

Méthodologie pour l'identification de tendances technologiques en utilisant des logiciels de la famille ISIS ayant comme source d'information le document de brevet.

VERA L. M. LELLIS

vlellis@int.gov.br

Instituto Nacional de Tecnologia, Av. Venezuela, 82, Rio de Janeiro, Brésil

www.int.gov.br

Mots-clés:

[*Famille ISIS*](#), bibliométrie, analyse automatique de l'information, traitement automatique de l'information, brevets.

Keywords:

ISIS Family, bibliometry, automatic analyze information, automatic treatment of information, , patents documents.

Palabras claves:

Familia *ISIS*, bibliometria, análisis automática de información, tratamiento automático de información, *WINISIS*, patentes.

Résumé

Le cycle technologique est de plus en plus court, c'est à dire, la technologie arrive à la désuétude avant de se détériorer et est substituée dans des délais de plus en plus courts. L'un de ces subsides pour cette évolution est sans aucun doute l'information technologique, et la demande pour ce produit est proportionnelle à la croissance de la science, de la technologie et favorise sans doute l'innovation. La croissance de l'intérêt pour de nouvelles technologies se reflète dans l'augmentation de R&D qui de son côté se reflète dans l'augmentation de dépôts de brevets. Parce qu'elles constituent un moyen de divulgation technologique de la plus récente publication, les bases de données des brevets ont été utilisées comme source de veille technologique, en présupposant qu'il est possible d'identifier des chemins innovateurs pour de nouvelles technologies, des méthodes et des processus, par l'analyse des modèles de demandes de brevets dans des domaines déterminés de la connaissance. Les résultats sont souvent présentés de façon quantitative, mais son utilisation dans le procès de décision permet des évaluations qualitatives. Le présent travail a comme objectif de montrer l'expérience d'un organisme d'action technologique brésilien dans l'utilisation de méthodologies de veille visant à identifier des thèmes prioritaires pour un programme de développement technologique. La première étape du programme a comme point de départ la veille d'information dans la base de données des brevets et l'emphase est donnée à l'utilisation de logiciels de la famille ISIS pour le traitement automatique de l'information dans le domaine des brevets. Le résultat préliminaire nous a permis d'obtenir des réponses aux questions technologiques qui peuvent définir les chemins d'innovation d'un programme de gouvernement.

Abstracts

The technological cycle is increasingly short, technology arrives at disuse before deteriorate and is substituted within increasingly short times. The subsidy for evolution is without any doubt technological information, and demand for this product is proportional to the growth of science, technology and undoubtedly supports for innovation. The growth of interest for new technologies is reflected in augmentation of R&D which on its side is reflected in augmentation of patent register. For constitute a technological means of release of the most recent publication, the data bases of the patents were used like technological source of forecast, by presupposing it is a possible identifier of the innovating ways for new technologies, methods and processes, by analyze of the models of patent applications in given fields of knowledge. The results are often presented in a quantitative way, but its use in the decision process allows qualitative evaluations. This work has like objective to show an experience in a Brazilian technological institute in use of methodologies of forecasting for aim identifying priority topics for a program of technological development. The first stage of the program has as starting point the forecasting information in the data base of the patents and emphasis is given to the use of ISIS family software's, to do automatic treatment of information in the field of the patents. The preliminary result enabled us to take answers to the technological questions which can define ways innovation for government program.

1 Introduction

Créer des programmes de développement technologique demande de la créativité, caractéristique qui doit être présente dans toute les études de cette nature, des groupes avec des compétences transdisciplinaires et une connaissance de méthodes prospectives propres à l'identification de tendances technologiques à partir de macros thèmes d'intérêt mondial.

