

S.F.B.A.

ÎLE ROUSSE

27 septembre au 1 octobre 1999

Les systèmes d'information élaborée

Bibliométrie – Linguistique – Information Stratégique – Veille Technologique
Intelligence économique



PROGRAMME

Journées d'études organisées par

La Société Française de Bibliométrie Appliquée

Du 27 Septembre au 1^{er} Octobre 1999

Avec le soutien de :

La Municipalité d'Île Rousse

L'Université d'Aix - Marseille (CRRM)

L'Université de Marne - la - Vallée

Le Centre de Documentation de l'Armement (CEDOCAR)

Le Centre d'Etudes Scientifiques de Défense (CESD)

COMITÉ SCIENTIFIQUE DU PROGRAMME

J.P. DEBREUIL (CEDOCAR)

H. DOU (CRRM)

B. DOUSSET (IRIT)

J. DUCLOY (CRIN-Nancy)

J.E. DUBOIS (Paris VII)

C. DUTHEUIL (Synthélabo)

J. GUYAUX (CESD)

P. LACOSTE (CESD)

D. LAURENT (U.M.L.V)

Y. LE COADIC (CNAM)

F. LEPINE (FED)

C. LONGEVIALLE (DGA)

B. MARX (INPI)

F. MARCOTORCHINO (IBM)

C. PAOLI (CESD)

L. QUONIAM (CRRM)

R. SOLER (CERE-CERAM)

W. TURNER (CERESI)

MARDI 28 SEPTEMBRE 1999

- 9h00 : Accueil des participants par la Municipalité d'Île Rousse
9h30 : Séance d'ouverture : H. DOU J.E. DUBOIS J. GUYAUX
10h30 : Pause

STRATEGIE & CONTEXTE

Président : H. DOU

- 11h00 : Etude de définition du contenu et du mode de fonctionnement « d'un pôle de valorisation » de la recherche en matière de bio-industries. p 6
P.Y BACLE, B. DULOU, H. DOU (GIE Santé Valor ; CRSSA ; CRRM)
- 11h30 : The Brazilian trajectory in the age of the Competitive Intelligence and the Knowledge Management. p 8
G. BATTAGLIA ; A.V WANDERLEY ; J.R JARDIM ; N.F LIMA ; L. BATTAGLIA (INMETRO ; FINEP)
- 12h00 : L'impact socio-économique des Nouvelles Technologies de l'Information et de la Communication (NTIC) – Menaces et Opportunités. p 10
B. DULOU ; R. CATERINI ; J. D'OLLIER ; J.P HAMON ; G. PELLON (CRSSA)
- 12h30 : Déjeuner

STRATEGIE & CONTEXTE (suite)

Président : J. GUYAUX

- 14h00 : Pour une culture de défense ou vers une culture de sécurité économique ? p 11
J.P HAMON ; B. DULOU ; H. DOU (CRSSA ; CRRM)
- 14h30 : Le déploiement de l'information au sein du centre STMicroelectronics de Tours. p 12
C. MARTY ; N. HOELTZI (STM ; CRRM)
- 15h00 : Identification de contraintes organisationnelles, humaines et techniques structurant un système d'information élaborée. p 14
Y. BIGOT ; S. ARTILLAN ; M. LALAUDE (Aérospatiale)
- 15h30 : Sensibilisation des cadres supérieurs du Ministère de la Défense à la veille stratégique et à l'intelligence économique dans les stages informatiques et management de l'information de la Direction de l'Administration Centrale. p 15
B. DULOU (CRSSA)
- 16h00 : Pause

STRATEGIE & CONTEXTE (suite)

Président : C. PAOLI

- 16h30 : Intelligence économique et décision. p 16
J. LINK-PEZET ; R. BERKANE (URFIST Toulouse ; LIHS Toulouse)
- 17h00 : De l'importance de la normalisation en bibliométrie. p 17
E. NAUER (LORIA/CNRS)
- 17h30 : Application des technologies Objets à la gestion de la Documentation d'une Entreprise. p 18
O. THIERY (LORIA/Université Nancy2)
- Représentation du complexe local et émergence d'un projet fédérateur. p 20
Y. BERTACCHINI ; P. DUMAS (IUT/Toulon ; CRRM)
- 18h30 : Management de données complexes et création d'information stratégiques. p 21
J.E DUBOIS ; A.F CUTTING-DEULLES ; K. FRAOUA (CODATA ; Université de Chambéry ; Université de Marne la Vallée)
- 19h00 : Fin de journée
- 19h30 : Réception en Mairie de l'Île Rousse.

