



HAL
open science

La bibliométrie comme instrument d'aide à la sélection et à l'expertise lors d'une étude de veille technologique. Comment réduire les coûts de collecte dans les bases de données de brevets tout en conservant la qualité d'une étude de veille

Hervé Rostaing, Alain Baroni

► To cite this version:

Hervé Rostaing, Alain Baroni. La bibliométrie comme instrument d'aide à la sélection et à l'expertise lors d'une étude de veille technologique. Comment réduire les coûts de collecte dans les bases de données de brevets tout en conservant la qualité d'une étude de veille. *Veille Stratégique, Scientifique & Technologique*, VSST'2010, Université Paul Sabatier, IRIT, Oct 2010, Toulouse, France. hal-01539876

HAL Id: hal-01539876

<https://hal.science/hal-01539876v1>

Submitted on 15 Jun 2017

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

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

LA BIBLIOMETRIE COMME INSTRUMENT D'AIDE A LA SELECTION ET A L'EXPERTISE LORS D'UNE ETUDE DE VEILLE TECHNOLOGIQUE.

Comment réduire les coûts de collecte dans les bases de données de brevets tout en conservant la qualité d'une étude de veille.

[Hervé ROSTAING](#) (*), [Alain BARONI](#) (**)
herve.rostaing@univ-cezanne.fr, alain.baroni@framatech.fr

(*) [LSIS UMR CNRS 6168](#), Université Paul Cézanne Aix Marseille III, Faculté des Sciences et Techniques, av. Escadrille Normandie-Niémen 13397 Marseille Cedex 20 (France),

(**) [FRAMATECH SA](#), Bâtiment Europarc D, Rue John Meynard Keynes, Technopôle Château Gombert, 13013 Marseille (France).

Mots clefs :

Veille technologique, bibliométrie, classification automatique des données, aide à l'expertise, étude technique, réduction des coûts

Keywords:

Competitive technical intelligence, bibliometrics, clustering analysis, expertise assistance, technical study, cost reduction

Palabras clave :

Vigilancia tecnológica, bibliometría, análisis cluster, ayuda al peritaje, estudio técnico, reducción de costos

Résumé

Cette communication a pour but d'exposer comment une démarche bibliométrique a été utile au processus de collecte et d'analyse de l'information brevet lors d'une étude de veille technologique. L'expérience relatée dans ce travail est tirée d'une étude de veille technologique réalisée par un professionnel de l'information pour le compte d'un client. Le professionnel de l'information, utilisant la base de données Derwent WPI pour cette étude, a été immédiatement confronté à la problématique suivante : comment respecter l'enveloppe budgétaire fixée par son client tout en garantissant la pertinence et la bonne couverture du sujet étudié par les données extraites de cette base de données payante. La démarche mise en œuvre fut de télécharger les notices au format gratuit répondant à la stratégie de recherche élaborée pour couvrir au mieux le domaine à étudier, puis d'accompagner la phase de validation/sélection de ces notices gratuites par l'assistance d'une démarche bibliométrique pour au final ne télécharger au format complet et payant que l'ensemble des notices les plus pertinentes pour l'étude technique des brevets du domaine étudié. Le traitement bibliométrique mis en œuvre pour la phase de sélection avait pour objectif d'élaborer une grille de lecture des 394 notices au format gratuites pour assister l'expertise et la sélection par un expert du domaine. Un regroupement thématique de ces notices a été obtenu par une méthode de classification automatique (K-means clustering), puis le résultat de ce regroupement thématique a été remonté sur un support facilement exploitable par l'expert et le professionnel de l'information pour rapidement converger une sélection de notices pertinentes. Cette démarche a permis à la fois de réduire de 50% le coût de collecte des ces informations brevets mais aussi de construire une nouvelle grille de caractérisation technique des brevets (regroupement des domaines techniques couverts par les brevets) très utile pour l'étude technique de ces brevets. La conduite de cette étude illustre également l'importance de la mise en commun de multiples compétences pour ce genre d'étude, à savoir un professionnel de l'information, un spécialiste du traitement automatique de l'information et un expert technique du domaine.

1 Introduction

Le besoin d'instruments de traitement automatique de l'information lors d'une étude de veille technologique est fréquemment avancé dans la littérature. La principale raison de ce besoin est la facilité d'accès à de très grands volumes de données scientifiques et technique à expertiser lors d'une telle étude grâce aux sources d'information que sont les bases de données bibliographiques.

La littérature évoquant l'application de la bibliométrie en veille technologique est très riche allant de la présentation des techniques, outils, logiciels pouvant être employés jusqu'à la présentation des résultats obtenus lors de la mise en œuvre de ces outils à des fins de veille technologique [1], [2], [3], [4], [5], [8].

Cette communication a plus la volonté d'exposer plus précisément comment une démarche bibliométrique a été utile au processus de collecte et d'analyse de l'information brevet lors d'une étude de veille technologique.

Cette expérience de confrontation de l'instrument bibliométrie aux besoins de traitement de l'information technique en veille technologique est tirée d'une étude de veille technologique réalisée par un professionnel de l'information pour le compte d'un client. Ce professionnel a été confronté à la difficulté d'optimiser la phase de collecte des données dans une base de données sur les brevets pour réduire le coût de cette collecte tout en garantissant la pertinence de l'étude finale.

2 Problématique

Les bases de données bibliographiques qui référencient les travaux scientifiques ou les demandes de brevet sont des sources incontournables lors d'une étude de veille technologique. L'exhaustivité thématique, géographique et temporelle ainsi que la qualité des données proposées par ces sources permettent de garantir la pertinence des résultats de l'étude de veille technologique à condition de savoir extraire le corpus de données représentant parfaitement bien le domaine scientifique ou technique à étudier.

La contrepartie d'une telle pertinence est le coût de telles sources d'information. Si les bases de données scientifiques restent à des prix relativement raisonnables, cela n'est pas toujours le cas pour les bases de données des brevets disponibles sur les serveurs professionnels d'informations scientifiques et techniques (Dialog, Questel, STN...).