Un regard holistique sur cet ensemble nous permet de connaître les alignements technologiques qui se forment à partir du développement de la science et de la technologie. Du point de vue technologique ces thèmes doivent répondre aux demandes du secteur industriel d'un pays et doivent concerner les Instituts de Recherche Technologique- IPT's, parce que ceux-ci ont pour mission de créer le lien entre l'université et le secteur industriel. En d'autres termes, on peut dire que les IPT's doivent transformer la science en technologie et accomplir son transfert pour le secteur productif.

Dans ce sens, il est important que les institutions publiques insérées dans le contexte de la science et de la technologie aient des objectifs liés à gestion stratégique nationale d'où sont créés les objectifs du plan directeur institutionnel.

L'objectif de cette étude est présenter la méthodologie de veille utilisée dans la construction d'un programme de développement technologique, d'un Institut de Recherche Technologique brésilien, que nous appellerons à partir de maintenant l'INT.

La méthodologie présentée est une composition de techniques et d'outils de veilles connus internationalement mais l'emphase est donné à l'utilisation de logiciels de la famille ISIS, dans l'étape du traitement automatique de l'information contenue dans le document de brevets.

Les résultats partiels de l'étude nous ont permis d'obtenir des réponses aux questions technologiques décisives pour la prise de décision ayant trait aux chemins d'innovation qui doivent être suivis dans un programme de développement technologique d'une institution gouvernementale, comme ils nous permettent de connaître le panorama international des thèmes choisis pour composer le programme de développement technologique. Ce cadre a été possible grâce aux recours offerts par la famille ISIS, qui a passé à intégrer le cadre des outils de traitement de l'information.

2 L'Importance de Prévoir les Tendances

D'après Berger [1], avant d'être une méthode ou une discipline, la veille est une attitude. La principale caractéristique de l'attitude prospectrice est de voir loin, de connaître ce qui va arriver, même si c'est une hypothèse.

Prévoir des tendances entraîne la systématisation des futurs capables d'influencer de façon significative une industrie, une économie ou une société comme un tout.

Créer des scénarios qui vont survenir est une activité qui est en train d'être développée non seulement par les gouvernements, mais principalement par n'importe quel secteur de la société de n'importe quel pays, comme moyen de préparation des acteurs en question pour affronter les opportunités ou les menaces et garantir sa compétitivité sur le marché, dans le cas du secteur productif, ou faire face aux adversités internationales, en garantissant un développement national de pointe, dans le cas des institutions publiques. L'identification des tendances ou des routes technologiques peut être obtenue à partir d'une série de méthodologies combinées avec des outils de TI dont les résultats dépendent des enquêtes faites dans les divers niveaux d'intérêt. Pour cela, il est commun de trouver des méthodes et des techniques développées pour des usages spécifiques, ce qui, parfois nous amènent à des résultats contestables dus à la difficulté inhérente au traitement des incertitudes du futur.

Une réflexion sur les différents abordages, les avantages et les désavantages, nous a amené à combiner les méthodes de *brainstorming*, l'analyse de brevets et de méthodes statistiques en bibliométrie.

3 Abordage Methodologique

Cette méthodologie a été développée en six étapes et a eu comme objectif la consolidation d'un document pour orienter les processus de recherche dans des bases de données structurées, et le traitement et l'analyse des résultats.

3.1 Planification de la Strategie de Recherche

Les thèmes ont été définis par les spécialistes des secteurs, car ils connaissent leurs caractéristiques et leurs particularités.

Ensuite, on recherche l'équivalence des thèmes dans le Classement International de Brevets – IPC, en visant à incorporer le sujet dans les catégories du système de classement des brevets, avec l'objectif d'obtenir une récupération totale à partir de termes consacrés et utilisés par la communauté internationale.

Une autre forme de validité des termes pour la stratégie de recherche est l'utilisation de logiciels que rendent possible la création de *rank*, favorisant la comparaison de listes de descriptifs récupérés sur internet au moyen d'un moteur de recherche. Les logiciels suivants peuvent être utilisés. Quoniam [2]:

- lustering engine;
- search vivismo;
- iboogie;
- gigablast;
- CM-E meta engine.