MERCREDI 29 SEPTEMBRE 1999

METHODES & OUTILS

Président : L. QUONIAM

- 9h00 : **Automatic treatment of information – a case study on the Brazilian theses database.** p 22
C. CANONGIA, H. CHAVES, S. MAFFIA (IBICT)
- 9h30 : **Elaboration d'un réseau cartographique de codes CIB à partir de commandes statistiques en ligne.** p 24
A. DA SILVA, S. COLLET, C. BAINIER-GIRARD, L. QUONIAM, H. ROSTAING (CRRM)
- 10h00 : **Analyse différentielle des co-signatures dans les publications des chercheurs belges francophones.** p 25
M. DORBAN (Université catholique de Louvain)

10h30 : *Pause*

METHODES & OUTILS

Président : H. ROSTAING

- 11h00 : **Entropie et distributions bibliométriques.** p 26
T. LAFOUGE, C. MICHEL (RECODOC/Université Claude Bernard Lyon 1)
- 11h30 : **Une nouvelle méthode de recherche des tendances thématiques fondée sur un réseau de variations terminologiques.** p 28
F. IBEKWE-SANJUAN, E. SANJUAN (IUTA . LMD)
- 12h00 : **Utilisation de la représentation hypertextuelle en veille technologique.** p 31
LEVEILLE V. ; ROSTAING H. (CRRM)

12h30 : *Déjeuner*

METHODES & OUTILS

Président : F. MARCOTORCHINO

- 14h00 : **The usage of Metadex modifying terms for bibliometric analysis : Study on Piston Ring.** p 32
I.L. De FARIA, A.M. HOFFMANN, A.R GREGOLIN, L. QUONIAM (NIT ; CRRM)
- 14h30 : **Text Mining et Knowledge Management.** p 33
C. HUOT (IBM)
- 15h00 : **ATRIUM : Automates pour un Traitement de l'Information Unifié et Multilingue.** p 34
T. REGNIER, L. MARCELLI (CESD, CEDOCAR)
- 15h30 : **Les métriques et l'analyse relationnelle : visualisation en 4 dimensions.** p 35
A. BENNAMAR, B. DOUSSET (IRIT/équipe SIG)

16h00 : *Pause.*

METHODES & OUTILS

Président : C. DUTHEUIL

- 16h30 : **Vers une méthode par prototypage pour la conception de produits d'intelligence économique.** p 36
T. ZID, M. SALLES (Université Toulouse 1)
- 17h00 : **Les séquences communes. Une nouvelle approche de lemmatisation.** p 38
B. ROUDANI, V. PILLET, L. QUONIAM (CRRM)
- 17h30 : **Perception infographique.** p 39
L. CORNU, J.M. DOU (CRRM)
- 18h00 : **L'attractivité, un concept remplaçant la compétitivité dans un nouveau contexte.** p 40
G.M COELHO, H. DOU (I.N.T, CRRM)
- 18h30 : **Nouvelle approche de décomposition de données.** p 41
J.E DUBOIS, K. FRAOUA (CODATA, Université de Marne la Vallée)
- 19h00 : *Fin de journée.*

JEUDI 30 SEPTEMBRE 1999

METHODES & OUTILS

Président : Y. LECOADIC

- 9h00 : **SAMPLER : définition et application d'une plate-forme de text mining.** p 42
O. JOUVE, S. LAURIN (CS-CISI ; ADIT)
- 9h30 : **Visualisation interactive de classification d'informations.** p 43
S. KAROUACH, B. DOUSSET, T. DKAKI (IRIT/équipe SIG)
- 10h00 : **La mise en convergence des usages locaux. Vers l'élaboration d'un outil de veille territoriale.** p 44
P.H DUMAS, E. BOUTIN, Y. BERTACCHINI (CRRM)
- 10h30 : *Pause*

ANALYSE TEXTUELLE

Président : C. LONGEVIALLE

- 11h00 : **Tesouro de quimica en lingua portuguesa-tesquimica, una ferramenta de recuperacao** p 45
M.G.B BATTAGLIA (FINEP)
- 11h30 : **VIGITEXT : un nouvel outil d'extraction d'informations textuelles pour la veille technologique.** p 46
B. GOUJON (Bureau Van Dijk - CAMS/équipe LaLIC)
- 12h00 : **Analyse automatique de texte libre issu de résumés de publications en génétique : identification des gènes cités.** p 47
A. INGOLD, B. JACQ, L. QUONIAM (CRRM)
- 12h30 : *Déjeuner*

