

Doc'CISMeF : un outil de recherche Internet orienté vers l'enseignement et la formation à distance en médecine

Stéfan Jacques DARMONI^{1,3}, Jean-Philippe LEROY¹, Magaly DOUYÈRE¹, Josette PIOT², Saïda OUAZIR², Benoît LACOSTE³, Christophe GODARD³, Isabelle RIGOLLE³, Martial BRISOU³, Stéphane VIDEAU³, Myriam QUÉRÉ¹, Eric GOUPY¹, Habib ABDULRAB³, Benoît THIRION²

Résumé *Contexte* : les ressources disponibles sur l'Internet concernant l'enseignement et la formation dans le domaine de la médecine sont en croissance exponentielle. L'objectif de Doc'CISMeF est de créer un outil de recherche générique fondé sur un modèle d'information qui encapsule le thésaurus MeSH. Pour indexer les ressources de qualité, Doc'CISMeF utilise 4 niveaux hiérarchiques : méta-terme, mot-clé, qualificatif et type de ressources. Deux recherches sont possibles dans Doc'CISMeF : une recherche simple qui permet à l'utilisateur de saisir un terme ou une expression ; si ce terme appartient à notre modèle d'information, il sera explosé, c'est-à-dire que Doc'CISMeF affichera toutes les ressources indexées avec des termes hiérarchiquement inférieurs. La recherche avancée permet des requêtes plus complexes, combinant avec des opérateurs booléens, méta-termes, mots-clés, qualificatifs et types de ressources, mais aussi tous les champs de Doc'CISMeF, comme par exemple la cible d'une ressource éducative, notamment le cycle d'études. Doc'CISMeF utilise deux outils standards pour organiser l'information : le thésaurus MeSH et le format de données et de méta-données Dublin Core. Les champs de Doc'CISMeF sont : titre, auteur, description, site éditeur, date, identifiant, format, langue, mots-clés et type de ressources, institution, ville, département ou province ou état, pays, public concerné, coût, parrainage. Doc'CISMeF est un des outils de recherche qui sera utilisé au sein de la future Université Médicale Virtuelle Francophone.

Mots clés Analyse et indexation ; base de données ; catalogage ; descripteur ; enseignement et éducation ; Internet ; stockage et recherche information ; vocabulaire contrôlé ; Support, Non-U.S. Gov't.

Summary *Context*: In the year 2001, the Internet has become a major source of health information both for the health professional and the Netizen. The objective of Doc'CISMeF (D'C) is to create a powerful generic search tool based on an information structure model which encapsulates the MeSH thesaurus. To index resources, D'C uses four levels of hierarchy in its information structure model: « meta-term », keyword, subheading, and resource type. Two levels of searches are available: simple search where the end-user can input a single term or expression. If this term belongs to D'C information structure model, it will be exploded. If not, a full-text search is performed. In the advanced search, complex searches are possible combining Boolean operators with meta-terms, keywords, subheadings and resource types. D'C uses two standard tools for organizing information: the MeSH thesaurus and the Dublin Core metadata format. Resources included in D'C are described by the following elements: title, author or creator, subject and keywords, description, publisher, date, resource type, format, identifier, and language. Doc'CISMeF is one of the search tools of the French-speaking virtual medical university.

Key words Abstracting and indexing; Cataloging; Database; Education; Information Storage and Retrieval; Internet; Subject Headings; Support, Non-U.S. Gov't; Vocabulary controlled.

Pédagogie Médicale 2001 ; 2 : 170-178

¹ Direction Informatique et des Réseaux,

² Bibliothèque Médicale, Centre Hospitalier Universitaire de Rouen,

³ Laboratoire Perception, Système, Information, Institut National des Sciences Appliquées, Rouen ;

Correspondance : SJ. Darmoni - CHU de Rouen - 1, rue Germont - F76031 Rouen Cedex - France - Tél : 02 32 88 88 29
Courriel : Stefan.Darmoni@chu-rouen.fr

Introduction

A l'aube du troisième millénaire, l'Internet est devenu une source majeure d'information dans le domaine de la santé, tant pour l'étudiant et le professionnel de santé que pour le cyber-citoyen¹. Les ressources disponibles dans le domaine de la médecine sur l'Internet concernant l'enseignement et la formation sont en croissance exponentielle. Elles émanent de différentes institutions et, récemment encore, rien ne semblait coordonner leur mise en ligne. Depuis un an, le projet Université Virtuelle Médicale Francophone (UVMF)², consortium de huit facultés de médecine française (Grenoble, Lille, Marseille, Nancy, Paris V, Paris VI, Rennes, Rouen) tente d'y apporter une réponse.