Un professionnel de l'information utilisant les bases de données sur les brevets peut rapidement être confronté au dilemme suivant : respecter l'enveloppe budgétaire fixée par son client tout en garantissant la pertinence et la bonne couverture du sujet par les données extraites des bases de données. Les systèmes d'interrogation de ces bases de données permettent de parfaire la stratégie de recherche en multipliant et combinant les descripteurs du domaine scientifique et technique à étudier. Par contre, la visualisation des notices identifiées par les requêtes étant très coûteuse, ces stratégies d'interrogation sont la plupart du temps construite à « l'aveuglette » et le corpus obtenu lors du téléchargement peut occasionner quelques mauvaises surprises et imposer alors une nouvelle interrogation de la base de données et donc un surcoût.

Une méthode alternative est de visualiser dans un format gratuit les notices obtenues par la stratégie d'interrogation. Chaque base de données offre au moins un format permettant de visualiser gratuitement les notices mais ce format est constitué d'un nombre très limité de champs bibliographiques. Comme exemple, la Figure 1 présente le format gratuit proposé par le serveur Questel pour la base de données brevet Derwent World Patent Index (DWPI).

```
1/394 DWPI - (C) The Thomson Corp
AN - 99-164791 [14]
XA - C99-048035
XP - N99-120761
TT - **THERMAL** **STORAGE** DEVICE DOMESTIC **WATER**
    **HEAT** AIR CONDITION HEAT TRANSFER WALL EXCHANGE
    LATENT HEAT **THERMAL** **STORAGE** MATERIAL LIQUID
    HEAT CARRY FLASH POINT MORE STORAGE MATERIAL MELT POINT
DC - J08 Q74
IC - C09K-005/06; F24H-007/00; F24H-007/02
MC - J08-D
NP - 1
NC - 001

2/394 DWPI - (C) The Thomson Corp
AN - 99-161228 [14]
```

Figure 1 : format gratuit d'une notice bibliographique de la base de données Derwent WPI sur Questel

Ce format gratuit peut être utilisé pour visualiser toutes les références répondant à la stratégie d'interrogation établie, pour ensuite tenter de vérifier la validité de ce corpus, d'affiner la stratégie et finalement extraire les notices pertinentes pour l'étude de veille technologique. Ce format gratuit de notices de DWPI privilégiant les champs bibliographiques retranscrivant le contenu technique de la demande de brevet (TT, DC, IC, MC)¹, il paraît envisageable d'effectuer cette sélection de notices pertinentes pour l'étude technique souhaitée.

Une telle démarche a été tentée au cours d'une étude de veille technologique dans le domaine technique des appareils de chauffe-eau. Le client de cette étude souhaitait faire l'état des technologies dans les domaines de l'isolation thermique et du stockage/conservation de l'énergie pour ce type d'appareils. L'étude devait tout particulièrement aborder les avancées techniques dans le domaine de l'énergie solaire et pouvoir identifier tous les acteurs internationaux et plus particulièrement les acteurs japonais développant ces technologies. Un traitement bibliométrique était envisagé pour étayer l'analyse technique et favoriser l'identification des acteurs majeurs.

La stratégie de recherche mise au point pour interroger la base de données DWPI a permis d'identifier 394 notices. Selon la contrainte de l'enveloppe budgétaire fixée pour la collecte des données, ce corpus devait être réduit de moitié pour pouvoir télécharger les notices complètes de DWPI. Ne sachant pas comment restreindre ce corpus de notices de moitié sans prendre le risque de passer à côté de brevets clefs du domaine technique, la démarche envisagée a été de télécharger ces 394 notices de la base de données DWPI au format gratuit (Figure 1) pour les soumettre à un expert du domaine pour sélection.

Ce corpus bibliographique en main, l'expert n'a pas été capable de sélectionner les notices qui semblaient être les plus pertinentes pour l'étude technique. Ce format bibliographique gratuit ne proposant pas le titre des demandes de brevet normalisé par Derwent et surtout n'offrant pas le résumé rédigé par Derwent², l'expert n'a pas pu se prononcer sur la pertinence de près de la moitié des notices du corpus. La seule présence des mots significatifs du titre (TT) n'était pas suffisante pour décider de l'intérêt ou non de certaines demandes de brevet. Il faut noter qu'à ce stade de l'étape de sélection/validation, l'expert n'avait pas encore connaissance de la signification des codes CIB, Codes manuels et Codes DWPI.

La problématique était alors de voir comment un traitement bibliométrique de ce corpus de notices au format gratuit pouvait assister l'expert dans l'étape de sélection des notices à étudier ultérieurement. Nous allons voir quelle démarche et quels instruments ont été élaborés pour aider à faire converger le point de vue de l'expert avec celui du professionnel de l'information devant réaliser cette étude de veille technologique. Nous verrons aussi comment cette phase de dialogue soutenue par des supports issus de traitement bibliométriques a permis de créer une nouvelle nomenclature de caractérisation technique des brevets déposés. Cette nouvelle nomenclature technique étant plus adaptée à l'objectif de l'étude technique souhaitée que les nomenclatures disponibles dans les notices de DWPI (CIB, DC, MC voire TT ou TI), l'étude bibliométrique finale des notices complètes n'en a été que plus pertinente.

3 Méthode employée

Le traitement bibliométrique mis en place pour assister l'expert dans sa phase de sélection devait permettre de structurer la lecture du corpus afin d'isoler plus facilement les notices potentiellement pertinentes des notices potentiellement non pertinentes. L'approche retenue pour élaborer cette grille de lecture a été la suivante :

¹ TT = les termes significatifs du titre du brevet normalisé par Derwent

DC = codes des classes DWPI attribués par Derwent

IC = codes de la Classification Internationale des Brevets attribués par les offices de dépôt

MC = codes manuels attribués par Derwent

² Derwent est un producteur de base de données sur les brevets qui ne propose pas les titres et résumés originaux des demandes de brevet (qui sont souvent laconiques voire volontairement vide de sens) mais il réécrit totalement ces titres et résumés à partir d'une analyse de la demande complète. Les titres et résumés Derwent sont donc excessivement plus représentatifs de la demande que leurs versions originales. C'est ce qui fait l'attrait des professionnels de l'IST pour les bases produites par ce producteur et qui explique le coût très élevé de ces produits bibliographiques.