APPLICATIONS

Président : P. DUMAS

- 14h00 : **Complémentarité de l'analyse statistique et de l'analyse de données dans la détection et l'études de collèges d'experts.** p 48
L. LEDUC, C. LONGEVIALLE, C. PAOLI (ESIEE ; CESD)
- 14h30 : **Système d'information Internet et environnement. Description du système d'Information Environnement : de l'information statique à l'information dynamique.** p 49
M. LEITZELMAN, H. DOU, C. REY (Marseille Innovation ; CRRM)
- 15h00 : **Analyse automatique des courriers de réclamation des clients EDF-GDF** p 50
GOUTTAS C. ; OTTHOFFER D. ; PIAT G. (ECAM IBM ; EDF-GDF)
- 15h30 : **Utilisation d'outils infométriques pour la mise en évidence de la structuration d'un secteur d'activité économique.** p 50bis
F. BARON, C. POUZENC, C. PAOLI (I2P ; CESD)
- 16h00 : *Pause.*

APPLICATIONS (suite)

Président : M.G.B. BATTAGLIA

- 16h30 : **The Ibero-American region scientific production on the corrosion and anti-corrosive protection area.** p 51
N.J SILVEIRA VARGAS, C.N DA SILVA, G.M COELHO, L. QUONIAM, H. DOU (INT,CRRM)
- 17h00 : **Competitive intelligence system to small and medium-sized companies.** p 52
J.A MILANI, C. CANONGIA (BRASPETRO ; MCT/CNPq/IBICT)
- 17h30 : **Analyse stratégique d'un domaine scientifique du CNRS par des méthodes d'analyses bibliométriques.** p 53
B. RING, R. RING (I.R.e.S ; CESD)
- 18h00 : **Les biotechnologies appliquées aux processus de fermentation pour l'industrie agro-alimentaire : méthodes d'analyses bibliométriques pour dresser un état de l'art.**
V. LONGY, C. LONGEVIALLE, C. PAOLI (CEDOCAR ; CESD)
- 18h30 : *Fin de journée.*

VENDREDI 1^{ER} OCTOBRE 1999

INTERNET & BASES DE DONNEES

Président : B. DOUSSET

- 9h00 : **Création automatique d'une base de données brevets à partir de données récupérées de l'Internet.** p 54
I.L. DE FARIA, J.A.R. GREGOLIN, B. MANNINA (NIT/Matériaux ; CRRM)
- 9h30 : **DirectProfil : de la navigation au partage de documents hypermédia.** p 55
E. GIRAUD, C. DOU (CRRM ; LEPONT)
- 10h00 : **Evaluation des sites Internet. Construction d'indicateurs à partir des réseaux de « citations »** p 56
H. ROSTAING, E. BOUTIN (CRRM ; LEPONT)
- 10h30 : *Pause*

INTERNET & BASES DE DONNEES (suite)

Président : J. KISTER

- 11h00 : **Création automatique de bases de données préformatées : Information issue de l'Internet pour la Veille Technologique et l'Intelligence Economique.** p 58
B. MANNINA, L. QUONIAM (CRRM)
- 11h30 : **La conception de bases de données infométriques : une comparaison des pratiques dans quatre pays européens.** p 59
L. GRIVEL (INIST/CNRS)
- 12h00 : **La dynamique de la technologie Intranet au service d'un système d'information avec BdD interactives.** p 61
C. DOU, C. PIRONNEAU, E. GIRAUD (CRRM)
- 12h30 : **DATA CIB : un logiciel de liens automatiques entre références bibliographiques scientifiques et information technique de brevet.** p 62
P. FAUCOMPRES, L. QUONIAM, H. ROSTAING (CRRM)
- 13h00 : *Déjeuner*
- 14h30 : *Synthèse des journées.*
- 15h00 : *Séance de clôture.*

Autres Communications et Posters

Communications :