On trouve sur l'Internet, pêle-mêle, cours en ligne, apprentissage par problèmes (APP), questions à choix multiples (QCM), banque de cas cliniques, banque d'images. Retrouver et utiliser une information adéquate en fonction de la question posée n'est pas si aisé sur l'Internet : d'où la profusion de moteurs de recherche et de catalogues³. Mais les catalogues comme Yahoo*, ou les moteurs de recherche comme AltaVista** ont montré leurs limites, respectivement dans la sélection des ressources et dans l'organisation de l'information.

Les cours en ligne, traitant par exemple de l'épilepsie et disponibles à notre connaissance sur au moins cinq sites universitaires, ne sont signalés ni sur Yahoo ni sur Nomade (le 31 octobre 2000). Les retrouver sur AltaVista suppose une gymnastique délicate parmi les 2 220 pages proposées suite à la requête « +épileps* +cours ».

Nous pensons qu'il est nécessaire de recenser, sélectionner, classer et indexer l'information disponible, à l'aide de thésaurus dédiés à la médecine, et de la décrire au moyen de normes communes. L'objectif de CISMeF⁴ (Catalogue et Index des Sites Médicaux Francophones) est d'assister le professionnel de santé pendant sa recherche d'informations disponibles sur l'Internet. CISMeF est un projet initié par le Centre Hospitalier Universitaire (CHU) de Rouen***. CISMeF a été créé en même temps que le site du CHU en février 1995.

Ce catalogue recense actuellement 10 190 ressources

(sites et documents) et indexe, en priorité, les sites institutionnels et les documents en émanant, avec trois axes prioritaires : l'enseignement et l'éducation des professionnels de santé et des apprenants, c'est-à-dire les étudiants en médecine, la médecine factuelle (recommandations pour la bonne pratique clinique et conférences de consensus), et l'information destinée aux patients et aux cyber-citoyens afin d'améliorer l'éducation sanitaire dans le monde francophone.

Le principal défaut de CISMeF résidait dans son faible niveau technique, du fait qu'il était construit uniquement en HTML statique qui ne permet pas d'interroger une banque de données. CISMeF ne permettait pas notamment d'effectuer des recherches complexes, comme par exemple les cours sur l'hépatite virale ou les recommandations pour la bonne pratique clinique en neurologie. D'où la décision de débiter la conception de Doc'CISMeF (D'C) en décembre 1999.

Le but de D'C est de créer un outil de recherche proposant des possibilités de recherche plus performantes et étendues que celles de CISMeF mais utilisant la même structure du modèle d'information qui encapsule le thésaurus MeSH. D'C est un projet également initié au CHU de Rouen****. Actuellement D'C contient plus de 5 700 ressources dont plus de 1 400 concernant l'enseignement (voir Tableau 1). Depuis août 2000, toute ressource incluse dans CISMeF l'est également dans D'C. Nous prévoyons que les deux outils auront le même nombre de ressources en septembre 2001.

Construction de Doc'CISMeF

Logiciel

Le prototype de D'C a été implanté en juin 2000 sur une machine Wintel. D'C est un site Web utilisant un serveur Apache et un Apache Jserv pour l'exécution de servlets. L'outil de recherche a été développé en Java car c'est une solution ouverte, bon marché, puissante et portable. La base de données utilisée est MS Access. Les servlets Java génèrent des requêtes SQL *via* des drivers ODBC qui lisent la base de données sous Access.

