas, de BRABANT Charles et GOMEZ Emmanuel  
IDT93

[GAZE98]  
*"Sés@me"*  
Marianne GAZEAU Consultants  
<http://www.opensesame.tm.fr>, 1998

[GUER98]  
*"Soyez plus productifs avec Internet"*  
GUERIN Cécile et MAHE Thierry  
Industries et Techniques n°789 janvier 1998 p. 50-57

[GUIL94]  
*"Guide de l'utilisateur Puzzle"*  
GUILLAUME P. et NEGRE C.  
Travail DESS MSIO, ESA, Grenoble, 1994

[HOHH94]  
*"Developping information systems for competitive intelligence support"*  
HOHHOF B.  
Lybrary Trends vol. 43 n°2 1994, p. 226-238

[IBM98]  
*"Data et Text Mining"*  
ECAM (Centre Européen de Mathématiques Appliquées d'IBM)  
<http://www.fr.ibm.com/france/ecam/expert.htm>, 1998

[IRIT95]  
*"Veille stratégique, scientifique et technologique : textes des communications"*  
Colloque IRIT-PEROMIP  
Institut de Recherche en Informatique de Toulouse, Toulouse, 1995

[JAKO92]  
*"Exemples commentés de veille technologique"*  
JAKOBIAK F.  
Editions d'Organisation, Paris, 1992

[JEAN97]  
*"Intelligence économique : la guerre de l'information est déclarée"*  
JEANNE François  
Le Monde Informatique n°705 janvier 1997

[LARD97]  
*"Recherche d'information sur l'internet : outils et méthodes"*  
LARDY Jean-Pierre  
ADBS Editions, 1997  
<http://www.adbs.fr/adbs/viepro/sinfont/lardy/risi.htm>

[LESC97]  
*"Des outils au service de la veille stratégique"*  
LESCA Humbert et ROUBAH K.  
Systèmes d'information et management n°2 vol. 2 1997, p. 101-132

[MAUR96]  
*"Les logiciels de clusterisation : un aperçu de leur utilisation"*  
MAURO Sylvia  
<http://www.ceveil.qc.ca/traitem/mauros/intro.htm>, 1996

[NAJM97]  
*"Des outils à la portée de l'entreprise"*  
NAJMAN M.  
Le Monde Informatique n°705 janvier 1997

[NOYE95]

*"Bibliométrie, scientométrie, infométrie"*

NOYER Jean-Max (sous la direction de)

<http://www.info.unicaen.fr/bnum/jelec/Solaris/d02/index.html>, 1995

[OBER97]

*"L'internet et l'intelligence économique - méthode de veille économique par l'internet"*

OBERSON Philippe

Editions d'Organisation, Paris, 1997

[PAST98]

*"Introduction à la bibliométrie"*

Bibliothèque de l'Institut PASTEUR

<http://www.pasteur.fr/units/biblio/Metrie/index.html>, 1998

[POLA93]

*"Analyse stratégique de l'information scientifique et technique - construction de clusters de mots-clés"*

POLANCO Xavier

Les cahiers du Leran - Sciences de la Société n°29 mai 1993, p. 111-126

[PRAX97]

*"Les nouvelles technologies au service de l'intelligence collective"*

PRAX J.Y et al.

IDT 97, p. 102-106

[REMI98]

*"Forum européen 97 de l'intelligence économique"*

REMIZE Michel

Archimag n°110 décembre janvier 1998, p. 21-22

[REYN90]

*"Le développement de l'entreprise par la veille technologique"*

REYNE Maurice

Hermès, Paris, 1990

[RINA97]

*"Technologies, services, stratégies, réglementation"*

RINAUDO J.

Internet et services online 31 mars 1997

[ROIR97]

*"Enquête sur la veille stratégique dans les entreprises britanniques"*

ROIRON Lucile et LESCA Humbert

Technologies Internationales n°32 mars 1997, p. 40-43

[ROST93]

*"Veille technologique et bibliométrie : concept, outils, applications"*

ROSTAING Hervé

Thèse de doctorat - Université Aix Marseille 3, 1993

[ROUS97]

*"Veille et informatique : des besoins aux solutions"*

ROUSSEAU-HANS Françoise et THIL Jérôme

Technologies Internationales n°39 novembre 1997, p. 33-36

[SCHW96]

*"Competitive intelligence primer"*

SCHWARTZWALDER R.

Database 19 (4) Aug/sept 1996, p. 89-91

[SFBA95]

*"Les systèmes d'information élaborée - Journées organisées par la SFBA"*

Société Française de Bibliométrie Appliquée

SFBA, Ile Rousse, 1995

[THIL98]

*"Une méthode nouvelle pour cartographier des coopérations"*

THIL Jérôme

Technologies Internationales n°41 février 1998, p. 25-31

[THIL97a]

"Veille stratégique : un système analyse les réseaux de collaboration"

THIL Jérôme

Technologies Internationales n°32 mars 1997, p. 26-31

[THIL97b]

"Veille et informatique : optez pour un médiateur"

THIL Jérôme

Technologies Internationales n°34 mars 1997, p. 33-36

[WERN98]

*"Une approche économique de l'information"*

WERNER E.

Technologies Internationales n°40 décembre-janvier 1998, p. 39-40

Ainsi que de nombreux documents publicitaires ou de présentation des outils, obtenus par l'intermédiaire de collègues de travail ou des fournisseurs.