Les mots-clés doivent être testés dans les outils ci-dessus et le résultat de cette étape est la présentation, sous forme de liste, des descriptifs qui sont les moins ou les plus utilisés sur internet. Cette information contribue aussi à la validité des termes qui seront utilisés dans la stratégie de recherche.

Il est également possible de soumettre les termes à la recherche dans les banques de données *Dialog et STN*[3]. Celles-ci, offrent gratuitement un *rank* de sujets par bases de données. Ainsi est-il possible de visualiser quelles sont les bases de données qui détiennent le plus grand volume de sujets consultés.

3.2 Collecte des Données

Avant de décider quelle source d'information nous utiliserons pour la réalisation des recherches il faut identifier quelles sont les bases qui ont le profil qui répond aux paramètres des recherches. Le profil d'une base de données se compose des informations suivantes:

- Couverture thématique de base;
- Quantité totale de documents indexés;
- Types de documents incorporés;
- Producteur de la base;
- Couverture temporaire des données.

Ces informations vont rendre possible la prise de décision quant aux sources d'information plus appropriées aux sujets que nous désirons rechercher.

Il est également important d'identifier quels sont les recours que la base de données offre à l'utilisateur. Dans le cas de la recherche d'informations pour la veille, il est fondamental que la base de données permette de faire des téléchargements de tout l'ensemble d'informations récupérées en une seule fois au lieu de présenter un registre à chaque fois.

3.3 Padronization d'Informations Pertinentes

La recherche d'information au moyen de nouveaux outils de TI, comme l'utilisation de vieux outils en ensemble avec de nouvelles méthodologies peuvent offrir des produits d'information d'une haute valeur ajoutée.

L'utilisation d'outils automatisés pour la récupération et l'analyse de l'information est une solution proposée pour le problème de l'hyper information. Certains outils de traitement de données utilisent des principes bibliométriques et permettent le traitement statistique de n'importe quel ensemble d'informations récupérées dans des bases de données bibliographiques, quel que soit le volume qui sera traité, au moyen de techniques proches de celles utilisées dans le traitement de recherches. Ils permettent aussi, d'obtenir une présentation des références bibliographiques sous forme de tableau pour une analyse postérieure des données.

Dans cette étape, les enregistrements récupérés sont exportés vers des archives créées avec l'extension *txt* en utilisant le mécanisme d'exportation des propres logiciels ou des bases de données électroniques. Ensuite, l'archive avec l'extension *txt* est envoyée pour un logiciel qui rend possible la réalisation d'une uniformisation des informations ou des ensembles des champs.

Un point important qui doit être pris en considération dans toutes les étapes, c'est la dénomination de l'extension des archives dans toutes les étapes du processus d'exportation d'un logiciel vers un autre. A chaque entrée et sortie de l'archive, en accord avec la sollicitation du logiciel, il faut attribuer un nouveau nom à ces archives, afin qu'elles puissent être lues et travaillées correctement par chaque outil.

En l'absence d'un logiciel qui permet de faire l'uniformisation des données, le *Word* réalise quelques modifications dans les archives selon la stratégie de purification ou d'uniformisation des données désirées. Au moyen du dispositif « *localiser/substituer* » il est possible, par exemple, de retirer des espaces en blanc, de mettre les accents sur les mots, de substituer ou supprimer des mots ou des caractères, de réunir ou exclure des champs un utilisant les caractères joker.

En utilisant ce premier filtre, on note que cela augmente considérablement la qualité des informations dans les archives en *txt*. La réunion d'enregistrements de différentes bases de données engendre un volume d'informations avec des modèles différents et des formats qui doivent être analysés et reformatés.

L'absence d'une politique d'information dans les bases de données, l'alimentation de données décentralisées et discontinues, la réunion de champs de différentes bases de données, entre autres choses, font que l'on doit faire un nettoyage dans les champs et dans les informations qui seront traités automatiquement..