- **Intranet : La base du système d'information (Exemple de Gemplus)** p 63
C. DOU, P. LAURI, E. GIRAUD (CRRM, GEMPLUS, Conseil Régional Provence-Alpes-Côte d'Azur)
- **Observations et diagnostic des usages locaux – Vers l'élaboration d'un outil de veille territoriale.** p 64
Y. BERTACCHINI (CRRM – LEPONT)
- **Contribution à la veille juridictionnelle : exemple d'application de la veille stratégique au contentieux fiscal de l'exonération de l'impôt sur le revenu ou de l'impôt sur les sociétés pour les entreprises nouvelles.** p 65
M. LEBRETON, H.J.M. DOU (Tribunal administratif de Marseille, Université Aix-Marseille III)

Poster :

- **Analyse stratégique du domaine de recherche LIPIDES par la méthode des co-citations** p 66
E. BASSECOULARD-ZITT, M. BLONDEL, C. BOULEY, G. GANDEMER, PH. GUESNET,
F. HUMBLLOT, M. ZITT (INRA, Groupe DANONE, Centre International de Recherche Daniel Carasso)

Élaboration d'un réseau cartographique de codes CIB partir de commandes statistiques en ligne.

A. DA SILVA, S. COLLET, C. BAINIER-GIRARD,
L. QUONIAM, H. ROSTAING
CRRM ; SNECMA

Résumé

L'information technologique contenue dans les brevets est une information de grande valeur pour les entreprises. Aussi, l'accès aux bases de données brevets en ligne est il cher, en particulier lorsqu'on souhaite se procurer toutes les notices se rapportant à un domaine. Pour des raisons de coût, il est alors préférable de procéder autrement si l'on veut réaliser une étude brevet.

L'objet de cette communication est de décrire une méthode faisant appel à la bibliométrie, ainsi qu'aux commandes statistiques en ligne (nettement moins coûteuses que les notices intégrales) disponibles sur les centres serveurs, pour élaborer une cartographie précise de l'activité technologique d'un domaine.

Cette méthode procède en différentes étapes successives de sélection et de validation de l'information recueillie, par les experts du domaine étudié. Nous verrons l'importance de rendre l'information brute plus accessible à ces mêmes experts. Nous découvrirons aussi que cette collaboration permanente est un facteur essentiel pour la pertinence des résultats obtenus.

Mots Clés : cartographie, brevets, codes CIB, statistiques en ligne, bibliométrie, experts, théorie des ensemble.

The usage of Metadex modifying terms for bibliometric analysis : Study On Piston Ring

I.L. De FARIA¹, A.M. HOFFMANN¹, A.R. GREGOLIN¹, L. QUONIAM²
¹NIT/Materiais ; ²CRRM

Résumé

The Metadex and the Engineered Materials Abstracts bibliographic databases are very useful in Materials Science and Engineering because they provide comprehensive coverage of international metals, polymer, ceramic and composite literature and gives a vision of the scientific and technological developements of metallurgy and materials. The Metadex and the Engineered Materials Abstracts keywords are necessary in the Thesaurus of Metallurgical Terms and Thesaurus of Engineered Materials, respectively. These Thesaurus include the same controlled list of terms which may be linked to primary index terms as modifying terms. All modifying terms fit in one of the four categories : processes, properties, effects or forms. The knowledge of the relations : structure-properties-processes is of great importance for the materials area. The subdivision of the Metadex and Engineered Materials Abstracts Descriptors field from its Modifying Terms, properties, processes, effects and forms, is a way of amplifying the possibilities of bibliometric analysis of information about materials. The inter-linkage of the Modifying Terms, among them and among others descriptors allowed to extract important information. The method was used for the information analysis on piston ring, specifically in relation to material used in their fabrication.

Mots Clés : Automated Information Treatment, Bibliometric Techniques, Technological Information, Piston Rings, Materials Engineering

Les séquences communes. Une nouvelle approche de lemmatisation.

B. ROUDANI¹, V. PILLET¹, L. QUONIAM^{1,2}
¹CRRM, ²IUT SRC

Résumé :

Une des principales difficultés que rencontrent les spécialistes de l'analyse de l'information textuelle est sans doute le problème des variantes graphiques que présentent les mots dans les corpus textuels. Cela implique que pour un mot donné, nous pouvons rencontrer plusieurs formes graphiques différentes. Le poids de ce concept est réparti sur plusieurs formes et son importance statistique est donc minimisée.

Avant d'entreprendre tout traitement statistique sur un corpus de texte, il est donc indispensable de soumettre les formes graphiques à une homogénéisation morphologique et sémantique.

Généralement, les outils informatiques utilisés pour ce type de traitement s'appuient sur des dictionnaires de termes préétablis. Cela implique que l'efficacité du traitement dépend de la richesse de ces dictionnaires. Le principal inconvénient de ces outils est donc leur forte dépendance de la langue et du domaine traité.