- regrouper les notices qui abordent les aspects techniques communs (regroupement thématique de notices par classification automatique)
- présenter le résultat de ces regroupements thématiques sur un support facilement exploitable par l'expert
- grâce à ce support dialoguer à l'expert afin de converger vers une sélection de regroupements thématiques de notices permettant de réduire par 2 le corpus

Regroupement thématique des notices au format gratuit

Après plusieurs tentatives de classification automatique (clustering), les résultats les plus satisfaisants au point de vue statistique³ ont été obtenus en représentant les caractéristiques techniques des demandes de brevet par les codes CIB tronqués à 7 digits, c'est-à-dire au niveau de la subdivision des groupes principaux de ce classement hiérarchique. A ce niveau de subdivision, les codes CIB restent encore à un niveau de description technique suffisamment fin pour bien caractériser la demande de brevet.

Cette conclusion est convergente avec de précédentes études bibliométriques mettant en jeu des problématiques de caractérisation technique des brevets par les codes disponibles dans les notices DWPI [6], [7].

Le tableau traité par la méthode de classification automatique est un tableau binaire (présence/absence) de la forme suivante (Figure 2) :

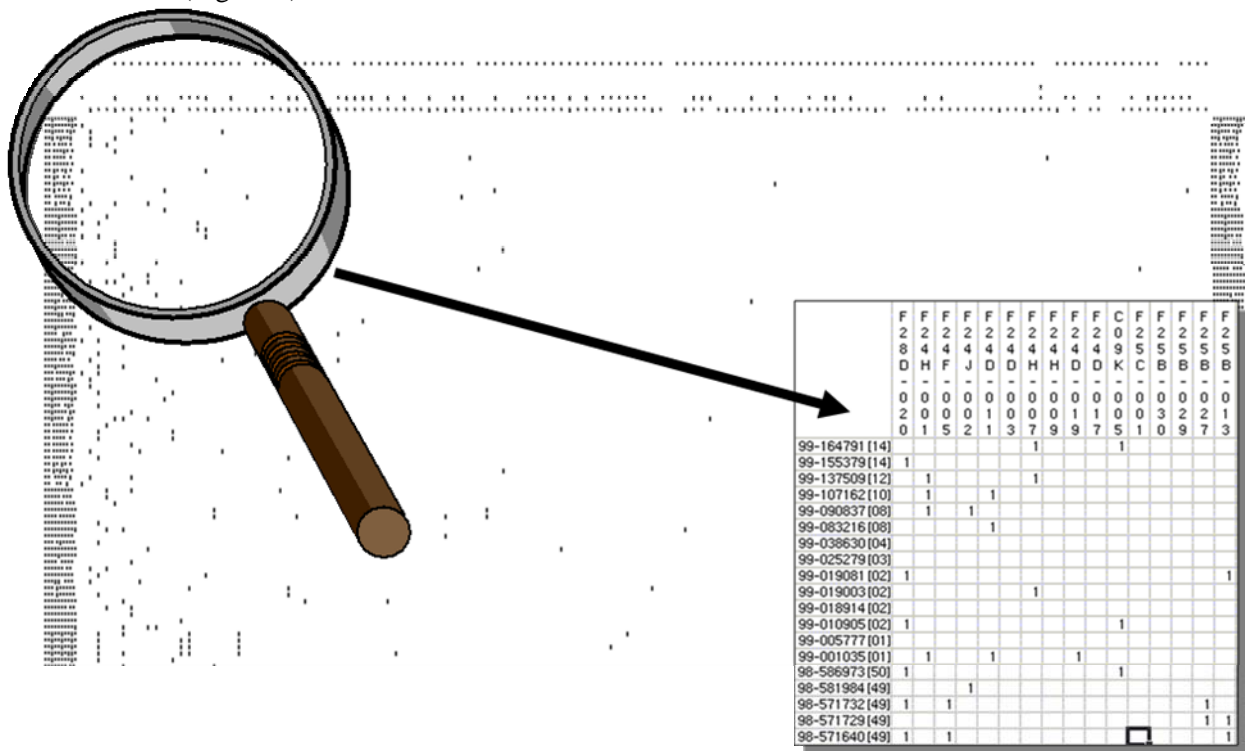


Figure 2 : tableau de données caractérisant les notices selon les aspects techniques exprimés par la subdivision des groupes principaux des codes CIB

Une ligne de ce tableau représente le profil technique d'une demande de brevet du corpus selon la description donnée par l'ensemble des codes CIB (tronqués à la subdivision des groupes principaux) affectée à sa notice bibliographique. Par exemple, la première notice du corpus, présentée par la Figure 1, est représentée dans le tableau ci-dessus par une ligne comportant que des valeurs 0 (non représentées dans la figure) sauf pour les colonnes F24H-007 et C09K-009 qui comportent la valeur 1 puisque ces 2 groupes principaux de la CIB caractérisent cette notice (Cf. Figure 1).

Les regroupements thématiques des notices ont été obtenus en analysant ce tableau binaire par une méthode statistique de classification automatique. Une méthode de classification automatique cherche à regrouper les

³ Nombre de clusters, taille des clusters, homogénéité de chaque cluster et grande dissemblance entre les clusters

lignes (ou les colonnes) d'un tableau qui ont les profils de description les plus ressemblants. Donc dans notre cas, la classification automatique regroupe les demandes de brevet selon les profils techniques les plus ressemblants, c'est-à-dire les combinaisons de présence de codes CIB les plus ressemblantes.

La méthode de classification automatique employée au cours de cette étude est celle des centres mobiles (K-means clustering). Cette méthode de classification automatique non hiérarchique permet d'obtenir un regroupement (partitionnement) des lignes du tableau proposé simplement en fixant le nombre de regroupements souhaités (classes). Lorsque certains regroupements ne semblent pas assez homogènes (les profils ne semblent pas suffisamment ressemblants) il est alors possible de relancer la méthode des centres mobiles uniquement sur ce regroupement en précisant le nombre de sous-regroupements souhaités.

Cela aura nécessité de faire des choix arbitraires sur la cohérence des regroupements vis-à-vis des mots significatifs du titre Derwent correspondant, tout en assurant une adéquation entre le nombre de regroupements à effectuer et le nombre de notices contenus dans ces regroupements.