L'exportation de données doit être faite dans le format *tag*, c'est à dire, chaque champ reste sur une ligne et les informations arrivent insérées entre un caractère de début et de fin de champ.

3.4 Formatage des archives avec le Systèmes ISIS

Le formatage avec le logiciel *ISIS* est fait en couvrant des étapes suivantes:

- Téléchargements des registres des bases de brevets dans le format *txt*;
- Réunion des registres dans une seule archive;
- Normalisation des données dans le logiciel *Word*, comme cité ci-dessus;
- Formatage de l'archive à travers *ISISAscii* pour que s'obtiennent le format ISO2790;
- Importation pour la base de données *WINISIS*.

Il est important de souligner que la base de données créée dans le *WinISIS* doit contenir les mêmes tags des registres qui seront importés.

3.4.1. WinISIS

Le system *WinISIS* [4,5,6,7,12,13] est la version du *CDS/ISIS* avec une interface graphique pour le système opérationnel *Windows*. Cette version est totalement compatible avec le *CDS/ISIS* pour *DOS*, signifiant que toutes les applications des deux versions peuvent travailler simultanément avec les mêmes bases de données

Le *WinISIS* permet une relation avec l'utilisateur de façon plus directe et intuitive. On cherche à travailler avec les principes d'une interface simple qui facilite aussi bien la convivialité avec la version *DOS* que la transition pour un milieu graphique.

Au-delà de tous les progrès qui rendent possible une interface graphique en *Windows*, le *WinISIS* permet la manipulation simultanée de plusieurs bases de données, de nouvelles ressources de langage de format comme le contrôle du type de source utilisée pour la formatation de caractères, les ressources d'hypertexte, les ressources de bases de données etc....

Des bases de données en *ISIS* reçoivent des archives de n'importe quelle extension, il suffit qu'ils soient convertis pour l'extension *txt* et au moyen de l'applicatif *ISISAscii* [8], transformer les ensembles de champs en archives avec l'extension *ISO/2709* [6] ainsi que le montre la figure ci-dessous].

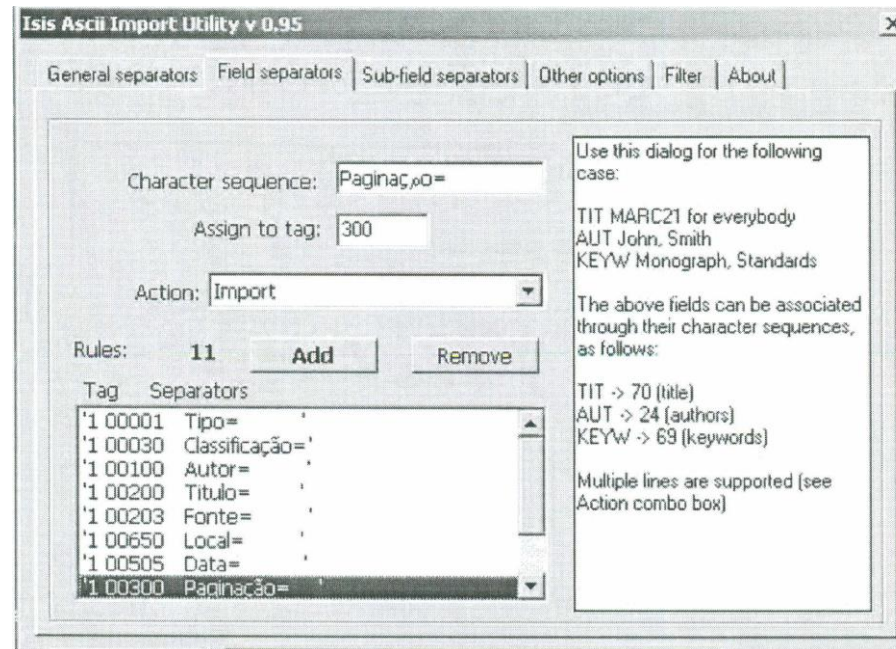


Figure 1: Écran de configuration des ISISACH

Au-delà de l'objectif de stocker les ensembles de champs pour la récupération, l'utilisation du *WinISIS* comme un outil d'analyse est possible grâce à sa caractéristique de gérant de base de données et sa facilité de créer des bases pour l'utilisateur.

