

ISTITUTO SUPERIORE DI SANITÀ

Convegno
Il MeSH parla anche italiano
nella ricerca, l'editoria, la promozione della salute

Istituto Superiore di Sanità
Roma, 27 giugno 2006

Atti a cura di
Maria Cristina Calicchia e Maurella Della Seta
Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali

ISSN 1123-3117
Rapporti ISTISAN
06/49

Istituto Superiore di Sanità

Convegno. Il MeSH parla anche italiano nella ricerca, l'editoria, la promozione della salute. Istituto Superiore di Sanità. Roma, 27 giugno 2006.

Atti a cura di Maria Cristina Calicchia e Maurella Della Seta
2006, 65 p. Rapporti ISTISAN 06/49

Dal 1978 il Settore Documentazione dell'Istituto Superiore di Sanità (ISS) è riconosciuto quale Centro di riferimento, a livello nazionale, per il sistema MEDLARS (*MEDical Literature Analysis and Retrieval System*) dalla *National Library of Medicine* (NLM) di Bethesda. Nell'ambito della collaborazione con l'NLM si realizza la traduzione italiana dei MeSH (*Medical Subject Headings*), nel quadro di un progetto internazionale UMLS (*Unified Medical Language System*), avviato dall'NLM nel 1986, che mira alla standardizzazione della terminologia biomedica e di sanità pubblica. L'edizione 2006 dei MeSH contiene 23.885 descrittori, che permettono di indicizzare, recuperare, organizzare e gestire oltre 14 milioni di citazioni di articoli per la base dati MEDLINE/PubMed. Il progetto di traduzione italiana dei MeSH è iniziato nel 1998 ed è stato completato nell'ottobre del 2004. La traduzione viene aggiornata regolarmente a livello annuale e risulta disponibile sul sito dell'ISS. Il Convegno ha offerto una ricognizione delle utilizzazioni della traduzione italiana dei MeSH nei settori della ricerca, dell'editoria scientifica e della promozione della salute.

Parole chiave: Medical Subject Headings, Traduzioni, Italia, Unified Medical Language System, Basi di dati

Istituto Superiore di Sanità

Meeting. MeSH speaks Italian in research, publishing, health promotion. Istituto Superiore di Sanità. Rome, 27 June 2006.

Proceedings edited by Maria Cristina Calicchia and Maurella Della Seta
2006, 65 p. Rapporti ISTISAN 06/49 (in Italian)

Since 1978 the Documentation Service has been the national Reference Center for MEDLARS (*MEDical Literature Analysis and Retrieval System*) of the *National Library of Medicine* (NLM), Bethesda, USA. In the framework of the collaboration between the ISS and the NLM, the Italian translation of *Medical Subject Headings* (MeSH) was conceived to join the *Unified Medical Language System* (UMLS) project for the standardization, at an international level, of biomedical and health-related terminology. 2006 MeSH release includes 23,885 main descriptors for indexing, retrieval, organising and management of over 14 millions of citations of articles in MEDLINE/PubMed. The Italian translation of Mesh started in 1998 and was completed by 2004. The translation, available since November 2004 on ISS website, has been yearly updated. The meeting offered an overview of utilization of the Italian translation of MeSH in research, scientific publishing and health promotion.

Key words: Medical Subject Headings, Translations, Italy, Unified Medical Language System, Databases

Direzione scientifica del Convegno

Maria Cristina Calicchia, Maurella Della Seta (Settore Documentazione) e Franco Toni (Settore Biblioteca).

Segreteria tecnico-scientifica

Alessandra Ceccarini, Italo Gentilini (Settore Documentazione).

Per informazioni su questo documento scrivere a: documentazione@iss.it

Il rapporto è accessibile online dal sito di questo Istituto: www.iss.it.

Presidente dell'Istituto Superiore di Sanità e Direttore responsabile: *Enrico Garaci*
Registro della Stampa - Tribunale di Roma n. 131/88 del 1° marzo 1988

Redazione: *Paola De Castro, Sara Modigliani e Sandra Salinetti*
La responsabilità dei dati scientifici e tecnici è dei singoli autori.

© Istituto Superiore di Sanità 2006

INDICE

| | |
|---|-----|
| Presentazione | iii |
| Introduzione al Convegno <i>Maria Cristina Calicchia</i> | 1 |
| Il linguaggio medico universale e la traduzione italiana dei <i>Medical Subject Headings</i> <i>Maurella Della Seta</i> | 4 |
| Soggettario della biblioteca dell'Istituto Superiore di Sanità: esperienze d'efficacia e precisione scientifica nell'impiego delle voci MeSH <i>Maria Alessandra Falcone, Paola Ferrari</i> | 12 |
| Implementazioni e sviluppi dell'applicativo web <i>Medical Subject Headings</i> – traduzione italiana <i>Maurizio Ferri, Paolo Roazzi</i> | 19 |
| L'applicazione del MeSH nella biblioteca digitale Azalea: dal linguaggio del paziente al linguaggio controllato <i>Gaetana Cognetti, Katuscia Dormi, Paola Seroppi, Ivana Truccolo</i> | 25 |
| L'utilizzo dei MeSH in italiano nell'evoluzione del progetto ELISIR <i>Paolo Gardois, Sandro Garoffolo, Diego Arduino, Michele Mannucci, Renata Riva, Marcello Valente</i> | 32 |
| Impiego dei MeSH nel processo editoriale: dalla <i>submission</i> alla pubblicazione <i>Paola De Castro, Federica Napolitani, Elisabetta Poltronieri</i> | 37 |
| Conoscenza e web semantico nell'area salute, sanità e prevenzione <i>Paola Capitani</i> | 44 |
| Aggiornamento delle modalità di ricerca in PubMed <i>Scilla Pizzarelli, Rosaria Rosanna Cammarano</i> | 49 |

PRESENTAZIONE

È con vivo piacere che introduco questo Convegno sulla traduzione italiana dei MeSH. Questo prodotto di grande importanza per la standardizzazione di una terminologia biomedica e di sanità pubblica è frutto di una collaborazione tra due Settori dell'Istituto Superiore di Sanità (ISS): il Settore Documentazione e il Settore Informatico, che precedentemente costituivano due Servizi separati, e che la nuova riorganizzazione dell'ISS ha unito in un unico Servizio, che comprende anche la Biblioteca e le Attività Editoriali, per riconoscerne la necessità di integrazione e sinergia.

Gli aspetti della traduzione italiana dei MeSH legati alla creazione del prodotto e alla sua applicazione sono tra gli obiettivi di questo Convegno, da parte mia in qualità di Direttore del Servizio sento opportuno informare che l'Istituto Superiore di Sanità sta riservando, ormai da tempo, una attenzione particolare alle problematiche inerenti a riservatezza, integrità, sicurezza dei dati informatici. Ciò è dovuto oltre che ad una necessità tecnologica al dovere imposto dalla vigente normativa (DL.vo 196/2003 sul trattamento dei dati personali e tutela della privacy).

Per prevenire e combattere azioni intrusive, l'ISS si è dotato di strumenti deputati a garantire un livello adeguato di sicurezza e i più importanti sistemi attivati sono i seguenti:

– *Antivirus enterprise e perimetrale*

I due antivirus agiscono, a diversi livelli, in modo da attuare un rilevamento tempestivo di ogni minaccia virale.

La sicurezza interna o di *enterprise* riguarda l'insieme degli accorgimenti idonei ad assicurare la protezione dei dispositivi *server* e *storage*, contenenti dati sensibili, da accessi indesiderati provenienti dall'interno della propria rete. La sicurezza perimetrale è relativa, invece, al governo delle interazioni tra la propria rete ed entità esterne quali collegamenti con enti terzi e/o reti pubbliche quali, per esempio, Internet. La sicurezza perimetrale viene tipicamente realizzata tramite l'adozione di sistemi *firewall* che permettono il controllo delle comunicazioni in transito tra la propria rete e il mondo esterno.

– *Virtual Private Network*

L'utilizzo consente una riduzione dei costi delle connessioni e una comunicazione sicura, perché criptata, pur utilizzando tratte di rete pubblica.

– *Wireless*

Le connessioni *wireless* e *laser* rappresentano una interessante evoluzione tecnologica. Sono stata implementate per i collegamenti tra la sede principale e quella remota dell'Istituto e in particolar modo nelle connessioni riguardanti le aule didattiche.

Auguro al Convegno il successo meritato e ringrazio tutti voi della presenza e dell'attenzione.

Claudio Di Benedetto

*Direttore del Servizio Informatico, Documentazione,
Biblioteca ed Attività Editoriali,
Istituto Superiore di Sanità, Roma*

INTRODUZIONE AL CONVEGNO

Maria Cristina Calicchia

Settore Documentazione – Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali

Istituto Superiore di Sanità, Roma

mariacristina.calicchia@iss.it

Il Convegno che ho l'onore di introdurre tratterà della traduzione italiana dei MeSH (*Medical Subject Headings*) – che è stata realizzata dal Settore Documentazione di questo Istituto – ma, soprattutto, vuole offrire una ricognizione delle utilizzazioni della traduzione nei settori della ricerca, dell'editoria scientifica e della promozione della salute.

Non mi soffermerò a parlare dei MeSH, ampiamente conosciuti a livello internazionale e che hanno accompagnato la nostra storia di ricercatori, ma trovo utile un richiamo a tempi e date in cui si sono registrati eventi che hanno portato a grandi cambiamenti e motivato la traduzione stessa.

Nel 1978, in seguito alla firma di un protocollo di collaborazione scientifica, tra il governo italiano e quello statunitense, inizia la collaborazione tra l'Istituto Superiore di Sanità (ISS) e la *National Library of Medicine* (NLM) di Bethesda e da allora, il Servizio Documentazione dell'ISS è riconosciuto quale Centro di riferimento, a livello nazionale, per il sistema MEDLARS (*MEDical Literature Analysis and Retrieval System*).

Nel 1980 erano ancora tempi in cui le connessioni avvenivano via modem, attraverso un terminale video/scrivente, e l'ISS coordinava contratti di sub-utenza MEDLARS per le istituzioni che ne facevano richiesta. Il sistema di recupero dell'informazione attraverso il MEDLARS era piuttosto complesso e richiedeva una formazione altamente specifica, il ricercatore doveva affidarsi all'intermediazione del documentarista e per dirla in breve l'approccio orientato a facilitare l'utente finale (*user oriented*) era ancora molto lontano e con esso anche la gratuità dell'informazione: l'uso del PubMed gratuito arriverà, diciassette anni dopo, nel 1997. I costi di connessione e i costi di interrogazione erano ancora piuttosto alti e la penuria di personale dedicato era presente anche allora: dalla Relazione sulle attività del Servizio Documentazione, del 1980, si conta un solo addetto per effettuare oltre 700 ricerche annue per gli interni e oltre 300 per gli esterni.

Risale al 1954 la prima lista ufficiale delle intestazioni di soggetto dell'NLM e ancora non si parla di MeSH ma di *Subject Heading Authority List*. Infatti, la prima edizione dei *Medical Subject Headings* viene pubblicata nel 1960 e presenta 4.400 descrittori, tre anni dopo la seconda edizione già ne contiene 5.700.