*URL (Universal Resource Locator) : <http://www.yahoo.com>

**URL : <http://www.altavista.com>

***Ses URL sont <http://www.chu-rouen.fr/cismef> ou <http://www.cismef.org>

****URL : <http://doccismef.chu-rouen.fr/>

Standards et modèle d'information

D'C et CISMef utilisent deux outils standards pour organiser l'information : le thésaurus MeSH (Medical Subject Headings) de la Bibliothèque Nationale Américaine de Médecine (NLM-US) et le format de données et de méta-données de l'initiative Dublin Core⁸⁻⁹. Le thésaurus MeSH, utilisé notamment par la base de données bibliographique Medline, contient 19 942 termes et 82 qualificatifs dans sa version 2001. Les termes français du MeSH sont ceux de la traduction de l'INSERM (Institut National de la Santé et la Recherche Médicale)*, et plus spécifiquement du réseau DISC-DOC.

Les qualificatifs du MeSH permettent de restreindre une recherche sur un sous-domaine d'un terme MeSH, comme, par exemple, la toxicité du chlore (noté chlore/toxicité ou chlorure/toxicity en anglais). Nous utilisons également la traduction française des qualificatifs**.

Du fait de la mise à jour de la base de données de Doc'CISMef plus facile que CISMef statique, les ressources incluses dans D'C sont décrites avec plus de termes MeSH (une moyenne de 4,4 termes MeSH dans D'C comparé à 1,3 dans CISMef au début 2000) et de qualificatifs que précédemment dans CISMef.

Le modèle d'information de D'C encapsule le thésaurus MeSH avec deux autres niveaux également organisés sous forme de thésaurus : les méta-termes [10] (N=57) et les types de ressources [11] (N=101). Le méta-terme est généralement une spécialité médicale ou biologique, comme la cancérologie ou la bactériologie. L'idée du concept « méta-terme » est né du fait de la limite relative des termes MeSH qui entraîne du silence quand on cherche « cours en cardiologie » ou « problèmes et exercices en virologie », ne retrouvant pas, par exemple, ce qui concerne la pathologie cardiologique ou virale. Quand nous recherchons des informations sur la cardiologie par exemple, le terme MeSH « cardiologie » ne concerne pas les pathologies cardiologiques ou les agents cardiovasculaires. Pour une recherche sur des « cours en cardiologie », il faut pouvoir obtenir des cours sur les cardiopathies ou les anti-hypertenseurs. Ce qui est aujourd'hui tout à fait possible avec ce concept de « méta-terme »***.

Tableau 1 : Types de ressources concernant l'enseignement *

- enseignement et éducation (6)
 - formation (13)
 - matériel enseignement (101)
 - cours (809)
 - film éducatif (2)
 - manuel enseignement (18)
 - mode d'emploi (0)
 - problèmes et exercices (33)
- structure enseignement (168)
 - établissement enseignement médical ou apparenté (32)
 - école d'infirmiers (13)
 - école dentaire (23)
 - école santé publique (6)
 - école vétérinaire (8)
 - faculté de médecine (71)
 - faculté de pharmacie (39)

* le nombre entre parenthèses indique le nombre de ressources indexées dans Doc'CISMef avec ce type de ressources sans explosion.

Dans CISMef, les méta-termes n'avaient des liens sémantiques qu'avec les arborescences ou les mots-clés MeSH. Par exemple, dans la page CISMef « cancérologie », apparaissent successivement : les sites généraux de cette spécialité indexés avec le mot-clé « oncologie médicale » et une liste d'hyper-liens vers les arborescences et les mots-clés de cette spécialité ; les arborescences sont : antinéoplasiques, oncologie médicale, marqueur biologique tumeur, et, tumeurs ; les termes MeSH sont : cancérogènes, service oncologie hôpital ; le qualificatif est : secondaire ; et le type de ressources est service oncologie hôpital.

En concevant D'C, nous avons ajouté des liens sémantiques avec les deux autres niveaux de notre modèle d'information : le qualificatif et les types de ressources. Ainsi, le méta-terme « cancérologie » a des hyper-liens vers le qualificatif « secondaire » et le type de ressources « service oncologie hôpital ».