La méthode des centres mobiles appliquée à ce corpus a permis de regrouper les notices en groupes fortement homogènes. La raison de ces regroupements a pu être déduite très intuitivement d'une simple représentation du tableau initial (Figure 2) réorganisé par la permutation de ses lignes selon l'appartenance à ces groupes. La Figure 3 ci-dessous présente une sous-partie de ce tableau réorganisé :

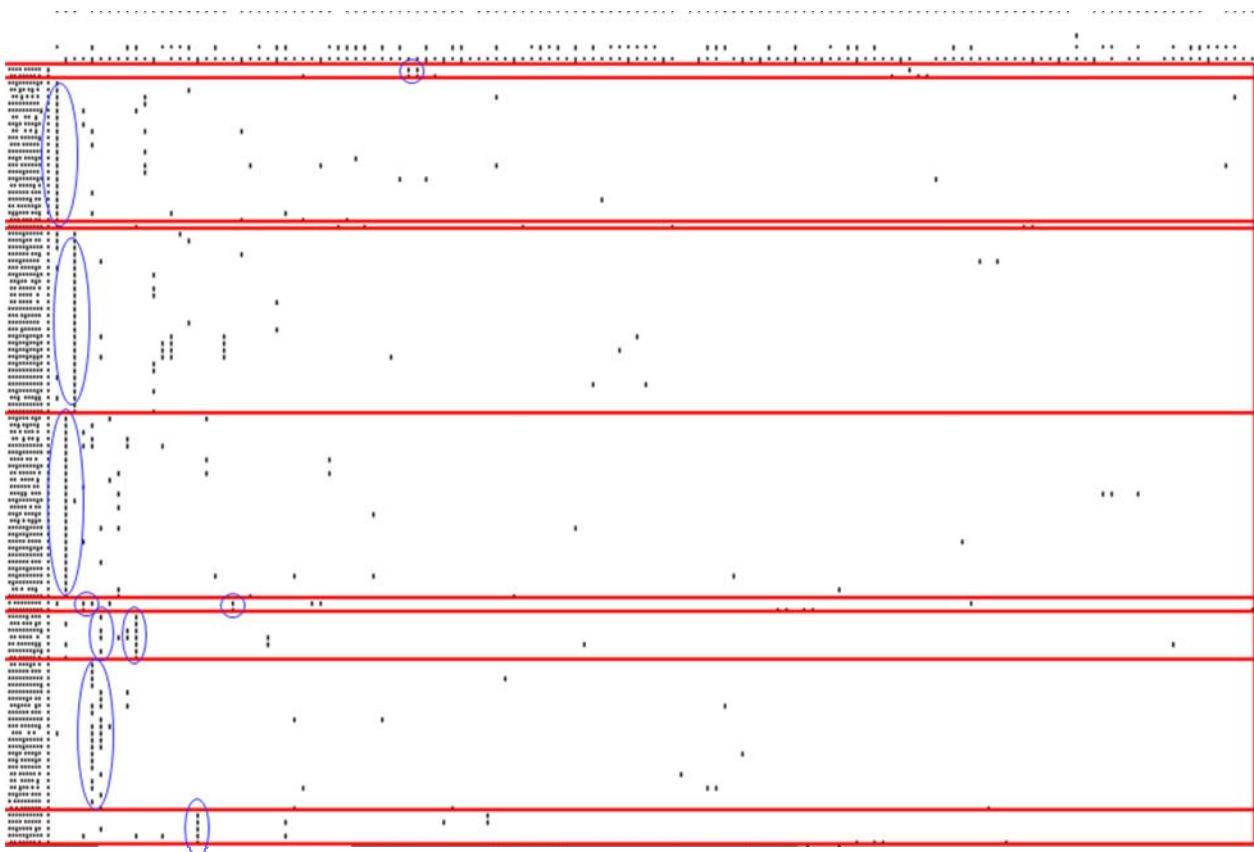


Figure 3 : tableau de données organisé selon le résultat de la classification par la méthode des centres mobiles

Les cadres rouges sur la Figure 3 représentent les différents groupes obtenus par la classification automatique. On peut remarquer que le nombre de notices (lignes) constituant chaque regroupement est très variable. Les combinaisons de codes CIB qui ont contribué à la création de chacun de ces regroupement sont mis en évidence par les ellipses bleues. Ce sont ces combinaisons de codes CIB qui ont discriminé les notices de chaque groupe par rapport à l'ensemble des notices. Ces combinaisons de codes sont communes à toutes les notices d'un même groupe alors qu'elles sont absentes de la plupart des autres notices du corpus. Donc, chaque groupe de notices peut être expliqué par une combinaison de codes très caractéristique. La

définition du groupe principal de chacun de ces codes CIB permet alors de connaître les aspects techniques spécifiques à chaque groupe de demandes de brevet.

Présenter à l'expert le résultat de ces regroupements thématiques sur un support adéquat

L'expert n'étant pas un professionnel de l'information, il est apparu préférable de lui faciliter l'appropriation de l'ensemble de ces résultats statistiques en les intégrant sur un support unique et facilement appropriable. Plutôt que de remonter toutes les données et les résultats statistiques dans un environnement offrant une navigation hypertextuelle, le choix s'est tourné vers l'intégration des principales informations utiles à la sélection sous la forme d'un tableau unique construit sous un tableur.

Au final, le support soumis à l'expert pour l'aider dans cette étape de sélection des notices pertinentes était de la forme suivante (Figure 4) :