L'edizione 2006 dei MeSH contiene 23.885 descrittori – che permettono di indicizzare, recuperare, organizzare e gestire oltre 14 milioni di citazioni di articoli scritti in 41 lingue e selezionati da oltre 4.800 periodici per la base dati MEDLINE/PubMed.

La traduzione italiana nasce nel quadro di un progetto internazionale UMLS (*Unified Medical Language System*) avviato dall'NLM, nel 1986. Questo progetto, che mira alla standardizzazione della terminologia medica, trova, nel 2000, una chiave di volta per facilitare le traduzioni dei MeSH in altre lingue. Infatti, viene modificata la struttura della base di dati, da una struttura basata su termini (*term-oriented*) si passa ad una struttura basata su concetti (*concept-oriented*): la base di dati costruita su descrittori e termini (*descriptors and terms*) varia in una che aggiunge i concetti e saranno proprio questi ultimi che consentiranno di individuare la migliore traduzione del singolo termine.

Le traduzioni dei MeSH, al settembre 2005, risultavano effettuate da 19 Paesi, anche se alcune risultano aggiornate annualmente e altre irregolarmente (1).

La traduzione italiana dei MeSH, oggetto di questo Convegno, inizia nel 1998 ed è stata completata nell'ottobre 2004, con risorse esclusivamente interne, viene aggiornata regolarmente a livello annuale e risulta disponibile, sul sito dell'ISS: <http://www.iss.it/site/Mesh/>.

Nel monitoraggio sull'uso della traduzione, nel periodo che va da novembre 2004 a novembre 2005, risultano visitate oltre 21.000 pagine. Per quanto riguarda l'affiliazione di coloro che si sono registrati abbiamo un certo numero di iscritti non identificabili per tale dato – questo perché l'informazione relativa all'affiliazione non era un campo obbligatorio nella scheda di registrazione. Tuttavia, ad un primo esame, tutte le istituzioni che possono beneficiare della traduzione sono presenti: Enti di ricerca, Università, Cliniche universitarie, strutture del Servizio Sanitario Nazionale, Ministeri, Organismi internazionali e privati.

La traduzione è disponibile, gratuitamente, previa registrazione, esclusivamente per una consultazione individuale, mentre sono previste tariffe diverse, sia in relazione alle diverse tipologie istituzionali relative, tuttavia, ad istituzioni non a fini di lucro, sia in relazione all'utilizzo legato a scopi interni o esterni. Nel caso di un utilizzo a scopo cosiddetto esterno, per esempio per distribuzione a terzi via Internet, occorre tenere presente che la traduzione italiana dei MeSH è coperta da copyright e, pertanto, i dati originali nel loro complesso non possono comunque essere resi disponibili ad altri utenti. Per ogni informazione concernente il copyright dei MeSH e della relativa traduzione italiana si può consultare il sito della NLM, progetto UMLS <http://www.nlm.nih.gov/research/umls/metaa1.html> sotto la voce MSHITA2006.

La traduzione è acquistabile, con procedure molto semplici, e il suo costo è da considerarsi in un certo senso irrisorio, rispetto ai vantaggi dell'utilizzazione: va da 180 euro, per tutte le istituzioni pubbliche e per scopi interni, a 450 euro per le stesse istituzioni pubbliche e scopi esterni, mentre l'aggiornamento annuale va, rispettivamente, da 45 a 110 euro e il costo di entrambe aumenta per istituzioni o enti privati. Informazioni più dettagliate sono reperibili al sito <http://www.iss.it/serv>.

Il costo che definisco irrisorio è tuttavia necessario per tentare di mantenere autosufficiente il sistema. Infatti, superato il problema della traduzione dei descrittori rimangono in piedi due importanti attività, legate alla dinamicità dei MeSH:

1. la prima attività riguarda l'aggiornamento annuale, in quanto un certo numero di termini viene introdotto annualmente dall'NLM, nel 2006 si calcolano circa 900 nuovi descrittori (1)
2. la seconda attività riguarda la revisione sistematica. Non mi riferisco, in questo caso, alla revisione annuale effettuata dall'NLM nella loro versione – che prevede un certo numero di cambiamenti e cancellazioni – bensì alla revisione sistematica della nostra traduzione italiana, in relazione alle modifiche nell'uso del nostro linguaggio scientifico. Inoltre, a partire dal 2006 si stanno introducendo anche i termini sinonimi.

In un'ottica di miglioramento del processo di revisione si è resa necessaria la costituzione di una rete di controllo semantico, ancora più ampia e, in tal senso, è stata avviata un'iniziativa di collaborazione con coloro che utilizzano la traduzione stessa.

Ritengo utile comunicare, inoltre, che lo sviluppo della traduzione italiana dei MeSH e l'importanza degli aspetti legati alla semantica ci ha permesso di collaborare al progetto REI: la Rete di Eccellenza dell'Italiano istituzionale, un'iniziativa del Dipartimento italiano della Direzione Generale della Traduzione della Commissione Europea a cui hanno aderito i cosiddetti "operatori istituzionali della lingua italiana" presenti all'interno di Dipartimenti,

Ministeri, Presidenza del Consiglio, ma anche Università e Enti di ricerca e, non ultima, l'Accademia della Crusca.

Strumenti di standardizzazione – quale il *Metathesaurus* del progetto UMLS di cui fa parte anche il *thesaurus* dei MeSH – sono stati prodotti per una migliore indicizzazione e quindi per un più efficace recupero e una migliore gestione dell'informazione, ma le traduzioni vanno considerate un ulteriore valore aggiunto a detti *thesauri*, poiché consentono l'interrogazione, nella propria lingua madre, di importanti basi di dati.

Un'ulteriore iniziativa dell'NLM, denominata BabelMeSH, vuole rendere ancor più integrato il sistema delle traduzioni esistenti creando, in un futuro, un'unica interfaccia di interrogazione che consenta immediatamente di effettuare la ricerca nella propria lingua madre ove esista, ovviamente, la traduzione.

Mi accingo a terminare questa introduzione, confermando che i competenti relatori di oggi illustreranno come la traduzione italiana dei MeSH possa essere proficuamente utilizzata non solo nella ricerca per l'interrogazione di basi di dati biomediche, ma anche a livello istituzionale e di ente per l'indicizzazione di risorse documentarie e di progetti, per la gestione di basi e banche di dati biomedici e di sanità pubblica e altro ancora.

Nel programma abbiamo voluto dare spazio anche ad interessanti comunicazioni libere, mentre nel pomeriggio avrà luogo una sessione dedicata alle ultime novità del PubMed e alla illustrazione del portale mondiale di tossicologia. Giovanni Arganese, Direttore della Biblioteca Medica, ha accettato di svolgere le funzioni di moderatore, e per questo ha la mia gratitudine.

Io, in qualità di Direttore del Settore Documentazione, ho ereditato importanti progetti – avviati sotto la direzione di chi mi ha preceduto: la dottoressa Adriana Dracos, che ci tengo a ringraziare, così come ringrazio il personale del Settore Documentazione che si è impegnato in questo importante lavoro e mi rivolgo a voi tutti per un sentito grazie per la vostra attenzione e partecipazione.

Bibliografia

1. Schulman J, Nelson S. *A model international cooperative project – the MeSH Translation Maintenance System*. Proceedings of the 9th World Congress on Health Information and Libraries. Salvador, Bahia, September 20-23, 2005.

IL LINGUAGGIO MEDICO UNIVERSALE E LA TRADUZIONE ITALIANA *DEI MEDICAL SUBJECT HEADINGS**

Maurella della Seta

Settore Documentazione – Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali

Istituto Superiore di Sanità, Roma

maurella.dellaseta@iss.it

Il progetto della traduzione italiana dei MeSH

La traduzione italiana del *thesaurus Medical Subject Headings* (MeSH) è stata realizzata nell'ambito di una collaborazione tra l'Istituto Superiore di Sanità (ISS) e la biblioteca statunitense *National Library of Medicine* (NLM), che cura la produzione e l'aggiornamento di questa terminologia biomedica, conosciuta a livello internazionale soprattutto come strumento per l'indicizzazione e la ricerca di articoli presenti nella banca dati PubMed (1). Quest'ultima, derivazione diretta di un repertorio a stampa, l'*Index Medicus*, pubblicato già a partire dal XIX secolo, considera tutti i settori della letteratura biomedica e indicizza gli articoli scientifici pubblicati in circa 4.800 periodici di oltre ottanta paesi; contiene attualmente oltre 16 milioni di riferimenti bibliografici dal 1950 ad oggi. In seguito ad un accordo stipulato nel 1978 tra il governo italiano e quello statunitense, l'ISS è centro di riferimento nazionale per il sistema MEDLARS (*MEDical Literature Analysis and Retrieval System*), di cui PubMed è l'elemento portante (2). In questa veste l'Istituto organizza corsi di formazione rivolti a professionisti dell'informazione, medici ed altro personale del Servizio sanitario nazionale, sulle caratteristiche e modalità di interrogazione delle basi dati prodotte dalla NLM. L'Istituto è inoltre centro di riferimento nazionale per il reperimento dei documenti non disponibili presso le biblioteche italiane attraverso il servizio DOCLINE della NLM (3). Nell'ambito dei progetti portati avanti congiuntamente alla NLM si è deciso nel 1997 di intraprendere l'oneroso lavoro di traduzione del *thesaurus* MeSH, analogamente a quanto fatto da altri Paesi, come la Germania, la Francia e la Spagna e di inserire la traduzione stessa all'interno di una base di dati multilingue.

Il *thesaurus* MeSH

Il *thesaurus* MeSH è uno strumento essenziale sia per l'indicizzazione delle riviste considerate in PubMed ed in altre basi di dati prodotte dalla NLM, sia per la ricerca su questi archivi, disponibili gratuitamente in Internet dal 1997 sul sito web della biblioteca stessa (4).

Il MeSH, consultabile liberamente tramite *MeSH Browser* (5) nella versione Internet, è composto da circa ventitremila descrittori principali (*main headings*), ottantatre sottodescrittori (*subheadings* o *qualifiers*) e oltre centocinquantamila voci supplementari (*Supplementary*

* Quest'articolo è stato pubblicato su *Biblioteche Oggi* 2006 24(9):37-42 (con il titolo *Un thesaurus bilingue per la biomedicina: la traduzione italiana dei Medical Subject Headings (MeSH)*) ed è qui riprodotto per gentile concessione dell'editore.

Concept Records), comprendenti nomi di sostanze chimiche, numeri di registro assegnati dal *Chemical Abstracts Service* e altre voci non accolte tra i descrittori principali. I descrittori MeSH sono riuniti per aree tematiche, ordinati gerarchicamente e suddivisi in sedici categorie (Figura 1), in modo che, all'interno di ciascuna categoria, ad un descrittore più generico segua uno più specifico, in una struttura definita ad albero. Uno stesso descrittore può far parte di diverse categorie secondo il punto di vista considerato (es. *Cysts* compare sia nella categoria *Neoplasms* che nella categoria *Pathological Conditions, Signs and Symptoms*) ed essere identificato quindi da più di un codice alfanumerico che indica la sua posizione all'interno della struttura ad albero.