Les types de ressources sont une généralisation du

* URL : http://dicdoc.kb.inserm.fr:2010/basismesh/frm_mesh2000.html

** URL : <http://dicdoc.kb.inserm.fr:2010/BASIS/elgis/fqmb/qualif/SF>

*** La liste des méta-termes est disponible à l'URL suivante : <http://www.chu-rouen.fr/ssfs/santspe.html>

concept « type de publication » de Medline. Nous avons notamment ajouté des types qui sont spécifiques des ressources disponibles sur l'Internet comme « association », « information patient » ou « réseaux coordonnés ». Les types de ressources de D'C sont organisés sous forme hiérarchique permettant leur explosion comme les méta-termes, les mots-clés ou les qualificatifs*. Les types de ressources décrivent la nature de la ressource alors que le mot-clé MeSH en décrit le sujet. Ainsi, par exemple, dans le cas d'une recommandation sur l'intoxication au monoxyde de carbone, le mot-clé MeSH est « intoxication au monoxyde de carbone » et le type de ressources est « ligne directrice pratique médicale ».

Description de la base de données de D'C et génération des notices

D'C partage avec CISMeF la même méthode pour construire le catalogue qui comprend quatre étapes : (1) le recensement des ressources, (2) le filtre ou la sélection, (3) la description et (4) l'indexation.

Etape 1 : Recensement

Une veille quotidienne est effectuée sur les annuaires multidisciplinaires Nomade, Toile du Québec, Yahoo...**. Douze pour cent des administrateurs de sites recensés dans D'C nous signalent également leurs réalisations. Nous effectuons enfin une veille complémentaire généralement hebdomadaire sur les sites producteurs de ressources (ministères, universités, sociétés savantes...). Cette étape est réalisée par une assistante documentaliste qui effectue le recensement et la veille informationnelle.

Etape 2 : Filtre ou sélection

Pour inclure uniquement des ressources de qualité dans D'C, nous utilisons les principaux critères du Net Scoring [12]. Le Net Scoring est une grille de 49 critères pour évaluer la qualité de l'information de santé sur l'Internet. Parmi ceux-ci sont, notamment, pris en compte l'identité de l'auteur ou de l'institution édi-

trice, la date de mise à jour pour un document et la navigabilité. Nous attachons une attention toute particulière au fait que les auteurs doivent être clairement identifiés, les ressources disponibles sur l'Internet concernant la formation et l'enseignement validées par les enseignants, et les cibles indiquées avec précision. Cela nécessite parfois quelques courriels avec les responsables des sites. Certaines ressources ne sont pas incluses dans D'C parce qu'elles ne respectent pas des critères essentiels, en particulier déontologiques. Cette étape est réalisée par le responsable de la bibliothèque, aidé en cas de doute par un comité d'experts.

Etape 3 : Description

Après le processus de sélection, chaque ressource est saisie dans la banque de données Access par trois assistantes documentalistes. Chaque ressource est décrite par les éléments suivants parmi les 15 du projet Dublin Core*** : titre, auteur, description, site éditeur, date, identifiant, format, langue, mots-clés et type de ressources. D'autres champs sont spécifiques de D'C : institution, ville, département ou province ou état, pays, public concerné, coût, parrainage.

Chaque ressource présente dans la base de données génère automatiquement : (a) la notice à inclure dans CISMeF, (b) la notice complète de D'C disponible au format HTML. Au sein de cette notice complète, il est possible de lancer une autre recherche en cliquant sur un mot-clé, un qualificatif ou un type de ressource. Il est également possible de cliquer sur un bouton à côté de chaque mot-clé qui donne accès à la page MeSH de CISMeF correspondante, incluant arborescences, relations « voir aussi » et synonymes en français. Ces différents éléments permettent d'élargir la recherche de départ.

Etape 4 : Indexation

Cette étape est réalisée par une documentaliste. Le responsable de la bibliothèque médicale est en charge de la « super-indexation », c'est-à-dire la vérification de cette indexation. Ce fonctionnement s'est inspiré par celui utilisé par la Bibliothèque Nationale Américaine

* La liste des types de ressources est disponible à l'URL suivante : <http://www.chu-rouen.fr/documed/typeressource.html>

** Liste complète disponible à l'URL : <http://www.chu-rouen.fr/documed/docum.html#VEILLE>

*** URL : http://purl.org/DC/about/element_set.htm

de Médecine (NLM-US) pour construire Medline. Enfin, nous organisons des réunions régulières entre le bibliothécaire médical et les deux médecins informaticiens pour une éventuelle seconde validation.