Le *WinISIS* rend possible la création de bases de données avec n'importe quel nombre de types de champs et de sous champs, selon les intérêts de la recherche. La base créée dans *WinISIS* dispose de nombreuses formes de récupération, facilitée grâce à ses archives inversées.. Le système permet la récupération au moyen de tout le champ, par sous champ ou champ, par sujet ou par préfixe de sous champ, par préfixe de sujet ou du descripteur, ou encore au moyen des opérateurs de recherche *booleanos*, pour combiner deux ou plus de sujets d'un seul ou plusieurs champs.

Le *WinISIS* rend aussi possible le décompte de la fréquence et l'occurrence de tous les mots du titre, du résumé ou des champs importés préalablement, comme montre la figure 2 et 3 ci-dessous.

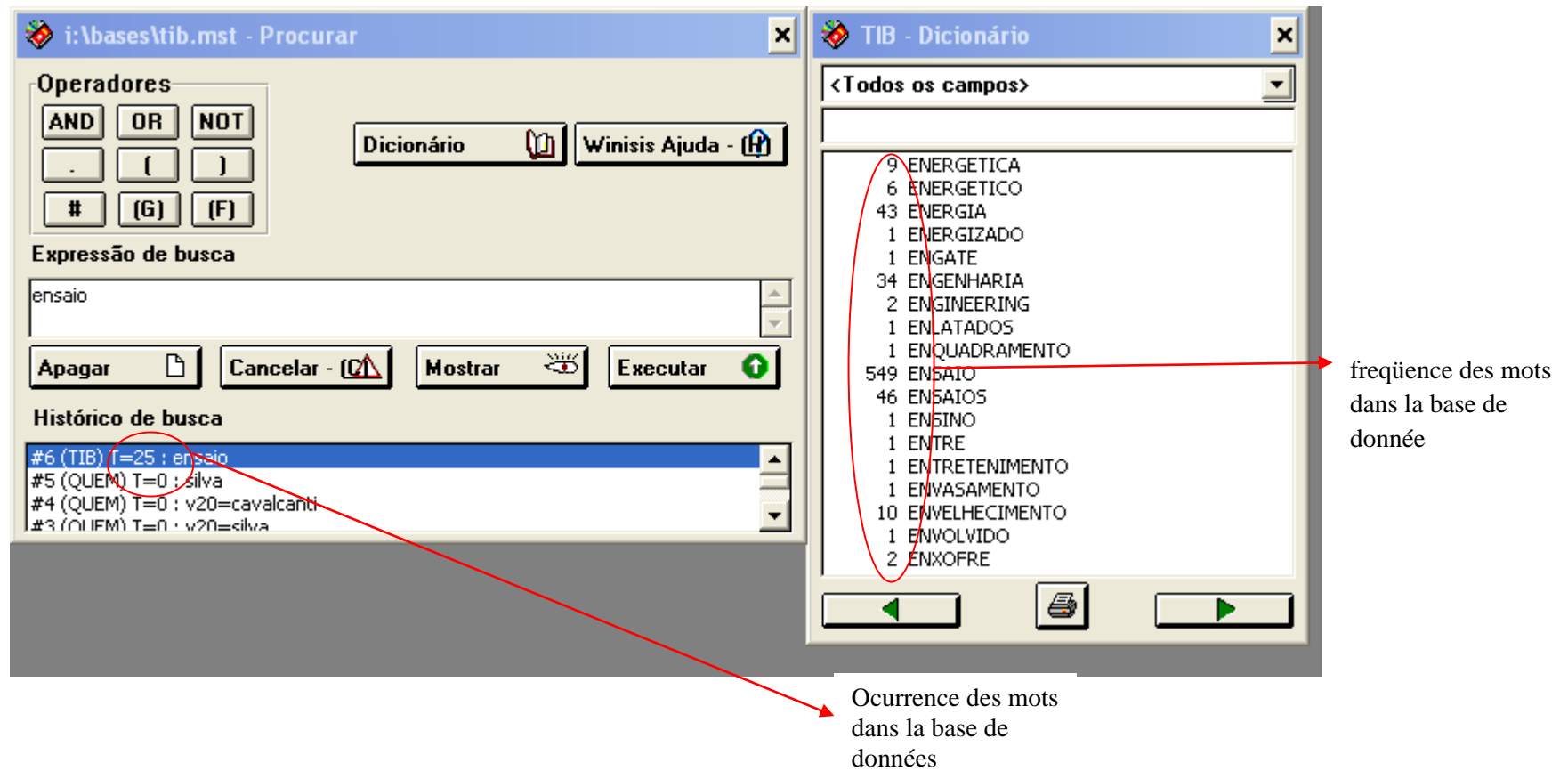


Figure2: Écran avec le dictionnaire du logiciel Winisis

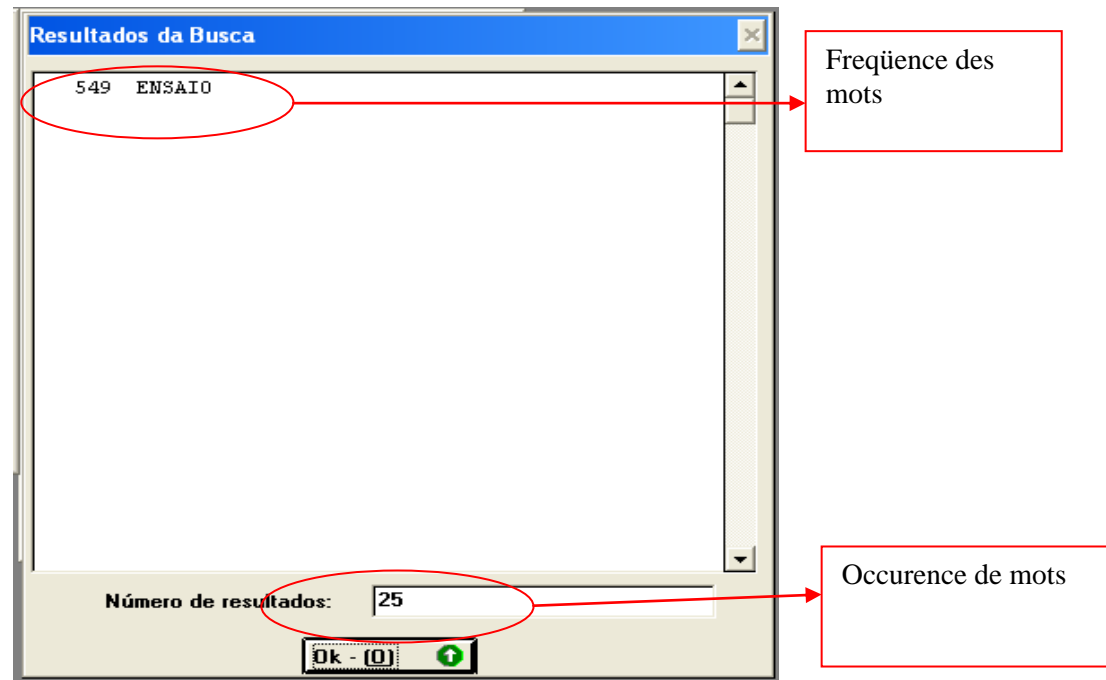


Figure 3: Écran de résultat de recherche

De cette forme, s'utilisant outils de prospection d'informations spécifique plus ce ensemble d'outils *ISIS* ce a été possible de construire le panorama actuel de l'utilisation de l'hydrogène, des matériels nanostructurés et des biocombustible dans contexte mondial à partir du traitement à près de 30.000 brevets.