Figura 1. Traduzione italiana delle categorie in cui sono raggruppati i MeSH

I descrittori MeSH sono corredati da relazioni semantiche (6); la griglia di connessioni concettuali consente, a partire da un dato descrittore e attraverso l'insieme dei termini ad esso collegati, di estendere o concentrare la ricerca di un documento. Si individuano a questo scopo tre fondamentali tipi di relazioni: la relazione di equivalenza, la relazione gerarchica e quella associativa.

La relazione di equivalenza rinvia da un termine di ricerca, non accolto, al termine adottato e regola l'uso dei sinonimi o quasi-sinonimi di un descrittore esistenti nel linguaggio naturale. I simboli che esprimono tale relazione sono: SEE = vedi e SEEN FROM = visto da (es. *Clinical Markers* SEE *Biological Markers*). Nell'interrogazione del PubMed la ricerca con un termine non accolto, sinonimo o quasi-sinonimo (*Entry Term*) sarà tradotta automaticamente nel descrittore MeSH preferenziale.

La relazione gerarchica collega ciascun termine MeSH ad uno di significato più generale o più specifico nell'ambito della medesima categoria. La struttura ad albero dei MeSH consente il meccanismo dell'esplosione, potente strumento di ricerca nel PubMed: inserendo un termine come *Anti-Bacterial Agents* il sistema ricercherà automaticamente anche gli articoli indicizzati con i descrittori *Antibiotics*, *Antitubercular*, *Antitreponemal Agents*, *Antitubercular Agents* e *Leprostatic Agents*, gerarchicamente inferiori (Figura 2).

La relazione associativa istituisce un richiamo tra due descrittori che appartengono alla stessa area semantica. Si tratta di una relazione reciproca che ha lo scopo di segnalare all'utente ulteriori chiavi di accesso all'informazione ricercata. Il simbolo che esprime tale relazione è

SEE ALSO = vedi anche (es. *Veins* SEE ALSO *Phlebography*; *Neoplasms* SEE ALSO *Antineoplastic Agents*). I nessi fra i due termini correlati possono essere di tipo organo-tecnica diagnostica, malattia-farmaco, come negli esempi citati, oppure di opposizione (*Fertility* SEE ALSO *Infertility*) o di coordinazione stretta (*Occupations* SEE ALSO *Workplace*).

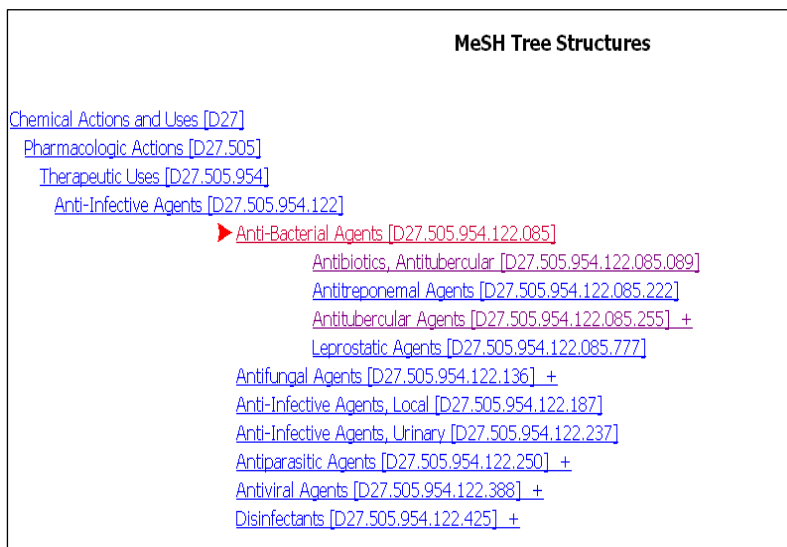


Figura 2. Struttura ad albero dei MeSH

Un altro tipo di relazione proprio del *thesaurus* MeSH è quella espressa dal simbolo CONSIDER ALSO TERMS AT che invita a considerare, nella ricerca di ulteriori chiavi di accesso all'informazione, le possibili radici latine o greche o anglosassoni dei descrittori: questa relazione è usata principalmente per i termini anatomici. Ad esempio *Liver* CONSIDER ALSO TERMS AT *hepat-*; *Kidney* CONSIDER ALSO TERMS AT *Glomerul-*, *Nephr-*, *Pyel-* and *Renal*.

Il *thesaurus* MeSH, attraverso il controllo della terminologia, rappresenta il punto d'incontro tra gli indicizzatori degli articoli (specialisti delle varie branche della medicina) e l'utente che interroga la base di dati PubMed, garantendo un alto grado di precisione nell'interrogazione dell'archivio.

I descrittori MeSH sono revisionati ed aggiornati annualmente sulla base dell'evoluzione delle scienze biomediche ed i nuovi termini sono esaminati e definiti all'interno del contesto già esistente da un'apposita commissione, che ne valuta l'eventuale inserimento nel *thesaurus*. Ogni anno è pubblicata una versione aggiornata dei MeSH: i nuovi descrittori per l'anno 2007 sono quasi cinquecento, mentre nel 2006 erano circa novecento. Generalmente, nell'annuale riesame del *thesaurus*, i revisori ne prendono in considerazione una o più aree semantiche e rivedono i termini che la compongono: negli ultimi due anni i maggiori cambiamenti sono stati effettuati all'interno della categoria D, Sostanze chimiche e Farmaci.

Il progetto *Unified Medical Language System*

La traduzione italiana dei MESH si inserisce all'interno di un progetto internazionale per la realizzazione di un linguaggio medico universale (*Unified Medical Language System - UMLS*) (7). L'UMLS, progetto intrapreso dalla NLM nel 1986, ha lo scopo di promuovere lo sviluppo di sistemi che aiutino i medici ed i ricercatori a reperire e ad integrare l'informazione in formato elettronico proveniente da fonti disparate e a superare le difficoltà derivanti dalle differenze terminologiche tra le varie basi di dati.

L'UMLS è articolato in tre sezioni:

1. Il *Metathesaurus*, oltre ad includere le traduzioni nelle varie lingue (italiano, francese, tedesco, spagnolo, portoghese, ceco e croato tra le circa ventisette che hanno iniziato o hanno completato la traduzione), riunisce ed integra l'informazione derivante da oltre cinquanta linguaggi controllati biomedici, impiegati nell'indicizzazione di basi di dati bibliografici e a testo completo, di cartelle cliniche, di basi di dati amministrativi in campo sanitario; attualmente comprende oltre due milioni di termini e dodici milioni di relazioni tra i termini stessi.
2. Lo *Specialist Lexicon*, contiene l'informazione lessicale con varianti morfologiche, relativa sia a molti termini del *Metathesaurus*, sia a parole e verbi che non vi sono inclusi;
3. L'UMLS *Semantic Network*, rete semantica contenente l'informazione relativa alle categorie cui tutti i termini facenti parte del *Metathesaurus* sono stati assegnati e le relazioni esistenti tra le categorie.

La traduzione italiana dei MeSH

Il lavoro di traduzione del *thesaurus* MeSH è iniziato nel 1997 (8), anno in cui i dati relativi ai descrittori sono stati prelevati dal sito della NLM ed inseriti all'interno di un database di tipo Microsoft Access, realizzato dal Settore Informatico dell'ISS.

Nel 1998 è iniziato il lavoro vero e proprio di traduzione, a cura di personale interno dell'ISS, a partire dalle categorie A (Anatomia) e C (Malattie). Dal 1999 al 2001 sono state tradotte tutte le altre categorie ed ogni anno sono stati inviati i dati alla NLM, per l'aggiornamento del *Metathesaurus*; dal 1999 ad oggi continua l'aggiornamento annuale dei dati (9). Dal 2000 la traduzione italiana è stata inserita nel *Metathesaurus* ed è disponibile sul già citato sito del progetto UMLS; per consultarla, unitamente alle altre traduzioni realizzate in varie lingue, è necessario registrarsi ed ottemperare alle norme sul copyright.

Dal mese di ottobre 2004 la traduzione italiana dei MeSH è consultabile gratuitamente, previa registrazione, sul sito dell'ISS (10), attraverso un'interfaccia web appositamente creata. Nel suo interno è disponibile una pagina di aiuto che fornisce alcuni suggerimenti per ottimizzare l'uso del database. È possibile adoperare l'operatore di troncamento "*" sia a destra che a sinistra, per effettuare ricerche su stringhe di testo; la ricerca può essere effettuata partendo sia da un termine italiano che da un termine inglese. In alternativa si può esplorare la struttura gerarchica ad albero tradotta in italiano. Aprendo il collegamento a partire dal termine in inglese, si accede alla pagina relativa all'interno del *MESH browser*, cioè del dizionario in rete della NLM, mentre aprendo il collegamento da un termine italiano si ottiene il codice alfanumerico corrispondente alla posizione nell'albero gerarchico (Figura 3).

I limiti di questa realizzazione sono principalmente dovuti alla mancata traduzione dei sinonimi (*Entry Terms*); tale carenza fa sì che, a differenza di quanto accade nel MeSH Browser o nel PubMed, non sia consentito il rinvio automatico da un termine non accolto al termine preferenziale: è opportuno dunque utilizzare sempre l'operatore di troncamento "*" sia all'inizio che alla fine della stringa di ricerca, per consentire un maggiore reperimento di termini.

| Medical Subject Headings - 2006 | | |
|----------------------------------|--|--|
| Traduzione Italiana | | |
| La ricerca ha reperito 89 record | | |
| Codice Mesh | NLM MeSH Browser | Albero in italiano |
| D018255 | Adenomatosis, Pulmonary | Adenomatosi polmonare |
| D011648 | Pulmonary Adenomatosis, Ovine | Adenomatosi polmonare ovina |
| D011650 | Pulmonary Alveoli | Alveoli polmonari |
| D011651 | Pulmonary Artery | Arteria polmonare |
| D008169 | Lung Abscess | Ascesso polmonare |
| D001229 | Aspergillosis, Allergic Bronchopulmonary | Aspergillosi allergica broncopolmonare |
| 1 2 3 4 5 6 7 8 9 10 ... | | |

Figura 3. Risultato della ricerca per la stringa *polmon*

Dal 2004 per il lavoro di traduzione è utilizzato un software sviluppato dalla NLM, il MeSH Translation Maintenance System (MTMS) (11). I termini tradotti sono memorizzati direttamente sul *server* della NLM, che, a richiesta, invia periodicamente i dati aggiornati ai vari centri di traduzione per le applicazioni locali. La traduzione dei sinonimi (termini non preferenziali) è espressamente prevista nel sistema MTMS, e anche l'ISS ha iniziato ad effettuarla, almeno per quanto riguarda l'aggiornamento annuale dei dati e, a poco a poco, anche per il retrospettivo.