Il existe finalement trois possibilités d'indexer les ressources disponibles sur l'Internet concernant l'enseignement et la formation au sein de D'C qui sont accessibles par le méta-terme Enseignement et Education* :

- par mots clés de la catégorie I2 (Enseignement et éducation)**
- par qualificatifs / ed, affiliés aux mots clés appropriés selon les recommandations de la NLM, c'est-à-dire à des termes appartenant aux catégories suivantes du thésaurus MeSH : F4 (Disciplines et activités comportementales), G1-2 (Sciences biologiques - Profession santé), I3 (Activités humaines) et M. (Individus).
- par type de ressources, lorsque les règles d'usage des qualificatifs ne permettent pas leur utilisation. Dans ce cas et suivant l'importance de la ressource, la spécialité est ajoutée, avec le qualificatif / ed.

Fonctionnement de Doc'CISMeF

D'C est un outil de recherche permettant de retrouver des ressources francophones en santé de qualité sur l'Internet. Ce site Web est principalement dévolu aux étudiants et aux professionnels de santé, mais le citoyen y a également accès : il n'existe aucune restriction d'accès à D'C.

D'C permet deux types de recherche : la recherche simple et la recherche avancée.

Recherche simple

La recherche simple permet de saisir un terme unique (par exemple, SIDA) ou une expression (par exemple, « hépatite virale humaine »). D'C permet une recherche bilingue en français (avec ou sans accents) ou en anglais, et dans les deux cas, en minuscules ou majuscules, si le terme saisi est un mot « réservé ». Un terme est défini comme réservé si et seulement s'il

appartient à au moins un des niveaux de notre modèle : méta-terme, mot-clé, qualificatif ou type de ressource. Dans ce cas, le terme réservé X est explosé, c'est-à-dire que toutes les ressources avec un terme situé hiérarchiquement sous ce terme X seront sélectionnées : par exemple, une recherche sur le terme « viroses » retrouvera les termes hiérarchiquement inférieurs comme « grippe » ou « SIDA ». Si le terme réservé X appartient à plusieurs niveaux de notre modèle d'information (par exemple, virologie est à la fois un méta-terme, un mot-clé et un qualificatif), D'C unifie le résultat des explosions de tous les niveaux selon la formule ci-dessous :

$$\bigcup_{i=1}^4 \text{exp}(x)$$

où i est le niveau du modèle d'information et exp est la fonction d'explosion.

Si le terme saisi n'est pas un mot réservé, une recherche en texte intégral est effectuée sur les champs principaux de D'C : titre, auteur, ville, pays, description, méta-terme, mot-clé, qualificatifs, type de ressource, site éditeur, date de publication et URL. Cette recherche en texte intégral s'effectue par défaut avec une double troncature, droite et gauche. Par exemple, l'étudiant cherchant de l'information sur le « zen » retrouvera plusieurs ressources : la première car l'un des auteurs s'appelle « Rozenberg » qui contient l'occurrence « zen » ; la seconde parce que cette ressource est indexée au mot-clé « benzène » qui contient également l'occurrence « zen ». Cet exemple illustre le bruit généré par ce type de recherche, très proche de celle effectuée sur les moteurs de recherche comme Google ou AltaVista. C'est pourquoi cette recherche en texte intégral n'intervient que si l'utilisateur n'a pas saisi un mot réservé.

Si la requête n'apporte aucune réponse, une recherche complémentaire est proposée sur notre moteur de recherche interne et vers les autres outils de recherche de l'UMVF.

*URL : <http://www.chu-rouen.fr/ssfform.html>

**Les mots-clés concernant l'enseignement, regroupés dans Medline au sein de la catégorie I2, sont, pour la majorité d'entre eux, accessibles à partir de cette page Enseignement et Education [<http://www.chu-rouen.fr/ssflanthropfr.html#enseignementeteducation>].