4 Considérations Finales

Les expériences d'utilisation de veille de l'information dans les unités de recherche scientifique et technologique brésiliennes sont très peues et pionnières. La dissémination de cette pratique au niveau national est apparue grâce à des évènements mais, n'en est encore qu'au début par rapport à son application dans les organismes gouvernementaux, au contraire de ce qui arrive dans le secteur privé.

Du point de vue technique, les résultats préliminaires nous amènent à affirmer que l'utilisation d'outils de TI doit être fait en cherchant toujours à découvrir toutes ses potentialités. Dans le cas de la famille Isis, son efficacité a été prouvée comme un logiciel de traitement automatique d'information de texte, étant un modèle de banque de données plus flexible, et avec un taux élevé de développement dans la récupération d'information dû à son efficacité dans les opérations de recherche d'un élément aléatoire.

Du point de vue institutionnel, les résultats futurs obtenus dans la première série du processus de construction d'un programme de développement technologique, confirme le panorama international déjà connu du spécialiste, tout comme il a montré des voies technologiques peu développées et méritant une plus grande réflexion.

5 Bibliographie

- [1] BERGER, G. *L'attitude prospective*. Revue prospective, n.1, 1958
- [2] QUONIAM, Luc. *Prospecção de Informação: Relatório de consultoria do Programa de Desenvolvimento Tecnológico - PDTI*. Rio de Janeiro, Brasil: INT, 2007. (Note Technique)
- [3] <http://www.dialogbr.com.br/>; <http://www.stn-international.de/>
- [4] ALBUQUERQUE, Frederico L.; LELLIS, Vera. L. M.; SILVA, Cícera. H. *Disponibilização da memória técnico científica do Instituto Nacional de Tecnologia. Relato de experiência*. Salvador, Brasil: 6 CIFORM, 2005.
- [5] CDS/ISIS FOR WINDOWS HANDBOOK. Disponible dans <http://www.unesco.org>. Accédé dans 02/2005
- [6] INTERNATIONAL STANDARD ORGANIZATION: ISO 2709. *Information and documentation – Format for information exchange*. Genève: ISQ, 1996
- [7] LELLIS, Vera Lúcia Maia; et al. *Análise bibliométrica automatizada de acervos, como suporte à política de aquisição e desenvolvimento de coleções*. TECBAHIA: Revista Baiana de Tecnologia, Camaçari, Bahia, Brasil, v.12, n. 2, p. 182-190, maio/ago. 1997.
- [8] TORRICELLA, Raul (Trad). *ISISASCHII utilitário para importación. Versio 0: manual del usuário*. In: Microsoft and visual basic are registered trademark for information exchange. Berna: ISO, 1996. Disponible dans <http://www.infolac.ucol.mx/microisis/new-tools-fr>. Accédé dans 02/2005
- [9] *MANUAL TXTISO*. Disponible em: <http://Jibrary.wur.nl/isis/keyser/txt2iso.htm>. Accédé dans 02/2005
- [10] ORTEGA, C. D. *Introdução ao Microisis*. 2° ed. Brasília: Briquet de Lemos, 2002.
- [11] UNESCO. *ISISASCHII: version 0.95*. Paris: UNESCO, 2004.
- [12] UNESCO. WINISIS 1.5. *Fonctions avancées*. Paris: Unesco, 2005. v. 3.
- [13] UNESCO. WINISIS 1.5. *Le langage de formatage et applications*. Paris: Unesco, 2005. v.2.