N° brevet	N° classe		CIB	Titre du brevet
93-018198 [02]	1	N	F02C-007	POWER PLANT INCORPORATE COMPRESS AIR STORAGE SATURATE COMBUST DRIVE TURBOGENERATOR PEAK LOAD EN
96-440583 [44]	1	N	F02C-006	LIQUEFY FUEL GAS EVAPORATION LATENT **HEAT** **STORAGE** METHOD GAS TURBINE POWER PLANT **THERMAL** **
89-309620 [42]	2	O		STORAGE HEATER PHASE CHANGE MEDIUM **HOT** **WATER** GAS CIRCULATE PASSAGE ALTERNATE RIGHT ANGLE **H
91-149908 [21]	2	?		COLD **HEAT** **STORAGE** SYSTEM SOLID STORAGE MATERIAL CONVECTION PREVENT LAYER STORAGE CHAMBER
92-089556 [12]	2	?		LONG TERM **HOT** **WATER** STORAGE UNIT POLYETHYLENE@ BODY FLANGE CONNECT HEAT EXCHANGE SIMILAR EN
92-278510 [34]	2	N		HEAT TRANSPORT APPARATUS TRAIN TRUCK SHIP HEAT NUCLEAR FOSSIL FUEL POWER ELECTRIC GENERATE STATION S
93-127312 [16]	2	N		ENERGY STORAGE SYSTEM EXCESS ELECTRIC ENERGY ELECTROLYTIC DISSOCIATE WATER OBTAIN GAS STORAGE POW
93-296404 [38]	2	N		THERMAL ENERGY STORAGE BUFFER SPACE HEAT INSTALLATION TRAVERSE PERFORATION RETURN PIPE INLET OUTLET
94-043741 [06]	2	O		NON CENTRE **HOT** **WATER** TANK LATENT **HEAT** **STORAGE** MATERIAL AGITATE MATERIAL CRYSTAL
94-137239 [17]	2	?		**HEAT** **STORAGE** SYSTEM ENGINE COLD START ADD **HOT** **WATER** STORAGE TANK FLOW FIRST PLACE SECO
94-241578 [30]	2	N	F28D-020	QUASI TERNARY SALT MIXTURE LATENT **HEAT** **STORAGE** CONTAIN PERCHLORATE HYDRATE SUIT **STORAGE** **
94-307949 [38]	2	?	F24D-011	**HEAT** **STORAGE** MATERIAL HEATER ARCTIC CLOTHING COMPRISE LATENT **HEAT** **STORAGE** AGENT RESIN SH
95-140259 [19]	2	?	C09K-005	THERMAL ENERGY STORAGE ALTERNATE SORPTION DESORB UTILISE THERMOCHEMICAL MOLECULAR SIEVE ENERGY DE
96-335466 [34]	2	?		ERY THRITOL **HEAT** **STORAGE** COMPOSITION USEFUL ENERGY SAVE HEAT APPLY STABILISED SILICIC ACID COMPO
96-504763 [50]	2	N		DEVICE **STORAGE** **THERMAL** ENERGY VESSEL FILLED HEAT LIQUID COMPRISE INTERCONNECT TUBE FLOW LIQUID C
96-508827 [51]	2	N		HEAT ENERGY STORAGE METHOD AIR CONDITION APPARATUS HEAT RECOVER PUMP SOLAR HEATER STORAGE SALT C
97-000709 [01]	2	?		**HEAT** **STORAGE** DEVICE RAPID WARM UP VEHICLE HEAT SYSTEM SOLID LIQUID PHASE CHANGE **HEAT** **STORA
97-120253 [12]	2	N		WASTE HEAT ELIMINATE SALT **WATER** **HEAT** **STORAGE** LIGHT COATING APPLY SURFACE LIQUEFY PHASE TRAN
98-000860 [01]	2	?		**WATER** **HEAT** **STORAGE** BARRIER LAYER FLOAT WATER SURFACE PREVENT ENTER OXYGEN **HEAT** **WATE
98-468294 [41]	2	N		**WATER** **HEAT** SYSTEM SOLAR ENERGY THREE HEAT EXCHANGE STORAGE TANK CONTROL BY-PASS OPERATE CO
98-586973 [50]	2	?		**THERMAL** **STORAGE** HEAT DISSIPATE METHOD ELECTRIC **WATER** **HEAT** ALCOHOL OBTAIN SUGAR SOURC
99-010905 [02]	2	O		LATENT **HEAT** **STORAGE** BODY FLEXIBLE CONFIGURATION FIBRE PANEL FORMING ADHESIVE COMPRESS CAPILLAF
99-019081 [02]	2	?		**THERMAL** **STORAGE** MECHANISM **THERMAL** **STORAGE** TANK REFRIGERATE BOILER HEAT PUMP CONNECT CO
99-155379 [14]	2	?		PROCESS WATER **STORAGE** **HEAT** EXCHANGE LOWER REGION GUIDE SYSTEM **HEAT** **WATER** UPPER REGION

Figure 4 : support soumis à l'expert pour l'assister dans l'étape de sélection des notices pertinentes

Ce tableau présente l'ensemble des notices du corpus initial classé selon le résultat de la classification automatique. Chaque ligne du tableau représente une notice du corpus. La seconde colonne intitulée N° classe indique le groupe auquel cette notice a été affectée par la classification automatique. L'appartenance des notices aux différents groupes est mise en exergue par un encadrement en rouge. La combinaison des codes CIB qui a discriminé les notices de chaque groupe est mentionnée dans la quatrième colonne. La dernière colonne reprend les termes significatifs du titre du brevet contenu dans le champ TT des notices. Et pour finir, la troisième colonne (sans intitulé) contient l'avis de l'expert qu'il avait précédemment donné sur la pertinence ou non de la notice pour l'étude. Les notices restées sans avis de l'expert sont repérables par un point d'interrogation dans cette colonne.

Un tel support a permis assez facilement à l'expert et au professionnel de l'information de converger vers une sélection de groupes de notices à télécharger pour l'étude technique finale. Les groupes de notices contenant principalement des avis négatifs ont été rapidement réévalués et souvent définitivement rejetés. Les groupes de notices contenant majoritairement des avis positifs ont souvent été confirmés comme pertinents. Les groupes de notices comportant majoritairement des points d'interrogation ou autant d'avis positifs que négatifs étaient alors analysés plus finement par l'expert. Le professionnel de l'information pouvait à tout instant assister l'analyse de l'expert, qui se basait essentiellement sur l'interprétation de la colonne des termes significatifs du titre des brevets, en complétant la caractérisation technique de ce groupe par la signification des codes CIB qui le caractérisaient. Pour cela, la définition des codes CIB avait été remontée sous forme hypertextuelle⁴ et cela jusqu'au niveau hiérarchique des groupes principaux (Figure 5).