Nous avons développé pour cela une nouvelle approche qui a abouti à la réalisation d'un outil "**DATALEM**". Celui-ci permet à partir d'un corpus de regrouper les formes graphiques ayant la plus longue chaîne (ou multichaîne) de caractères communes (préfixe, suffixe, sous-chaînes ou multiparties communes). Son référentiel de base est le corpus à traiter plutôt qu'un dictionnaire terminologique. Cette approche nous permet d'éviter la non-exhaustivité des dictionnaires dans un domaine précis et d'être indépendant de la langue. Elle permet aussi d'identifier les formes graphiques mal orthographiées. A l'aide d'un expert du domaine, cet outil permet donc d'homogénéiser les formes graphiques.

Nous exposerons dans cet article les problèmes liés à la dispersion du vocabulaire ainsi que les différentes approches existantes pour résoudre ce problème. Nous présenterons ensuite la réalisation d'un nouvel outil de lemmatisation "**DATALEM**". Puis nous comparerons les résultats obtenus par cet outil avec une autre méthode dépendant d'un dictionnaire terminologique.

Mots Clés : Lemmatisation, analyse linguistique, séquences communes, dictionnaire terminologique, distribution statistique.

Analyse automatique de texte libre issu de résumés de publications en génétique : identification des gènes cités.

A. INGOLD¹, B. JACQ², L. QUONIAM¹
¹CRRM, ²CNRS-Luminy

Résumé

Nous présentons une étape d'un processus d'extraction de connaissances à partir de textes. Ces textes sont des résumés d'articles scientifiques répertoriés dans une base de donnée médicale. Le but ultime est de construire une base de données sur un phénomène biologique précis : les interactions génétiques. Des faits concernant ce phénomène sont énoncés dans les textes que nous analysons. Les phrases impliquées doivent être détectées puis analysées. Pour réaliser cette tâche, il faut être capable de reconnaître les gènes cités. Ce travail, actuellement conduit chez la mouche du vinaigre, sera étendu à la souris et à l'homme.

Le lexique du domaine scientifique que nous étudions est très étendu : on compte 25000 noms de gène différents. De plus, la terminologie est complexe puisqu'un même gène peut avoir plusieurs noms (synonymie) et un même nom peut désigner plusieurs gènes (homonymie). Cela rend difficile l'identification des gènes dans du texte. La situation est compliquée par la présence de mots vides dans le lexique : certains noms de gènes sont aussi des mots du vocabulaire courant.

Afin de détecter ces mots vides dans le lexique, nous avons utilisé un corpus de texte où les noms de gènes sont désambiguïsés. Le marquage du sens nous a permis de calculer un indicateur d'ambiguïté pour chaque terme du lexique. Cet indicateur estime la probabilité qu'a le terme de désigner un gène. Nous avons calculé la répartition des valeurs prises par cet indicateur. L'indicateur nous sert à isoler les noms de gènes à problème, sur lesquels nous appliquons des méthodes de désambiguïsation.

Certaines de ces méthodes sont basées sur la redondance d'informations dans les textes que nous étudions : quand un gène est le sujet d'étude, il est cité plusieurs fois et, en général, de plusieurs manières (pour des motifs qui sont propres au domaine).

D'autres méthodes utilisent le fait qu'un nom de gène est souvent (mais pas toujours) suivi du terme *gene*, *locus* ou *protein*.

Notre travail nous semble facilement transposable à d'autres domaines que la génétique, la chimie particulièrement, ainsi que toute discipline où il existe une terminologie étendue, complexe et évolutive.

Mots Clés : traitement automatique de l'information, statistique textuelle, application en biologie, extraction de connaissances, dictionnaire spécifique du domaine, ambiguïté sémantique, détection automatique de mots vides, désambiguïsation par un corpus d'exemples, statistique de l'ambiguïté sémantique, désambiguïsation par le contexte.

The Ibero-American region scientific production on the corrosion and anti-corrosive protection area

N.J. SILVERIA VARGAS¹, C.H. DA SILVA¹, G.M. COELHO¹
L. QUONIAM², H. DOU²
¹INT, ²CRRM

Résumé :

Preliminary study aiming the mapping of Ibero-American literature productivity on corrosion and anti-corrosive protection area in order to analyse the competitiveness of the countries that integrate the Ibero-American Technological Information Network on Corrosion - RICORR. It was used the technological monitoring methodology that is based in bibliometric analyses and employs automatic analysis tools. It was possible to build an overview of the area in the period 1980-1997, from an American database point of view.