La liste des mots-clés utilisés dans D'C concernant l'enseignement est disponible dans le tableau 2. Mais, respectant en cela les recommandations de la NLM, le qualificatif / enseignement et éducation (/ ed) peut aussi être affilié à un mot-clé, par exemple, ophtalmologie / enseignement et éducation

Tableau 2 : Mots clés MeSH concernant l'enseignement

- corps enseignant
 - corps enseignant infirmier
 - corps enseignant médical
- enseignement non professionnel
 - éducation sanitaire
 - éducation malade
 - éducation sexuelle
 - enseignement spécialisé
 - éducation spécialisée personne handicapée mentale
 - intégration scolaire enfants handicapés
 - étudiant
 - formation professionnelle
- enseignement professionnel
 - enseignement infirmier
 - formation continue infirmier
 - enseignement médical
 - enseignement médical post-universitaire
 - enseignement médical premier cycle
 - internat et résidence
 - enseignement pharmacie
 - enseignement pharmacie post-universitaire
 - enseignement préparatoire diplôme
 - formation continue
 - enseignement médical post-universitaire
 - enseignement pharmacie post-universitaire
 - formation continue infirmier
- établissement scolaire
 - établissement enseignement médical ou apparenté
 - école d'infirmiers
 - école dentaire
 - école santé publique
 - école vétérinaire
 - faculté de médecine
 - faculté de pharmacie
- étudiant
 - étudiants profession médicale ou paramédicale
 - élève infirmier
 - étudiant médecine
 - étudiant pharmacie
- matériel enseignement
 - matériel audio-visuel
- pédagogie
- programme études
 - intégration scolaire enfants handicapés
- stage pratique guidé

Recherche avancée

La recherche avancée permet une requête sur les principaux champs de D'C en utilisant les opérateurs booléens, ET, OU SAUF. L'explosion s'effectue par défaut mais elle peut être désactivée ; ainsi, l'utilisateur choisira « hôpital » comme type de ressource mais désactivera l'explosion pour éviter les ressources indexées avec des types de ressources hiérarchiquement inférieurs à « hôpital » comme « centre traitement toxicomanie » par exemple.

Une recherche avancée peut être limitée selon deux axes : (a) l'indication du niveau de preuve, qui a été défini par un groupe de travail mixte Ministère de la Santé/Conseil National de l'Ordre des Médecins comme le critère principal de qualité du contenu médical disponible sur l'Internet ; (b) les ressources textuelles qui permettent d'effectuer une recherche sur le noyau dur de D'C. Dans ce cadre, D'C construit une encyclopédie de santé accessible librement sur l'Internet.

Propriétés communes aux recherches simples et avancées

Pour les deux recherches, le résultat d'une recherche s'affiche comme une liste d'enregistrements. Quatre modes d'affichage plus ou moins détaillés sont possibles. Quel que soit le choix d'affichage, un accès à la notice complète est proposé.

Le nombre de notices affichables est également paramétrable : 10, 20, 50 (choix par défaut) et 100. Les notices sont affichées soit (a) par date de publication, (b) dans un ordre alphabétique des titres des ressources, (c) par type de ressources et (d) par pays et villes.

Si le terme ou l'expression recherché(e) appartient au thésaurus MeSH, une recherche complémentaire sur PubMed est proposée. Une aide à l'utilisation de D'C est disponible à l'URL suivante :

<<http://doccismef.chu-rouen.fr/aide.html>>. Enfin, pour toujours faciliter la recherche aux non-initiés à la recherche d'information, en nous inspirant de l'exemple d'Ovid*, nous proposons des requêtes pré-programmées sur les champs suivants : auteurs, cycles des études médicales, départements français, mots-clés

MeSH, pays, publics concernés, sites éditeurs, et types de ressources. Ainsi, l'utilisateur final pourra avec un clic de souris consulter, par exemple, toutes les ressources produites par le Sénat français ou celles indexées avec le type de ressources « hôpital enfant ». A l'inverse de PubMed, l'utilisateur saisissant dans la recherche simple un terme MeSH, D'C n'affiche que les ressources indexées avec ce terme et n'effectue pas une recherche en texte intégral complémentaire. N'ayant pas dans D'C d'équivalent à PreMedline (articles inclus dans Medline avant indexation), toutes les ressources de D'C sont indexées dès leur inclusion. Nous avons refusé d'introduire du bruit avec une recherche en texte intégral ; celle-ci est néanmoins possible dans la recherche avancée en choisissant l'option « tous les champs », ce qui est utile quand la recherche simple ne ramène que peu de résultats.