⁴ au format HLP propres aux fichiers d'aide de Windows avant sa mise à disposition en fichiers HTML

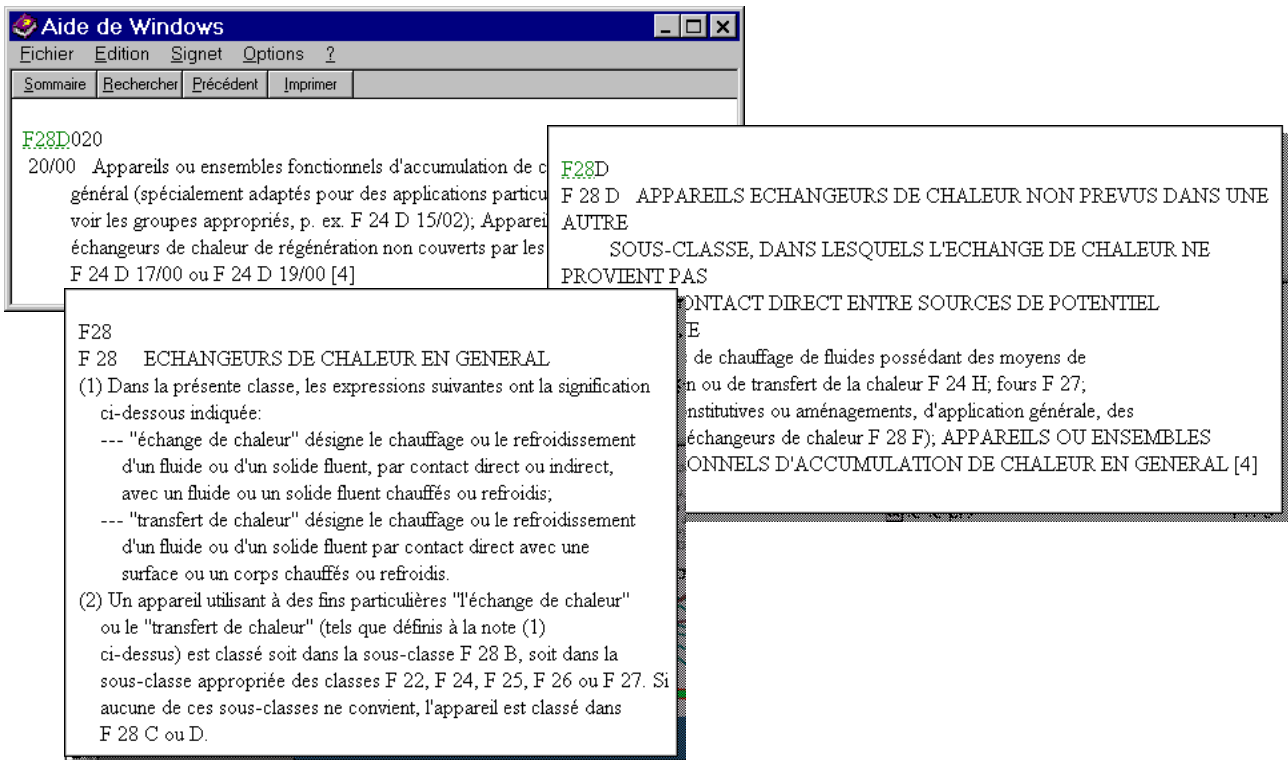


Figure 5 : navigation hypertextuelle dans la signification des différents niveaux hiérarchiques des codes CIB

Cette démarche a permis de rapidement sélectionner les demandes de brevet pertinentes pour l'étude car l'avis à donner n'était plus à réfléchir notice par notice mais par groupes de plusieurs notices. De plus, cet avis était conforté par la cohérence de ces regroupements de notices et par la confirmation de la validité des aspects techniques couverts grâce à la définition des codes CIB.

Cette étape de sélection/validation a abouti à l'identification d'un ensemble de 182 demandes de brevet qui semblaient les plus pertinentes pour l'étude.

Création d'une nouvelle nomenclature technique pour caractériser les demandes de brevet

Le travail commun de l'expert et du professionnel de l'information pour valider la pertinence des notices a aussi permis de très facilement nommer les groupes de notices sélectionnés selon les caractéristiques techniques abordées par les demandes de brevet. Très naturellement, les sous-domaines techniques du domaine étudié ont émergé de l'interprétation des caractéristiques techniques de ces groupes de notices. Et en définitive, chaque groupe de notices sélectionné a été nommé selon l'un des sous-domaines techniques suivants (Tableau 1) :

Tableau 1 : sous-domaines techniques identifiés par l'analyse et l'interprétation des groupes thématiques

Accumulation de chaleur
Changement d'état
Chauffage de l'eau
Gestion énergétique d'eau chaude
Isolation de tuyauterie
Isolation d'emballage
Isolation des chaudières
Isolation pour l'eau chaude
Mesure de ballon d'eau chaude
Régulation de ballon d'eau chaude
Stockage de froid
Stockage d'eau chaude
Système de stockage thermique

Cette nouvelle nomenclature technique a ainsi permis de mieux caractériser les demandes de brevet selon la vision souhaitée pour l'étude technique à effectuer par la suite. Il a donc été décidé d'employer cette nouvelle nomenclature pour caractériser toutes les notices téléchargées au format intégral identifiées par cette étape de sélection/validation. L'introduction de cette nomenclature dans chaque notice intégrale a pu être automatisée puisque l'affectation de ces 182 notices à chaque groupe thématique était déjà connue.

4 Résultats de l'étude technique

Le corpus des 182 notices au format intégral complété du sous-domaine technique de la nouvelle nomenclature a alors pu faire l'objet d'une étude bibliométrique. Tous les champs bibliographiques des notices DWPI ont pu être analysés (exemple Figure 6) et combinés entre eux (exemple Figure 7) mais aussi avec la nouvelle nomenclature technique offrant ainsi une grille d'analyse de ces brevets sélectionnés particulièrement adaptée à l'étude technique souhaitée (Figure 8, Figure 9 et Figure 10).