Problematiche terminologiche emerse nel corso della traduzione

L'analisi dei MeSH condotta a livello linguistico rivela come il *thesaurus* non costituisca nella sua versione originale un *corpus* omogeneo. Si può, pertanto, formulare l'ipotesi che l'opera, come del resto la sua traduzione italiana, sia il frutto del lavoro di più esperti, portato avanti nel corso degli anni. Infatti, considerando termini che richiederebbero una coerenza espressiva, come ad esempio quelli appartenenti alla Categoria A *Anatomia*, si nota che i descrittori sono stati introdotti indifferentemente nella forma singolare o plurale (*Eye/Eyelids* – *Breast/Mammary Glands*, *Human/Nipples*, rispettivamente per *Occhio/Palpebre* – *Mammella*, *Ghiandole Mammarie Umane/Capezzoli*). Nella Categoria C *Malattie*, alcune patologie sono espresse con terminologia derivata dal greco o dal latino (*Avian Leukosis* per *Leucosi aviaria*), altre nella loro forma volgare anglo-americana (*Influenza in Birds* per *Influenza aviaria*).

Il riflesso di questa disomogeneità traspare anche dagli aggiornamenti e revisioni annuali della NLM, che traggono spunto dall'adeguamento a terminologia più avanzata e corrente per ristrutturare ed uniformare dal punto di vista lessicale intere aree semantiche. Ad esempio, *Feet* è stato recentemente sostituito dal termine preferenziale *Foot (Piede)*, per uniformare al singolare le parti del corpo. Il lavoro di traduzione è stato svolto al massimo della fedeltà, traducendo i descrittori secondo l'impostazione lessicale originale. Solo laddove quest'ultima non avesse riscontro nella lingua italiana, si è optato per una diversa formulazione. Ad esempio, l'introduzione nei MeSH 2006 del descrittore *Hepatic Insufficiency* gerarchicamente superiore a *Liver Failure*, dove in italiano non esiste una distinzione netta concettuale, ha portato ad una scelta differente secondo l'uso, e cioè rispettivamente *Riduzione della funzionalità epatica* e *Insufficienza epatica*.

L'introduzione dei sinonimi anche in lingua italiana potrà andare ben oltre la risoluzione di questi problemi. Indipendentemente dalla scelta originale, l'italiano designerà come descrittore il termine più tecnico, quello, cioè, che l'esperto del settore andrebbe a ricercare. Seguirà, ove disponibile, il termine comune e di utilizzo finalizzato a possibili categorie mirate di utenti (bibliotecari, indicizzatori, utenza non specializzata). Così, dall'impostazione originale anglo-americana, si delinearanno dei veri e propri lessici paralleli. Il conseguente sviluppo del software di ricerca consentirà di risalire al termine originale in lingua inglese, collegando tutti gli altri sinonimi.

Da non sottovalutare, infine, il particolare interesse che potrebbe risvegliare l'accoglienza di terminologia anglo-americana scientifica nel lessico italiano, con l'introduzione di neologismi lessicali, come ad esempio il difficilmente traducibile *Dna Fingerprinting*.

Alcune applicazioni della traduzione italiana dei MeSH

La traduzione italiana dei MeSH, così come il suo inserimento all'interno del *Metathesaurus*, è stata motivata, oltre che dalla volontà di partecipare ad un progetto internazionale di ampio respiro e di sicuro interesse sia per la comunità medica che per quella documentaria, dalla necessità di fornire all'utenza italiana uno strumento di standardizzazione per la catalogazione e l'indicizzazione della documentazione biomedica, comprendente, tra l'altro, quella indirizzata ai cittadini nell'ambito di progetti di informazione ed educazione sanitaria.

Altri paesi, prima di noi, infatti, hanno realizzato da tempo repertori di risorse Internet in campo medico, utilizzando per l'indicizzazione delle stesse i descrittori MeSH o la loro traduzione. Si ricordano, ad esempio, il progetto *Intute: Health and Life Sciences* (già *Biome*) <<http://www.intute.ac.uk/healthandlifesciences/biomelost.html>>, realizzato nel Regno Unito, e il suo analogo francese *Cismef - Catalogage et Indexation des Sites Médicaux Francophones* <<http://www.chu-rouen.fr/cismef/>>.

In Italia il progetto AZALEA, biblioteca digitale in oncologia per malati, familiari e cittadini, è stato il primo utente esterno all'ISS a richiedere l'autorizzazione ad utilizzare la traduzione italiana dei MeSH per l'indicizzazione delle risorse elettroniche e cartacee presenti nel proprio catalogo (12).

Varie sono state le applicazioni informatiche per cui la traduzione italiana dei MeSH ha trovato utilizzazione all'interno dell'ISS. La Biblioteca adopera la terminologia MeSH tradotta in italiano per la soggettazione del materiale documentario presente nel proprio catalogo (13). La relazione annuale di rendiconto delle attività istituzionali si serve della traduzione dei MeSH, per l'indicizzazione di pubblicazioni, progetti di ricerca, attività di controllo. Analogamente, il Settore per le Attività Editoriali dell'ISS indicizza con i MeSH in italiano ed inglese le pubblicazioni scientifiche dei ricercatori dell'Istituto, raccolte in una base di dati bibliografici

interna. Il Settore Documentazione ha utilizzato un sottoinsieme della traduzione come punto di partenza per sviluppare il Tesoro italiano di bioetica (14, 15). La terminologia MeSH, opportunamente adattata, è servita come base per preparare lo schema di classificazione usato per indicizzare i periodici del PubMed analizzati all'interno del sito PIRAMIDE (16).

Uno degli scopi fondamentali della traduzione in italiano e nelle altre lingue, forse il più interessante per l'utenza costituita dal personale medico, resta quello di consentire l'interrogazione del PubMed nella lingua madre degli utenti, tramite la funzione di *mapping* (rinvio) ai MeSH in lingua inglese, gestita dal *Metathesaurus*. È in questa direzione che si sta muovendo la NLM attraverso la realizzazione di metamotori e di interfacce multilingue, per ora realizzati in versione sperimentale (17).

Bibliografia e note

1. <http://pubmed.gov>
2. Per informazioni sui centri MEDLARS si veda U.S. National Library of Medicine. *International Medlars Centers Fact Sheet* Disponibile all'indirizzo: <http://www.nlm.nih.gov/pubs/factsheets/intlmedlars.html>; ultima consultazione 22/12/06
3. Le attività del Settore Documentazione dell'ISS sono descritte con maggior dettaglio nel sito <http://www.iss.it/sdoc>
4. www.nlm.nih.gov
5. <http://www.nlm.nih.gov/mesh/MBrowser.html> . Si veda anche, per maggiori informazioni sui MeSH, U.S. National Library of Medicine. *Fact Sheet: Medical Subject Headings (MeSH®)*. Disponibile all'indirizzo: <http://www.nlm.nih.gov/pubs/factsheets/mesh.html>; ultima consultazione 22/12/06.
6. Nelson SJ, Johnston WD, Humphreys BL. *Relationships in Medical Subject Headings*. In: Bean, Carol A, Green R (Ed.). *Relationships in the organization of knowledge*. New York, Kluwer Academic Publishers, 2001, p. 171-184.
7. U.S National Library of Medicine. *Unified Medical Language System*. Disponibile all'indirizzo: <http://www.nlm.nih.gov/research/umls/umlsmain.html>; ultima consultazione 22/12/06
8. Della Seta M, Dracos A. La traduzione italiana del Thesaurus Medical Subject Headings (MeSH). In: *Seminario di studio "Reti per la salute"*. Atti, Perugia, SEDES, 1999, p. 57-62.
9. Ceccarini A, Della Seta M. The Italian translation of NLM MeSH: a collaboration between NLM and ISS. *Newsletter to European health librarians* 2004;68:40-4. Disponibile all'indirizzo: <http://www.eahil.net/newsletter/68/68.pdf> ; ultima consultazione 22/12/06.
10. <http://www.iss.it/site/mesh/>
11. Schulman JL, Nelson SJ. *A model international cooperative project. The MeSH translation maintenance system*. ICML 9, 9th World Congress on Health Information and Libraries, Salvador-Bahia, Brazil. Disponibile all'indirizzo: <http://www.icml9.org/program/track1/activity.php?lang=en&id=17>; ultima consultazione 22/12/06
12. <http://www.azaleaweb.it/>. Il sito alla data di stesura di questo articolo (ottobre 2006) risulta non più attivo.
13. Falcone MA, Ferrari P. Le nuove scelte di catalogazione semantica della biblioteca dell'Istituto superiore di sanità. *Bollettino AIB* 2002;42(3):333-6.
14. Gruppo di studio Istituto Superiore di Sanità "Tesoro italiano di bioetica". *Tesoro italiano di bioetica*. Roma: Istituto Superiore di Sanità; 2006. (Strumenti di Riferimento 06/S3). Disponibile all'indirizzo: <http://www.iss.it/binary/sibi/cont/thesa.1151313180.pdf>; ultima consultazione 22/12/06.

15. Della Seta M, Napolitani F, Pizzarelli S, Poltronieri E. Thesaurus Italiano di Bioetica: un esempio di lessico specializzato nel settore biomedico. In: Corino E, Marellò C, Onesti C (Ed.). *12. Congresso internazionale di lessicografia (EURALEX). Atti; 6-9 settembre 2006; Torino*. Alessandria: Edizioni dell'Orso; 2006. p. 825-30.
16. www.iss.it/site/piramide
17. <http://babelmesh.nlm.nih.gov/>

Dizionari consultati

- Delfino G, et al. *Il nuovo medicina e biologia. Medicine & Biology. Dizionario enciclopedico di scienze mediche e biologiche e di biotecnologie italiano-inglese inglese-italiano*. 1ª ed. 1996. Bologna: Zanichelli; 2003.
- De Mauro T. *Grande dizionario italiano dell'uso*. Torino, UTET, 2003.
- Il McGraw-Hill-Zanichelli Dizionario enciclopedico scientifico e tecnico inglese-italiano italiano-inglese* (1 ed. italiana 1980). Bologna: Zanichelli; 2004.
- Galli MC, Cappelli AM (Ed.). *Farmacopea europea: guida alla traduzione e glossario di termini utilizzati nelle monografie dei prodotti biologici e relativi metodi*. Roma: Istituto Superiore di Sanità; 1999.
- Mafatka Z. *Terminologia, definizioni e criteri diagnostici in endoscopia digestiva*. Bad Homburg: NORMED-Verl; 1986.

SOGGETTARIO DELLA BIBLIOTECA DELL'ISTITUTO SUPERIORE DI SANITÀ: ESPERIENZE D'EFFICACIA E PRECISIONE SCIENTIFICA NELL'IMPIEGO DELLE VOCI MESH

Maria Alessandra Falcone, Paola Ferrari

Biblioteca – Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali

Istituto Superiore di Sanità, Roma

alessandra.falcone@iss.it; paola.ferrari@iss.it

Il progetto di tradurre in lingua italiana le voci del *Medical Subject Headings* (MeSH), *thesaurus* gestito e prodotto dalla *National Library of Medicine* (NLM) di Bethesda per indicizzare principalmente gli articoli di oltre 4.800 riviste biomediche, è stato intrapreso con successo dal Settore Documentazione e dal Settore Informatico dell'Istituto Superiore di Sanità (ISS). La suddetta traduzione ha costituito, a partire dal 2001, un valido e autorevole supporto all'avvio operativo di un programma generale di revisione strutturale e semantica dell'intero soggetto della Biblioteca ISS

La Biblioteca è specializzata nella documentazione scientifica in ambito biomedico-sanitario con più di 7.000 titoli di periodici chiusi e più di 2.000 correnti, circa 200.000 libri e una ricca raccolta di documenti su CD-ROM e online.