Intérêts et limites de Doc'CISMeF

En accord avec le financement partiel en 1998 par l'Agence Universitaire de la Francophonie, l'équipe CISMeF s'est engagée à aider les professionnels de santé des pays francophones en voie de développement, dans leur quête d'informations de santé de qualité. Concernant les ressources disponibles sur l'Internet concernant l'enseignement et la formation, une collaboration a été instituée avec la commission pédagogique de la CIDMEF, Conférence Internationale des Doyens de facultés de Médecine d'Expression Française** 5. D'C est un des outils de recherches qui sera utilisé par la future Université Médicale Virtuelle², en compagnie d'Ariane⁶ développé à Marseille, MAOUSSC⁷ développé à Rennes, Medicadoc de France Telecom et les moteurs de recherche de Keo, avec lesquels nous avons programmé une interopérabilité.

D'C est un outil générique. Il est utilisable en anglais et dans chaque langue dans laquelle le MeSH a été traduit (allemand, espagnol, finnois, italien, portugais et russe¹³).

La force principale de D'C réside dans son modèle d'information. D'C encapsule le MeSH (arborescence, mot-clé, qualificatif) avec deux niveaux supplémentaires : l'un au-dessus (méta-terme) et l'un au-dessous (type de ressources). Ce modèle partagé avec CISMeF

*Société commerciale qui propose un accès intégré à des bases de données bibliographiques et des journaux électroniques en texte-intégral (URL : <http://www.ovid.com>)

**URL : <http://www.cidmef.u-bordeaux2.fr/>

a été récemment amélioré : à présent, un méta-terme a des liens hiérarchiques avec tous les niveaux inférieurs. A notre connaissance, D'C est le premier outil de recherche utilisant la principale qualité du thésaurus MeSH (l'explosion) sur chacun de ces niveaux.

Les utilisateurs de D'C doivent avoir une connaissance minimum du MeSH pour l'employer avec un maximum de pertinence, d'où la nécessité de former les utilisateurs. Notre objectif à long terme concernant D'C est d'être aussi riche en termes fonctionnels que les outils d'accès à Medline qu'ils soient publics ou privés. Des améliorations doivent être apportées à D'C :

1. une fonction d'adéquation automatique comme celle existante dans le site Web d'Internet Grateful Med* pour permettre aux utilisateurs inexpérimentés de taper un mot ou une expression quelconque et de voir proposer une liste de termes MeSH voisins.

2. un module d'interrogation pas à pas, comme il existe dans Ovid ou PubMed** permettant des recherches encore plus complexes que celles possibles avec la recherche avancée (incluant, par exemple, plus de quatre termes MeSH).

Une recherche par méta-terme est pertinente puisqu'elle inclut tous les mots réservés de cette spécialité sans que l'utilisateur connaisse ces derniers. Cette recherche par méta-terme est surtout utile quand la recherche par mot-clé MeSH donne peu de résultats.

Nous supposons que la recherche simple est plutôt utilisée par le cyber-citoyen et le professionnel de santé néophyte alors que la recherche avancée est plutôt le fait des bibliothécaires et des professionnels de santé formés à D'C mais une évaluation formelle devrait confirmer ou non cette intuition.

D'C peut également être utilisé comme un traducteur de requêtes français-anglais pour effectuer une recherche sur Medline via PubMed. Malheureusement, peu d'internautes, y compris chez les bibliothécaires connaissent le MeSH en français. Cette fonctionnalité de D'C sera donc utilisée par hasard, du moins avant formation, sur le thésaurus MeSH en français et en anglais. A l'inverse, de nombreux bibliothécaires connaissant le MeSH en anglais peuvent interroger D'C dans cette langue.

Conclusion

Le nombre de ressources concernant l'enseignement de la médecine est en croissance permanente et il est essentiel, pour les étudiants comme pour les médecins en exercice, d'avoir accès à un outil de recherche performant spécialisé dans le recensement des ressources médicales de qualité.

Conçu sur un modèle d'information très structuré intégrant notamment le thésaurus MeSH de la NLM, Doc'CISMeF offre, d'ores et déjà, de larges possibilités d'interrogation. Son intérêt majeur réside dans les nombreuses options de recherche qui permettant des requêtes complexes incluant ou non « l'explosion » des arborescences de mots-clés.