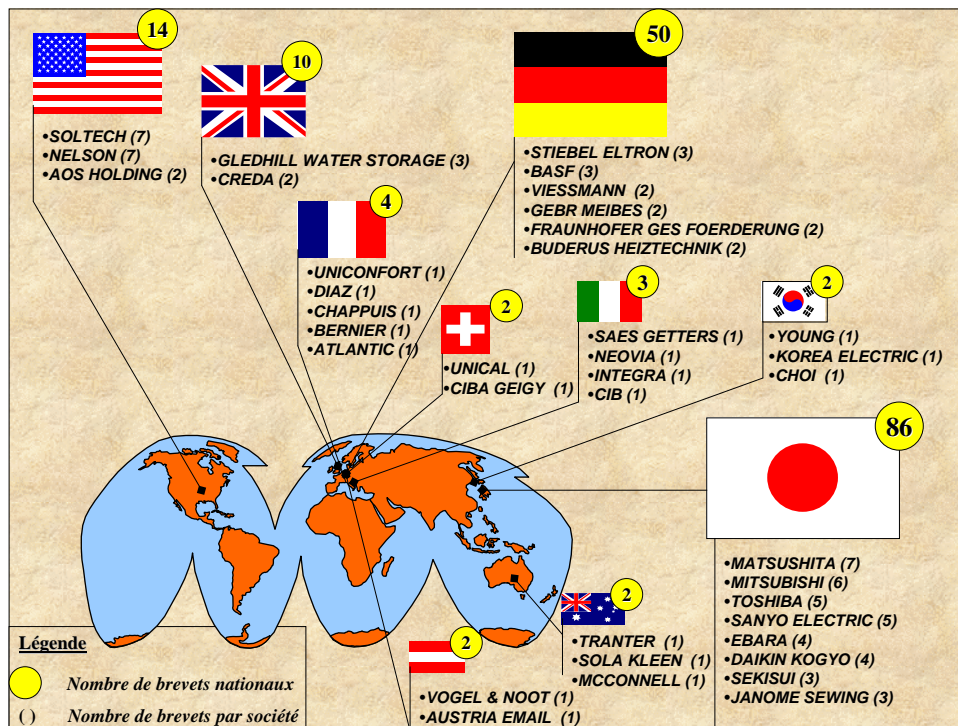


Figure 6 : répartition mondiale des sociétés déposants des brevets dans tous les domaines techniques étudiés

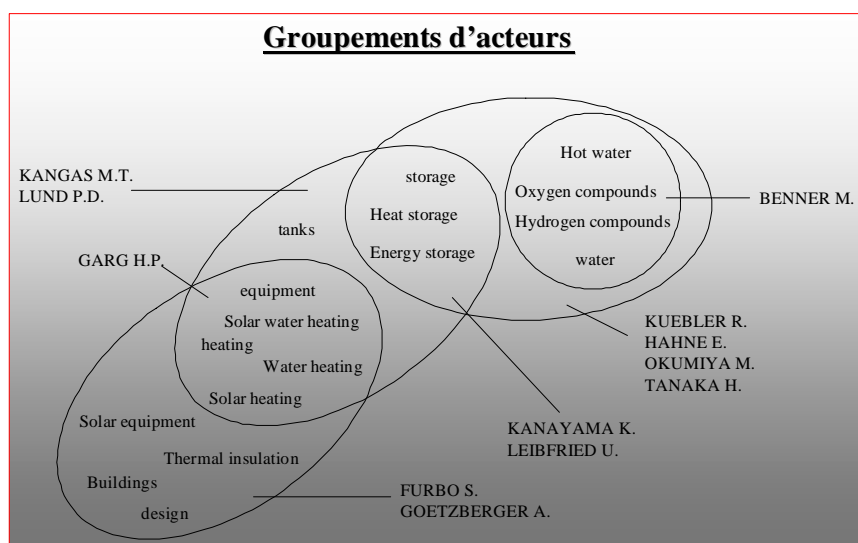


Figure 7 : Domaines de compétences (caractérisés par les termes du champ TT) de certains inventeurs

La Figure 8 présente, par exemple, la part de ces sous-domaines techniques protégés par les demandes de brevet de ce corpus. Certains domaines techniques de la nouvelle nomenclature ont été regroupés sous le nom de domaines techniques plus génériques pour encore améliorer la grille d'interprétation du domaine étudié.

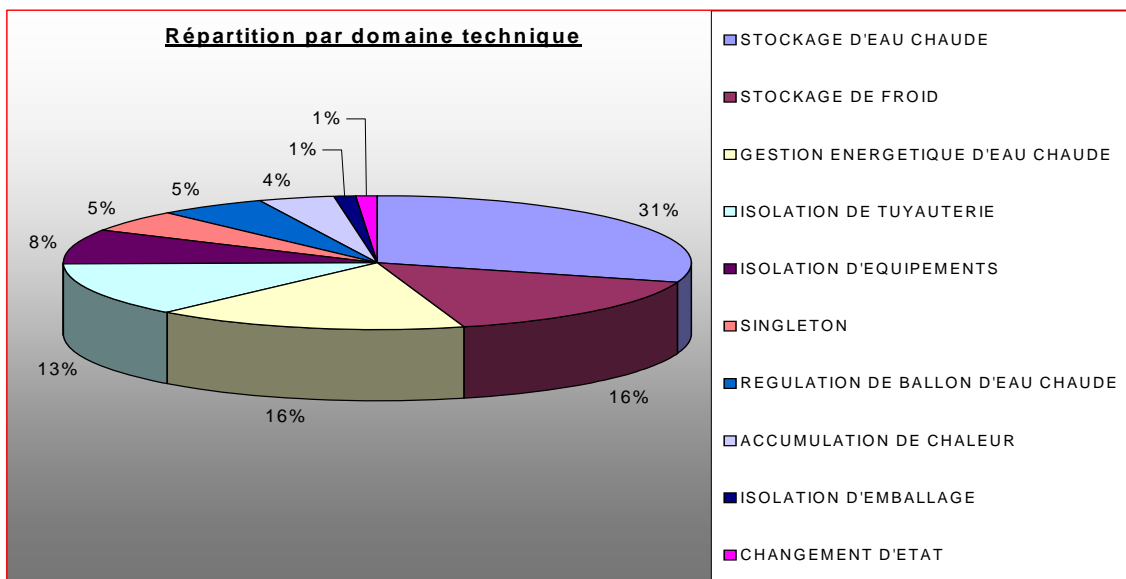


Figure 8 : répartition de la part des demandes de brevet selon les sous-domaines techniques identifiés

La Figure 9, quant à elle, présente l'évolution annuelle de la répartition des sous-domaines techniques protégés par les demandes de brevet concernant le domaine du stockage/conservation de l'énergie.