Nel corso degli ultimi anni, si era più volte presentata la necessità di riflettere sui criteri strutturali e metodologici applicati alla scelta e costruzione delle relazioni sintattiche tra i termini presenti nell'archivio automatizzato della nostra biblioteca.

Le maggiori difficoltà, connesse in passato all'utilizzo delle voci MeSH, erano infatti, spesso determinate dalla non formalizzazione in lingua italiana dei termini inglesi tecnico-scientifici, a volte di non agile resa, ostacolo superato dalla recente traduzione della quale oggi ci serviamo per la costruzione del nostro soggetto espresso, da sempre, per tradizione nella nostra madrelingua.

Inoltre, l'evoluzione terminologica e disciplinare in campo scientifico, l'affermarsi di nuovi strumenti d'interrogazione online e la reale esigenza dell'utenza di effettuare una ricerca il più possibile tempestiva e inerente alla specificità delle proprie richieste, hanno contribuito a motivare nuove scelte di soggettazione.

Il nostro intervento si propone di presentare l'utilità e la precisione scientifica acquisita dal soggetto della Biblioteca ISS, con l'impiego delle voci MeSH e della relativa traduzione italiana, attraverso i seguenti obiettivi:

- fornire strumenti idonei per un'adeguata interrogazione semantica;
- evidenziare la validità scientifica del soggetto della Biblioteca ISS;
- sottolineare l'alta specializzazione della terminologia MeSH in grado di consentire la selezione di soggetti pertinenti, rendendo più efficace ed efficiente la ricerca semantica e il recupero dell'informazione;
- effettuare ricerche mirate e complete con il minimo "rumore".

La completa realizzazione di tali obiettivi passa attraverso un complesso di principi catalografici quali l'interrogazione semantica e la soggettazione:

- l'interrogazione semantica ha, essenzialmente, la funzione di analizzare il contenuto concettuale di ogni singola unità bibliografica;

- la soggettazione rappresenta una tra le funzioni più significative di indicizzazione, volta ad individuare gli elementi concettuali contenuti nei documenti descritti, formalizzandoli in una espressione normalizzata, così che un soggetto sia sempre indicato allo stesso modo nel catalogo.

Per effettuare, quindi, una corretta procedura di ricerca semantica bisognerebbe prima di tutto individuare l'oggetto della propria ricerca, selezionare e circoscrivere i temi base, tradurre i concetti in termini normalizzati, consultare il soggettario estendendo successivamente l'indagine ai sinonimi e ai termini affini.

Il soggettario, strumento indispensabile nelle fasi di catalogazione e di interrogazione semantica, contraddistinto da un elenco alfabetico delle voci di soggetto e da una rete di collegamenti semantici tra termini, costituisce un valido strumento di normalizzazione specifico di una singola biblioteca o condiviso con altre.

I termini (soggetti) che costituiscono qualunque soggettario rappresentano:

- l'oggetto di studio dell'autore;
- un tema all'interno del documento stesso;
- l'argomento di cui tratta un documento;
- l'espressione linguistica usata per descrivere il contenuto di un documento;
- l'elemento descrittivo e ricognitivo del testo;
- l'enunciato espresso da un termine (descrittore) o da una combinazione di più termini (stringa).

Il soggetto è, quindi, lo strumento di indicizzazione primario che consiste nell'analisi concettuale di un documento, il cui contenuto verrà formulato sinteticamente con termini normalizzati, che, una volta ordinati e strutturati, costituiranno il soggettario.

I principi che regolano la soggettazione, validamente rappresentati dalla terminologia MeSH, sono volti a garantire la specificità, l'univocità, l'uniformità e la pertinenza dei termini scelti.

Tra le tecniche di rappresentazione del contenuto concettuale dei documenti si distinguono due importanti sistemi:

1. *Sistema pre-coordinato*

in cui le singole voci del soggettario vengono combinate tra loro nella fase di indicizzazione dal soggettore che, ordinando tali termini secondo precise regole sintattiche, formula una stringa di soggetto che verrà utilizzata dall'utente al momento della ricerca (es. Tabacco – Fumo – Combustione – Residui – Tossicità – Prevenzione e controllo – Italia);

2. *Sistema post-coordinato*

in cui i termini del soggettario o descrittori, esistenti come singole unità ordinate gerarchicamente, vengono concettualmente combinati tra loro dall'utente solo al momento della ricerca; esempio: Acqua potabile, Controllo di qualità, Microbiologia dell'acqua, Manuali di laboratorio.

Inizialmente le nostre attenzioni si sono concentrate su una politica di radicale cambiamento di impostazione strutturale del soggettario della Biblioteca dell'ISS, indirizzando le nostre scelte verso l'utilizzo di un sistema di indicizzazione pre-coordinato, qual era quello proposto dal Gruppo di ricerca sull'indicizzazione per soggetto (GRIS), rappresentando questo, un valido strumento di indicizzazione in campo semantico.

Tale metodo prevede tre fasi distinte:

1. analisi del documento;
2. costruzione della stringa di soggetto;
3. controllo terminologico.

L'aspetto innovativo di tale procedura è da ricercarsi soprattutto nella costruzione della stringa di soggetto. Questo processo prevede, infatti, una serie di operazioni analitiche per la

costruzione di una stringa definita come sequenza logica di termini che esprimono il soggetto del documento in modo univoco ed esaustivo attraverso un ordine di citazione regolato da principi logico-sintattici (concetto chiave, azione, agente ecc.).

Dopo più di un anno di sperimentazione e applicazione pratica del metodo GRIS, sono emerse carenze e problematiche che hanno imposto un ulteriore approfondimento e una soluzione che, nel tempo, avesse potuto soddisfare le peculiari esigenze della nostra utenza.

La scelta si è quindi indirizzata verso un sistema d'indicizzazione post-coordinato, in quanto più idoneo, tra l'altro, alla tipologia di utenza e alle raccolte librerie scientifiche che caratterizzano la Biblioteca ISS.

Le difficoltà più evidenti da superare erano determinate dall'analisi del materiale bibliografico posseduto, costituito da testi per lo più a carattere pluridisciplinare e marcatamente specialistico.

Per contro, l'uso di stringhe pre-coordinate regolate da relazioni logiche tra termini, specie per questo tipo di pubblicazioni scientifiche, non si presentava tale da eludere le difficoltà oggettive di esaustività descrittiva, dando spesso luogo, nella fase di consultazione del catalogo, ad ambiguità semantiche dell'enunciato di soggetto. Risultava oggettivamente complesso racchiudere in un'unica stringa e in modo esaustivo, il contenuto di ciascun documento.

La NLM, consentendo la consultazione in rete del proprio catalogo e offrendo una descrizione bibliografica completa dei documenti posseduti, costituisce oggi per noi, una valida traccia per l'ordine citazionale dei termini scelti che descrivono in modo esaustivo il contenuto concettuale del documento da indicizzare.

Un altro fattore, di non secondaria importanza, che ha influito sul cambiamento di scelta, riguarda l'interrogazione con gli operatori logici booleani, effettuabile tramite sistemi automatizzati e utilizzata per lo più da un'utenza esperta prevalentemente di biomedicina. Con un sistema post-coordinato è possibile attuare una ricerca più precisa e mirata, essendo l'utente parte attiva nell'individuazione dei descrittori più idonei, mentre la visualizzazione di stringhe, precedentemente formulate dal catalogatore e non sempre così inerenti alla specificità dei contenuti, ne disorientava e limitava, a volte, la ricerca.

La scelta di un'indicizzazione di tipo post-coordinato è stata inoltre motivata dalla possibilità di poter usufruire online della traduzione italiana del *thesaurus* MeSH, vocabolario di termini controllati utilizzati per la rappresentazione sintetica del contenuto concettuale di documenti biomedici. Composto da una lista strutturata di termini corrispondenti ad un concetto o ad una classe di concetti, il MeSH viene utilizzato dall'utenza come strumento fondamentale per il reperimento dell'informazione biomedica. Esso costituisce una fonte autorevole di termini di carattere tecnico-scientifico selezionati da specialisti delle singole discipline.

La revisione, il continuo aggiornamento, l'introduzione di nuove voci con note di ambito complete, l'eliminazione o modifica di termini, nonché l'aggiunta di rinvii a cui è sottoposto il MeSH, da parte di un'apposita commissione di specialisti della materia, ne garantiscono l'attualità, la validità scientifica e la verifica qualitativa della terminologia.

È stato quindi avviato un lavoro parallelo di puntuale revisione e aggiornamento anche del nostro soggettario che, altrimenti, avrebbe inevitabilmente diminuito il suo grado di precisione (Figura 1). Il nostro intervento, tuttora operativo, ha dunque interessato l'analisi delle relazioni tra i soggetti attraverso una revisione costante e completa di tutte quelle voci che, nel vecchio catalogo e nelle precedenti stringhe, apparivano spesso come singoli termini o come suddivisioni, mentre con la nuova soggettazione su base MeSH, i medesimi concetti vengono espressi con un unico termine. Esempio:

- Incidenti / prevenzione e controllo vedi Antinfortunistica (*Accident Prevention*)
- Acqua/Contaminazione radioattiva vedi Contaminazione radioattiva dell'acqua (*Water Pollution, Radioactive*)

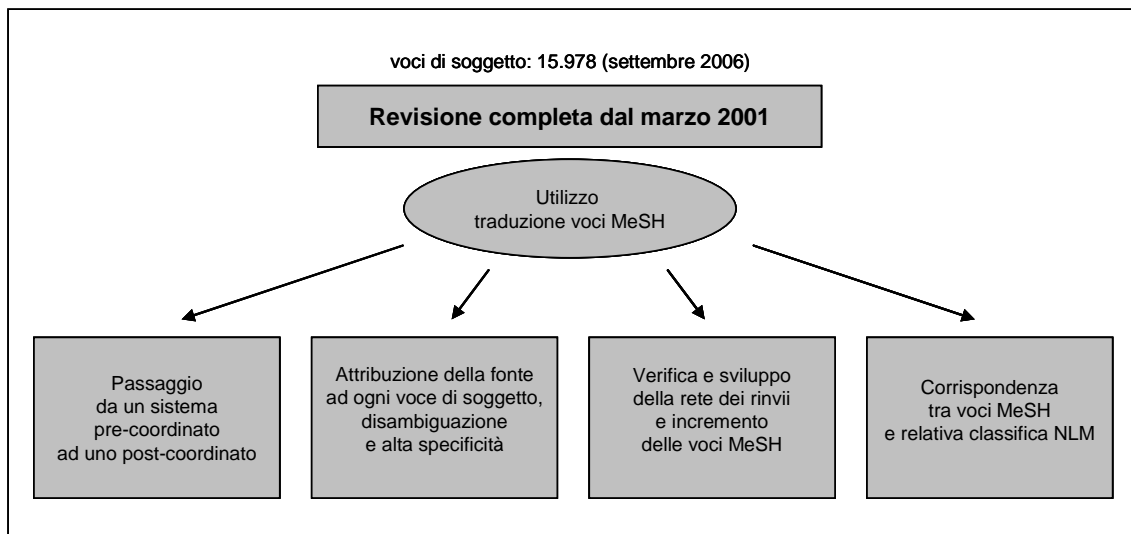


Figura 1. Principali fasi della revisione del soggettario della Biblioteca ISS

Il *thesaurus* MeSH è stato, dunque, lo strumento più completo ed efficiente che ha permesso la riorganizzazione del nostro soggettario, vista la ricchezza terminologica che lo caratterizza:

- circa 23.000 descrittori principali (*main heading*);
- quasi 800 sottodescrittori (*subheading o qualifier*);
- più di 100.000 voci supplementari (*supplementary concept record*).