Mots Clés : Bibliometrics, Corrosion Literature, Ibero-american Region, Automatic analysis, Technological monitoring.

Création automatique d'une base de données brevets à partir de données récupérées de l'Internet

L.I.L. DE FARIA^{1,2}, J.A.R. GREGOLIN¹,
L. QUONIAM², B. MANNINA²
¹NIT/Materiais, ²CRRM

Résumé :

Depuis 1992, la base de données brevets brésiliennes Braspat, est accessible sur le site Internet de l'INPI Brésil - Instituto Nacional da Propriedade Industrial. La Braspat contient des données bibliographiques concernant les brevets déposés au Brésil et chaque brevet peut être téléchargeable via l'Internet.

Toutefois, la Braspat ne permet pas le téléchargement massif des enregistrements. La récupération manuelle d'un brevet rend difficile, voire impossible une étude bibliométrique.

L'objectif de ce travail a été de récupérer les données de la base Braspat et de les réorganiser en une base de données locale. Environ 50.000 brevets publiés par l'INPI Brésil, concernant la période 1992-1997 ont été récupérés. Le téléchargement de ces données a été fait automatiquement, grâce au logiciel AURESYS - AUtomated REsearch SYstem, réduisant ainsi considérablement le temps nécessaire à la construction de la base de brevet locale.

Mots Clés : internet, base de données brevets, traitement bibliométrique

**Création automatique de bases de données préformatées :
Information issue de l'Internet pour la
Veille Technologique et l'Intelligence Economique**

B. MANNINA, L. QUONIAM
CRRM

Résumé :

Le réseau Internet connaît depuis quelques années une croissance exponentielle de sa quantité d'informations. Rapide et efficace, Internet offre à ses utilisateurs un moyen de communication et surtout de récolte d'information. Internet est donc en train de s'imposer comme source d'information incontournable pour les spécialistes.

Avoir accès à toute l'information disponible sur Internet ne suffit plus aux spécialistes. Ce foisonnement d'informations est très difficilement exploitable. La récolte et le reformatage des données s'avère essentiel pour une bonne veille technologique. L'utilisation de logiciels, appelés « Agents Intelligents », est nécessaire pour la phase de récolte mais peu d'entre eux structurent les informations recueillies afin d'avoir un corpus de documents directement analysable.

Nous allons dans cet article vous montrer comment, grâce au logiciel Auresys, il est possible de récolter et formater automatiquement les informations issues de l'Internet.

Mots clés : Internet, agents intelligents, informations, traitement, collecte, récolte, veille technologique, analyse automatique.

DATA CIB : un logiciel de liens automatiques entre références bibliographiques scientifiques et information technique de brevet

P. FAUCOMPRES, L. QUONIAM, H. ROSTAING
CRRM

Résumé :

A travers cette communication, nous présentons un nouveau logiciel, DataCIB, destiné à créer automatiquement un lien entre information scientifique et information technique. Le principe de ce programme consiste à créer dans des références bibliographiques de type scientifique un nouveau champ documentaire contenant des codes issus de la Classification Internationale des Brevets. Son fonctionnement est basé sur l'exploitation des mots clés décrivant le contenu thématique de ces références et leur rapprochement avec les index de catchwords de la classification spécifique à la Propriété Industrielle.

Une première étude, présentée lors des précédentes Journées, en avait exposé la faisabilité technique et ses grands principes méthodologiques. Cette étude restait cependant limitée à l'utilisation d'un seul langage documentaire, à savoir les termes d'indexation contrôlée de la base Pascal de l'INIST. Le produit qui sera présenté en juin marque une évolution importante dans la méthode adoptée puisqu'il s'affranchit de toute construction préalable de tables de correspondance. Par conséquent, la même procédure d'établissement de liens peut être appliquée aux références issues de n'importe quelle banque de données bibliographiques, sans consentir aucun investissement préalable. Des exemples concrets illustreront cette démarche et montreront comment une simple présomption de liens entre science et technologie présente un intérêt potentiel important dans le cadre d'actions d'intelligence économique.

Mots-clés : Banques de données bibliographiques, Brevets d'invention, Catchwords, Classification Internationale des Brevets, Correspondance automatique, DataCIB (logiciel), Indexation automatique, Science, Technologie.