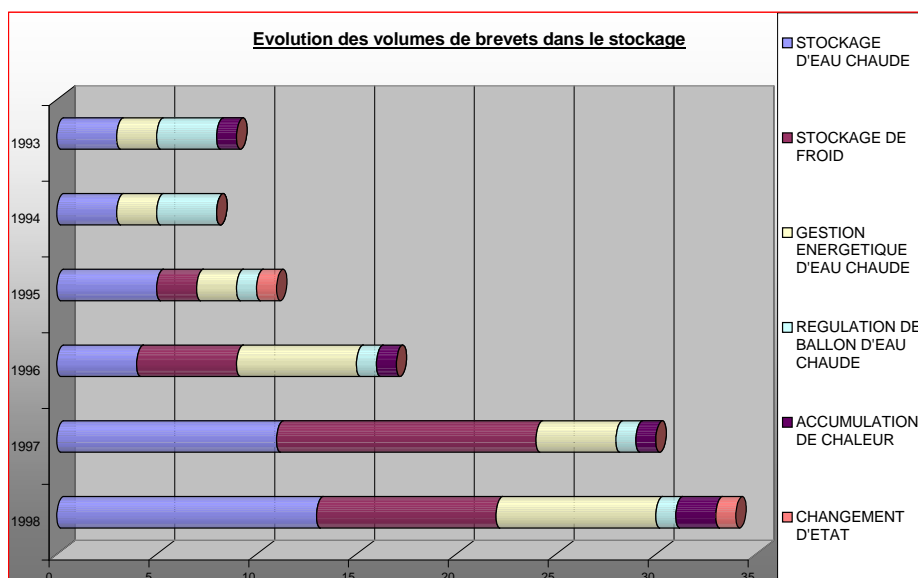


Figure 9 : évolution du nombre de dépôt de brevet dans les sous-domaines techniques du stockage/conservation de l'énergie au cours d'une période de 6 ans

La Figure 10 présente le profil thématique du portefeuille brevets des principales sociétés qui déposent des brevets dans les domaines techniques du stockage/conservation de l'énergie.

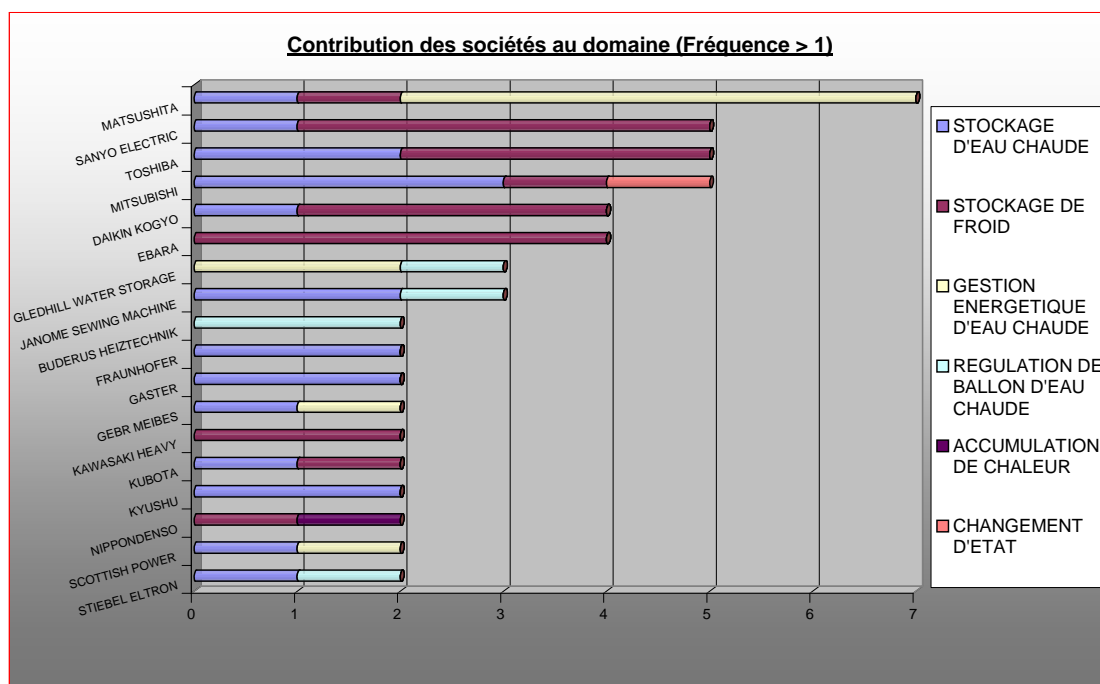


Figure 10 : profil des compétences techniques détenues par les principaux acteurs déposants des brevets dans des domaines techniques du stockage/conservation de l'énergie

5 Conclusion

Au cours de cette étude de veille technologique, les traitements bibliométriques ont démontré leurs utilités à la fois dans la phase de sélection et validation des demandes de brevet à étudier et par la suite dans la phase d'analyse du corpus de l'ensemble des demandes de brevet sélectionnés.

Elle montre également la mise en œuvre d'une combinaison de méthodes et d'outils associés à la coordination de plusieurs ressources humaines complémentaires permettant ainsi une maîtrise des coûts lors d'une telle étude de veille technologique. La conduite de cette étude illustre l'importance de la mise en commun de multiples compétences lors d'une telle étude, à savoir un professionnel de l'information, un spécialiste du traitement automatique de l'information et un expert technique du domaine.

6 Bibliographie

- [1] DEVALAN P, CANDORET JP, BOUVET C, LION JC, « La bibliométrie. Un outil de veille technologique pour l'entreprise », CETIM-Informations, N°116, p.89-95, 1990
- [2] DOU H, « Maîtriser la prolifération de l'information grâce à la bibliométrie », Technologies internationales, n°9, p. 42-44, 1994
- [3] DOU H, QUONIAM L, ROSTAING H, NIVOL W, « L'analyse des données au service de la bibliométrie. Outil de veille technologique à la dimension des moyennes entreprises », Revue Française de Bibliométrie, Paris, n°8, p. 27-67, 1990
- [4] ESCORSA CASTELLS P, RODRIGUEZ SLAVADOR M, MASPONS BOSCH R, « Technology mapping, business strategy and market opportunities », Competitive intelligence review, Vol.11, N°1, p.46-57, 2000
- [5] KOSTOFF R N, *Text Mining for Global Technology Watch*, in: DRAKE M A, *Encyclopedia of Library and Information Science*, Ed: Marcel Dekker Inc, NY, 2003
- [6] QUONIAM L, HASSANALY P, BALDIT P, ROSTAING H, DOU H, « Bibliometric analysis of patent documents for R&D management », Research Evaluation, England, Vol 3, N°1, p. 13-18, Avril 1993

- [7] ROSTAING H, NIVOL W, QUONIAM L, BÉDÉCARRAX C, HUOT C, « L'exploitation systématique des bases de données: des analyses stratégiques pour l'entreprise », Les Cahiers de l'ADEST, Paris, p. 7-22, Juillet 1993
- [8] TRIPPE A J, « Patinformatics :Tasks to tools », World Patent Information, n°25, p 211-221, 2003