Il MeSH costituisce oltretutto un sicuro ausilio per l'utente in quanto riporta una serie di termini non-MeSH, cioè di descrittori che, pur essendo parti essenziali della struttura ad albero, poiché passaggi di divisione necessari, è improbabile che siano utilizzati come termini di indicizzazione o di ricerca. Questi termini quindi rinvieranno a termini semanticamente correlati consentendo una ricerca a più ampio spettro.

Il MeSH non si serve dei codici di collegamento convenzionali quali BT, NT, RT, USE FOR ecc., bensì di un insieme di segni espliciti realizzati specificatamente per il sistema MEDLARS (insieme di fonti di informazione in formato elettronico in ambito biomedico prodotto dalla NLM di Bethesda).

Nel nostro soggettario sono stati quindi introdotti legami del MeSH, conformemente al sistema automatizzato Sebina, attualmente in uso presso la Biblioteca ISS, che gestisce il materiale bibliografico:

- *vedi (see)* che rimanda da termini non preferiti a termini scelti (CANCER *see* NEOPLASMS);
- *vedi anche (see related)* che indica la presenza di altri descrittori correlati concettualmente alla voce in esame (BLOOD PRESSURE *see related* HYPERTENSION / HYPERTENSION XR BLOOD PRESSURE).

La presentazione dei termini che indicizzano ciascuna unità bibliografica, segue inoltre, come per il MeSH, un ordine gerarchicamente strutturato in base al grado di rilevanza rispetto al contenuto concettuale del documento.

La scheda di descrizione bibliografica proposta dalla NLM e compilata utilizzando la terminologia MeSH, rappresenta un valido supporto alla catalogazione semantica, sia per quanto riguarda l'ordine citazionale gerarchico che per quello formale delle voci di soggetto. L'inserimento delle voci MeSH ha inevitabilmente comportato un conseguente aumento numerico

di termini, specialmente di quelli relativi al settore della patologia medica, della biologia cellulare e dei composti chimici, che ha reso necessario un successivo e attento lavoro di bonifica dei documenti precedentemente correlati a termini più generici. Questo lavoro di revisione terminologica dell'archivio dei soggetti, è stato ovviamente più consistente nei primi anni successivi al 2001 per colmare una mancanza di uniformità dei termini stessi. Ogni anno si introducono nuovi termini MeSH nel soggettario, i cui dati sono i seguenti:

- nel 2003 3793;
- nel 2004 2476;
- nel 2005 572;
- nel 2006 165.

La politica di indicizzazione del nostro soggettario, ha riguardato essenzialmente il rispetto del criterio di specificità nella selezione del descrittore MeSH più idoneo, la disambiguazione, l'assegnazione di tutti quei termini necessari a descriverne in modo esaustivo il contenuto, nonché la relazione tra due o più descrittori.

È inoltre da osservare come si sia venuta a semplificare la metodologia di ricerca, a seguito di un'accertata maggiore consuetudine e familiarità della nostra utenza ad utilizzare i descrittori MeSH per l'interrogazione di basi di dati quali ad esempio PubMed che, tra le basi di dati bibliografiche, rappresenta il sistema più diffuso a livello biomedico.

Anche in passato, nella nostra biblioteca, si era fatto uso del MeSH, ma esso non costituiva il principale *thesaurus* di riferimento, essendo molteplici le fonti alle quali il soggettatore attingeva per l'individuazione dei descrittori e delle suddivisioni, con un'inevitabile disomogeneità dei termini thesaurali presenti nell'archivio.

Per garantire un recupero esaustivo dell'informazione, particolare attenzione, è stata da noi dedicata alla traduzione in lingua italiana e al successivo inserimento nell'archivio dei soggetti, delle voci di rinvio e dei sinonimi MeSH.

L'utilizzo, in fase di ricerca, dei rinvii ai termini esclusi offre certamente maggiori garanzie di esattezza e completezza del soggettario, sebbene venga a volte da noi operata una scelta terminologica rispetto al MeSH, per non incorrere, nel caso ad esempio dei sinonimi, in inutili ripetizioni determinate dalla non sovrapponibilità dei termini inglesi con quelli italiani corrispondenti, che appesantirebbero inutilmente l'archivio (Figure 2 e 3).

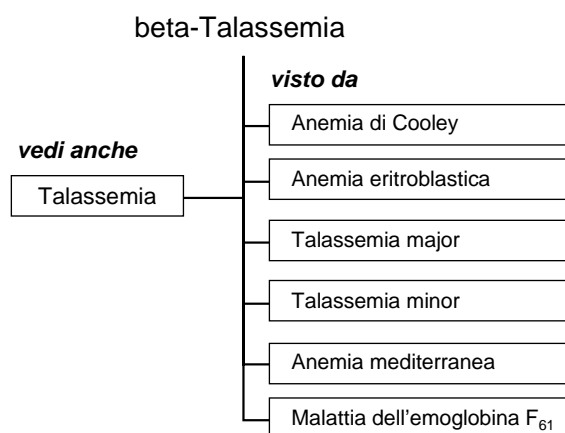


Figura 2. Esempio di voci MeSH con relative voci di rinvio presenti nel soggettario Biblioteca ISS

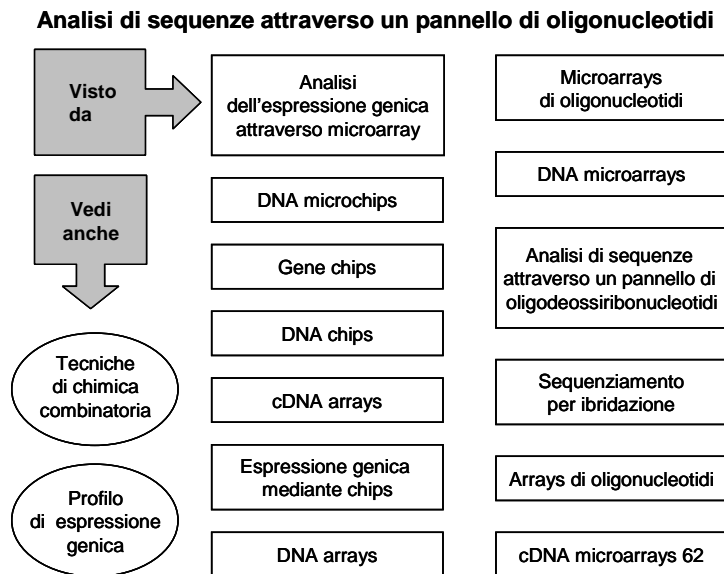


Figura 3. Esempio di voci MeSH con relative voci di rinvio presenti nel soggettario Biblioteca ISS

L'utilizzo del *thesaurus* MeSH garantisce ulteriore conformità ai soggetti prescelti per l'indicizzazione di un determinato documento e la classificazione NLM corrispondente.

In passato, invece, la scelta di termini non sempre appartenenti al *Medical Subject Headings* comportava discordanza con la classe assegnata. Il sistema di classifica, in uso presso la nostra biblioteca, è attualmente il più possibile conforme a quello della NLM per quanto concerne importanti aree disciplinari della biomedicina (Figura 4).

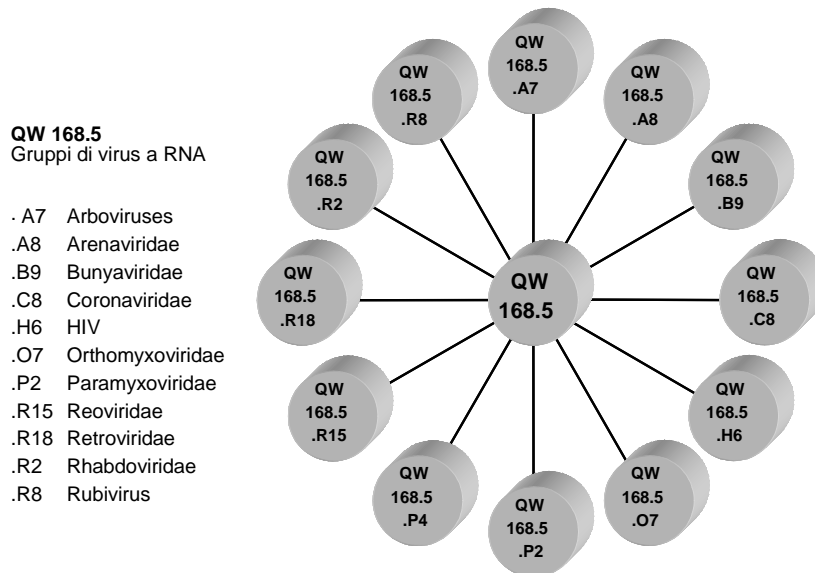


Figura 4. Specificità e corrispondenza tra classe e soggetto MeSH

Per argomenti riguardanti, ad esempio, la psicologia, la sociologia, la geografia, la tecnologia, la medicina veterinaria, la fisiologia animale nonché la biblioteconomia, si è resa necessaria l'integrazione con il sistema di classificazione della *Library of Congress*, che viene utilizzato come supporto sostanziale.

Il lavoro di aggiornamento ha riguardato anche una revisione sistematica e complessiva di tutti i soggetti presenti nell'archivio, con inserimento in nota della fonte, che molto spesso risulta essere il MeSH.

In alcuni casi e per documenti di ambito disciplinare non pertinente alla biomedicina, la fonte è stata evinta da ulteriori e autorevoli soggetti come ad esempio quello della *Library of Congress* o, nel caso di documenti italiani, prevalentemente dal soggetto della Biblioteca Nazionale Centrale di Firenze, utilizzato principalmente dal Servizio Bibliotecario Nazionale. L'indicazione della fonte, riportata per ciascun soggetto nella nota catalografica, è stata da noi introdotta in quanto conferisce al nostro soggetto maggior esattezza e validità scientifica, oltre ad offrire la possibilità di poter verificare la corrispondenza tra una voce e il *thesaurus* da cui è stata desunta (Figura 5).