A court terme, il est notamment prévu d'augmenter le nombre de ressources concernant la médecine factuelle et l'enseignement. D'C a d'ailleurs été choisi comme un des outils de recherche de l'Université Médicale Virtuelle Francophone. Une collaboration européenne avec d'autres outils de recherche (DDRT-de, HON-Ch, OMNI-Uk) est également à l'étude. Enfin, il semble primordial de mettre en place un système d'évaluation de l'utilisation de D'C chez les étudiants, les professionnels, et le grand public, l'objectif premier de D'C étant d'assister les internautes dans leurs recherches d'informations de qualité sur Internet.

Remerciements

Doc'CISMeF est financé en partie par le programme Educnet [<http://www.educnet.education.fr/res/bprojets.htm>] du Ministère de la Recherche (2000). CISMeF a reçu le prix Albert Sézary de l'Académie Nationale de Médecine (2000).

*URL : <http://igm.nlm.nih.gov/>

**URL ; <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>

Références

1. Schatz BR. *Information Retrieval in Digital Libraries : Bringing Search to the Net*. Science 1997 ; 34 : 275-327.
2. LeBeux P, Duff F, Fresnel A, et al. *The French Virtual Medical University*. In : *Proceedings of MIE 2000, Sixteenth International Congress of the European Federation for Medical Informatics, 2000 ; 554-561*.
3. Flannery MR. *Cataloging Internet resources*. Bull Med Libr Assoc 1995 ; 83(2) : 211-215.
4. Darmoni SJ, Leroy JP, Thirion B, Baudic F, Douyère M, Piot J. *CISMeF : a structured Health resource guide*. Meth InfMed 2000 ; 39(1) : 30-35.
5. Denef J.-F. *Nouvelles technologies et professionnalisation médicale*. In : *Actes du Colloques XIIIèmes Journées Universitaires Francophones de Pédagogie Médicale, 1999 ; 14-16*.
6. Fieschi M, Aymard S, Mobarhan H, Fieschi D, Volot F, Joubert M. *Le projet ARIANE : un aspect de l'interopérabilité sémantique pour l'accès aux bases de connaissances*. In : *Actes du colloque CRISTAL'S, Partage d'information en santé 2000 ; 100-108*.
7. Burgun A, Bodenreider O, Denier P, Delamarre D, Botti G, Oberlin P, Leveque JM, Bremond M, Fieschi M, Le Beux P. *A collaborative approach to building a terminology for medical procedures using a Web-based application : from specifications to daily use*. Medinfo 1998 ; 9(1) : 596-599.
8. Weibel S, Juha H. *DC-5 : The Helsinki Metadata Workshop ; A Report on the Workshop and Subsequent Developments*. D-Lib Magazine 1998 February. Disponible sur Internet : <<http://www.dlib.org/dlib/february98/02weibel.html>>.
9. Darmoni SJ, Thirion B, Leroy JP, Douyère M, Piot J. *The Use of Dublin Core Metadata in a Structured Health Resource Guide on the Internet*. Bull Med Libr Assoc 2001 ; (in press).
10. Thirion B, Darmoni SJ. *Simplified access to MeSH Tree Structures on CISMeF*. Bull Med Libr Assoc 1999 ; 87(4) : 480-481.
11. Darmoni SJ, Thirion B. *A standard metadata scheme for health resources*. J Am Med Inform Assoc 2000 ; 7(1) : 108-109.
12. Centrale Santé. *Net Scoring : critères de qualité de l'information de santé sur l'Internet 20 Apr 2000 [Document Web ; visité le 6 Octobre 2000]*. Disponible sur Internet : <<http://www.chu-rouen.fr/netscoring>>.
13. Nelson SJ, Schopen M, Schulman JL, Arluk N. *An Interlingual Database of MeSH Translations*. 20 Aug. 2000 [Document Web ; visité le 6 Octobre 2000]. Disponible sur Internet : <http://www.nlm.nih.gov/mesh/intlmesh.html>
14. National Library of Medicine. *Fact Sheet UMLS Metathesaurus*. 12 Aug 1998 [Document Web ; visité le 6 Octobre 2000]. Disponible sur Internet : <<http://www.nlm.nih.gov/pubs/factsheets/online-indexing-system.html>>