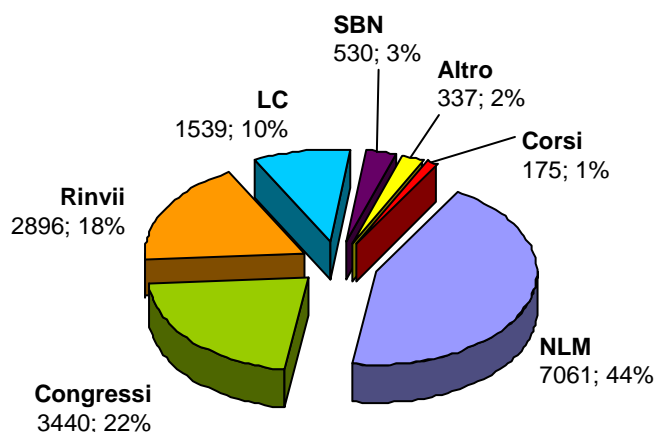


Figura 5. Numero delle voci di soggetto suddivise per fonte (Biblioteca ISS)

Nell'archivio dei soggetti del catalogo della nostra biblioteca, sono anche presenti in numero considerevole voci di congressi e corsi con relativa indicazione del luogo e della data del loro svolgimento. Queste voci offrono all'utente, un ulteriore recupero dell'informazione scientifica in gran parte presente e aggiornata proprio negli atti dei convegni.

A conclusione di questo nostro intervento, ci sembra opportuno sottolineare come frequenti studi di valutazione e una costante applicazione a livello operativo della traduzione italiana dei MeSH, conferiscano al soggetto della Biblioteca ISS, un'elevata specializzazione e una reale validità scientifica costituendo un efficace strumento di ricerca e recupero dell'informazione. Il MeSH rappresenta dunque un efficiente *thesaurus*, al quale il nostro soggetto sempre più si sta allineando con l'intento di poter costituire, in un prossimo futuro, un valido esempio di scelte terminologiche-descrittive pertinenti, utili all'indicizzazione del materiale documentario di altre biblioteche scientifiche.

IMPLEMENTAZIONI E SVILUPPI DELL'APPLICATIVO WEB *MEDICAL SUBJECT HEADINGS* – TRADUZIONE ITALIANA

Maurizio Ferri, Paolo Roazzi

*Settore Informatico, Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali
Istituto Superiore di Sanità, Roma
maurizio.ferri@iss.it, paolo.roazzi@iss.it*

L'applicativo web *Medical Subject Headings* – Traduzione italiana, consultabile dalla sezione del sito ISS dedicata al Settore Documentazione, è stato realizzato allo scopo di offrire, all'utenza interessata all'argomento, la possibilità di consultare in tempo reale la traduzione in lingua italiana dei termini MeSH contenuti nella banca dati della *National Library of Medicine* (NLM) di Bethesda (Stati Uniti).

Dal momento che il sito della NLM è indubbiamente di grande interesse per utenti di tipo medico e scientifico in tutto il mondo, si è pensato che nella realtà italiana sarebbe stato altrettanto utile mettere a disposizione queste informazioni tradotte da personale specializzato dell'ISS. La costruzione di questa web *Application* è risultata abbastanza lineare; le informazioni che transitano sulla rete ad ogni interrogazione sono piuttosto snelle e quindi anche i tempi di attesa non appaiono lunghi; gli oggetti utilizzati sono appropriati anche se ne esistono già di nuovi e quindi questo aspetto è migliorabile.

Le uniche difficoltà incontrate sono legate fundamentalmente alla implementazione di strutture dette ad albero, di cui parleremo nel paragrafo dedicato alle specifiche tecniche, che hanno fatto lievitare il costo in termini di lavoro. L'obiettivo che è stato sinora raggiunto è comunque da ritenersi il passo iniziale verso una finalità più ampia e in evoluzione.

Non esiste attualmente un disegno definitivo, ma l'intento è quello di rimanere aperti a suggerimenti e idee progressive per cercare di ottenere uno strumento che potrà essere modificato al variare delle esigenze dell'utenza. In questa ottica nei prossimi capitoli, dopo una breve panoramica sulle funzionalità attualmente disponibili, verranno illustrate le implementazioni da realizzare in tempi medio-brevi e alcune idee per il futuro.

Funzionalità attuali

Accesso

Allo stato attuale l'accesso alle pagine dell'applicativo web viene consentito ad utenti interni ed esterni tramite registrazione. Per il momento si tratta solo di una formalità atta ad avere informazioni circa la tipologia di utenza che frequenta il sito.

Si prevede, però per il futuro una politica degli accessi di diversa complessità i cui dettagli verranno illustrati in seguito. Per il momento dati necessari alla registrazione sono: Cognome, Nome, Affiliazione (tipologia di struttura di appartenenza), Denominazione (nome della struttura) e le informazioni che la riguardano, l'indirizzo e-mail se possibile, e naturalmente *UserID* e *Password*.

Ricerca

La ricerca può essere effettuata sia in lingua italiana che in inglese. In entrambi i casi si arriva a visualizzare le informazioni desiderate opportunamente ordinate in una griglia. Si può ricercare inserendo in una casella di testo un termine preciso oppure una stringa delimitata da uno o più asterischi. Questi si possono porre in testa, in mezzo o in coda, o in tutte e tre le posizioni, dando la possibilità all'utente di passare da una ricerca puntuale (termine esatto) ad un'altra estremamente generica e flessibile.

Navigazione all'interno del sito della NLM

Scegliendo invece un termine in inglese esiste la possibilità di collegarsi al sito della NLM puntando direttamente alle informazioni che lo descrivono. Naturalmente una volta entrati nel sito si può consultarlo a piacimento tornando alla *home page* o semplicemente ricercando altri termini. Tutto questo può essere evidentemente fatto anche entrando direttamente nel sito NLM, ma la *facility* consiste appunto nel passare da un sito all'altro avendo come parametro l'informazione a cui si è interessati.

Esplorazione della struttura ad albero

I termini visualizzati all'interno della griglia hanno anche funzione di *hyperlink*, e quindi da essi si può accedere ad informazioni di vario tipo. Con un *click* del mouse sul termine in italiano, ad esempio, si può esplorare la sezione della traduzione italiana che contiene il termine stesso. La struttura dei dati viene detta "ad albero" anche a causa delle ramificazioni ottenute navigandovi all'interno. Le informazioni sono infatti organizzate in modo gerarchico: esistono dei termini "radice" che sono i più generici, quelli che esprimono i concetti più estesi; e questi sono i primi ad essere visualizzati (Figura 1).

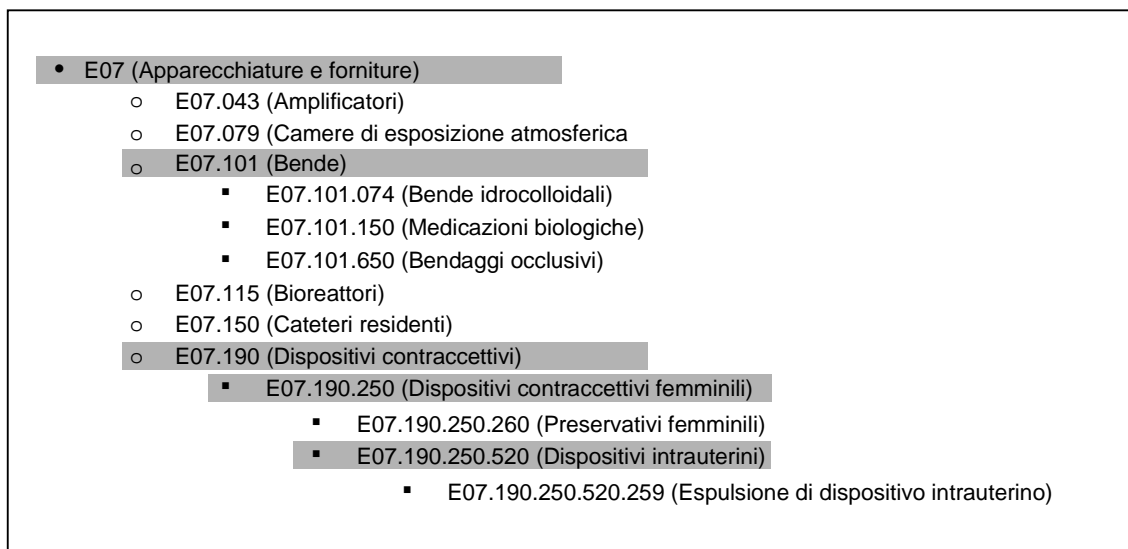


Figura 1. Modello di struttura ad albero: in grigio sono i rami con uno o più livelli al di sotto di essi, gli altri non ne hanno e quindi sono denominati "foglie"

Al di sotto di essi si può sviluppare un numero variabile di livelli detti “rami”. Infine il termine che risiede all’ultimo livello dell’albero viene denominato “foglia”. Uno dei vantaggi di una consultazione di dati così strutturati (quando è possibile organizzarli in questo modo) risiede nel fatto che partendo da una lista di elementi estremamente compattata si può arrivare a visualizzare una grande quantità di informazioni ordinate, verificando in tempo reale l’appartenenza di un termine ad un gruppo o sottogruppo. Naturalmente i livelli si possono aprire e richiudere dinamicamente, in modo da non generare confusione nell’utente con liste eccessivamente lunghe. Con l’applicativo MeSH ISS è possibile sia navigare all’interno di una specifica sezione, sia esplorare l’albero dalla radice ossia a partire da un elenco di macrocategorie.

Implementazioni

Formazione di tipologie di utenza

Come sopra descritto, attualmente accedono alla banca dati due diverse tipologie di utenti: da un lato coloro che usufruiscono delle informazioni in essa contenute, dall’altro il personale che esamina i vari termini e ne inserisce la traduzione. Al fine di allargare il contributo di tutti coloro che in qualche modo vogliono collaborare per migliorare la qualità del *DataBase* (DB) dei MeSH si ritiene importante cercare di diversificare le tipologie di utenti che nel tempo la utilizzeranno. Questo non solo per cercare di capire chi si collega e quindi concentrare gli sforzi in direzione delle necessità tipiche di chi usufruisce del dato, ma anche e soprattutto al fine di sfruttare competenze specifiche già esistenti in modo adeguato. L’idea di base è di prevedere quattro tipi di utenza; tale ipotesi di partenza non è però chiusa a suggerimenti che potranno essere valutati ed eventualmente verranno accolti nel tempo.

Schematicamente gli utenti si dividono in:

- *utente generico*
che può accedere a tutte le informazioni in sola lettura, il quale oltre ad utilizzarle per gli ovvi scopi lavorativi può dare suggerimenti per il miglioramento della applicazione stessa;
- *utente specializzato*
che può invece fornire un contributo inerente al proprio campo di specializzazione; esprime quindi un parere da prendere attentamente in considerazione al momento di tradurre un termine. Anch’egli accede in sola lettura al database, ma potrebbe avere un’area riservata nella quale scrivere i propri suggerimenti e le proprie osservazioni;
- *utente traduttore*
rappresentato da chi è autorizzato a tradurre o inserire termini. Questa è la categoria operativa abilitata a memorizzare i termini tradotti all’interno della base dati;
- *validatore*
ovvero colui che approva in maniera ufficiale l’ingresso di un termine nel database. Oltre a ciò analizza e valuta i suggerimenti degli utenti determinando in definitiva la qualità delle informazioni messe a disposizione.

Abbiamo già accennato al fatto che lo schema suddetto è ovviamente implementabile e quindi nulla esclude che le categorie possano essere di più o diverse, o che alcune di esse possano sovrapporsi. Dal punto di vista informatico questo si realizzerebbe impostando una politica degli accessi (sempre tramite *login* e *password*) che inserisca ogni utente in un gruppo opportunamente configurato. Al gruppo, che potremmo per comodità definire ad esempio nel

terzo caso *Mesh Traslator_Users*, verrebbero assegnati privilegi di lettura, scrittura e modifica dei dati e la possibilità di navigare attraverso le pagine dell'applicativo. Alcune pagine però potrebbero essere rese invisibili a questo tipo di utente e altre potrebbero contenere oggetti non disponibili a tutte le categorie di utenza. Arrivando fino al dettaglio del singolo oggetto si possono ottenere risultati davvero interessanti nel diversificare ruoli, competenze e privilegi di ogni singolo gruppo. Una volta definiti i gruppi diventa banale aggiungere un utente al gruppo stesso in modo da fargliene ereditare le caratteristiche. Da notare che un utente può appartenere a più gruppi contemporaneamente ed estendere così i propri privilegi.

Inserimento dei sinonimi nella banca dati della traduzione italiana

Un altro punto fondamentale riguarda l'inserimento dei sinonimi legati ai termini tradotti. La realizzazione di questa implementazione è probabilmente quella che presenta i problemi di maggior rilevanza.

Oggi la traduzione in italiano riguarda solo i termini principali e non i sinonimi. L'inserimento dei sinonimi nel DB comporta due ordini di problemi: uno di tipo semantico e uno di tipo tecnico. Dal punto di vista semantico, infatti, è evidente che non è sempre possibile tradurre univocamente un termine sinonimo di un altro in lingua inglese con un termine in lingua italiana che sia contemporaneamente sinonimo del termine principale anglosassone considerato; ovvero un sinonimo in una lingua non lo è necessariamente in un'altra. Agendo meccanicamente si avrebbe la certezza di inquinare la base con dati assolutamente non significativi e di non considerare invece la reale natura di un legame tra un termine tradotto e il proprio sinonimo. Inoltre è ragionevole ipotizzare che in italiano vi siano dei sinonimi del termine principale italiano senza corrispettivo in inglese.

La ricerca, nel momento in cui anche i sinonimi saranno integrati nel DB verrà lievemente modificata. Sarà sempre consentito ricercare in italiano o in inglese un termine principale, ma si potrà anche scegliere di esaminare direttamente i sinonimi. È anche prevista una ricerca combinata sui termini principali (TP) e sui sinonimi (SI) contemporaneamente, come si può vedere nell'esempio riportato in Figura 2.

Inserisci il termine o stringa di ricerca

Paroti*

termine principale
 sinonimo
 entrambi

Cerca in italiano Cerca in inglese

Possibile risultato

| Codice Mesh | Inglese | Italiano |
|-------------|------------------------|-------------------------------|
| D010305 | Parotid Diseases (TP) | Malattie della parotide (TP) |
| D010306 | Parotid Gland (TP) | Ghiandola parotide (TP) |
| D010307 | Parotid Neoplasms (TP) | Neoplasie della parotide (TP) |
| D035421 | Parotid Region (TP) | Regione parotide (TP) |
| D010309 | Parotitis (TP) | Parotite (TP) |
| D010309 | Parotiditis (SI) | |

Figura 2. Future funzioni di ricerca da implementare sulla Traduzione italiana dei MeSH

Nella figura precedente viene simulata la ricerca non puntuale di un termine in inglese; essa oltre ai termini principali estrae anche un sinonimo che non ha corrispettivo in italiano. Naturalmente un risultato perfettamente simmetrico si otterrebbe in caso di ricerca di un termine in italiano.

L'identità di codice MESH ci fa capire che il sinonimo e il termine principale si riferiscono ad un unico termine. Selezionando il termine "Parotitis" si ottiene il risultato espresso dalla Figura 3, dalla quale si evidenziano i termini con le rispettive traduzioni. Tutte le informazioni chiave saranno visualizzate sotto forma di *hyperlink* e attraverso di esse sarà possibile collegarsi al sito della NLM (o eventualmente altri siti) nonché esplorare il "Tree number" (quindi la struttura ad albero) sia in inglese che in italiano.

| | English | Italiano |
|-----------------------------|--|-------------------------------------|
| MeSH Heading | Parotitis | Parotite |
| Tree Number | C07.465.815.470.800 | C07.465.815.470.800 |
| Tree Number | C07.465.815.793.500 | C07.465.815.793.500 |
| Annotation | in translations use "parotitis", not "parotiditis" despite form of vernacular | ? |
| Scope Note | INFLAMMATION of the PAROTID GLAND . | ? |
| Entry Term | Parotiditis | traduzione |
| Allowable Qualifiers | BL CF CI CL CN CO DH DI DT EC EH EM EN EP ET GE HI IM ME MI MO NU PA PC PP PS PX RA RH RI RT SU TH UR US VE VI | ko |
| Unique ID | D010309 | invariante |

Figura 3. Schermata con le informazioni in forma tabellare relative al termine selezionato nel corso di una specifica ricerca

Scendendo su un piano puramente tecnico la gestione dei sinonimi comporta una serie di modifiche:

1. aggiornamento della procedura di *download* dei dati a partire da un file XML;
2. trasformazione delle informazioni ricevute in dati tabellari atti ad implementare il DB relazionale;
3. opportuna modifica delle tabelle della base di dati MeSH in relazione alle nuove specifiche;
4. ridisegno delle pagine attive di consultazione e di manutenzione dell'applicativo web al fine di rendere disponibili le nuove informazioni;
5. valutazione e realizzazione delle funzionalità da implementare nel nuovo sito.

Collegamento al servizio Pubmed della NLM

Un nuovo collegamento viene considerato di notevole importanza: quello al servizio PubMed della NLM, con le stesse modalità utilizzate per accedere al MeSH *Browser*, arrivando direttamente al termine puntuale partendo dalla base dati ISS. Qualora ci fosse interesse, questo tipo di collegamento si potrebbe estendere anche ad altri siti, sempre accordandosi sui parametri da passare per raggiungere l'informazione desiderata.

Brevi specifiche tecniche

La base di dati MeSH è stata realizzata utilizzando un *database* relazionale in ambiente *SQL Server*. I dati sono quindi organizzati in tabelle messe opportunamente in relazione tra di loro. Queste vengono annualmente popolate partendo dai dati messi a disposizione dalla NLM, dopo elaborazione da parte del personale ISS. Attualmente l’NLM fornisce un file XML, contenente anche la traduzione italiana che viene inserita da personale ISS mediante collegamento al sito della NLM stessa. L’utilizzo del formato XML, acronimo di *eXtensible Markup Language*, si può definire un testo strutturato e comporta indubbiamente vantaggi da un punto di vista di controllo della qualità formale dei dati per evitare appunto le ambiguità strutturali. Inoltre i dati sono delimitati da *tag* che ne indicano il significato; se uno di essi viene dimenticato oppure non è corretto, tutto il file diviene inutilizzabile e una applicazione che lo utilizza segnala un errore. La dimensione maggiore rispetto ad un file binario (per la presenza dei *tag* e altre informazioni caratteristiche del formato), non costituisce ormai un problema. L’unico svantaggio reale nel caso specifico consiste nel fatto che il file XML va convertito in dati tabellari da acquisire nel DB MeSH e ciò comporta una elaborazione preliminare. I dati residenti nel DB sono comunque esportabili in tutti i formati comunemente utilizzati, in particolare quelli previsti dall’ambiente Office (Access, Excel) oltre naturalmente all’XML. L’applicativo web MeSH – Traduzione Italiana è stato realizzato costruendo pagine attive (*Active Server Pages*, ASP) e pagine HTML (*Hyper Text Mark-Up Language*) con tecnologia Microsoft ASP.NET 1.1 in ambiente Visual Studio 2003. Questi strumenti, validissimi per la costruzione di siti e applicativi sfruttando le potenzialità della programmazione ad oggetti, sono però da un punto di vista di versione da considerarsi obsoleti. In effetti alcuni passaggi si sono rivelati particolarmente onerosi, in special modo quelli riguardanti la costruzione della struttura ad albero. Per realizzarla si è dovuto ricorrere a routine di Javascript per la gestione visuale e ad un motore implementato in Visual Basic.NET per la cattura e la gestione delle informazioni. Questo perché l’ambiente di sviluppo (nella sezione dedicata alle web *Application*) non prevedeva al momento un oggetto con proprietà e metodi adatti a risolvere il problema. Tale oggetto (*TreeView*) ora è stato sviluppato ed è disponibile nel *framework* Microsoft ASP.NET 2.0 utilizzabile con il Visual Studio 2005. Questi offrono la possibilità di utilizzare oggetti più adeguati, di usare interfacce grafiche che permettono di scrivere molto meno codice ottenendo risultati spesso migliori e snellendo le procedure, e maggior facilità di stratificazione delle componenti applicative (es. separazione della componente di accesso ai dati da quella di presentazione). Inoltre hanno al loro interno meccanismi di validazione che rendono più confortevole adeguarsi agli standard di accessibilità e usabilità necessari alla costruzione di siti e applicativi ben strutturati; standard ai quali le Pubbliche Amministrazioni sono e saranno in futuro sempre più vincolati.

Conclusioni

Ciò che emerge in modo più evidente dalle affermazioni fatte in precedenza è la opportunità di tendere ad una collaborazione tra ISS e coloro che usufruiscono della banca dati. Cercare quindi di accoglierne, ove possibile, i suggerimenti atti a migliorare sia la qualità delle informazioni messe a disposizione sia lo strumento deputato alla loro consultazione. Emerge anche la necessità di implementazioni applicative in linea con lo sviluppo delle tecnologie informatiche che permetta di migliorare le *performance*, di abbreviare i tempi di realizzazione e manutenzione delle applicazioni e di conformarsi il più possibile agli standard di riferimento sulla materia.

L'APPLICAZIONE DEL MESH NELLA BIBLIOTECA DIGITALE AZALEA: DAL LINGUAGGIO DEL PAZIENTE AL LINGUAGGIO CONTROLLATO

Gaetana Cognetti (a), Katuscia Dormi (a), Paola Seroppi (b), Ivana Truccolo (b)

(a) *Biblioteca Scientifica Centro di conoscenza "Riccardo Maceratini", Istituto Nazionale Tumori
Regina Elena, Roma*

(b) *Biblioteca Scientifica, Centro di riferimento oncologico, Aviano
cognetti.bib@ifo.it*

Cenni su biblioteche e informazione ai pazienti in Italia

I sistemi informativi rivolti ai pazienti, ancora di più di quelli finalizzati agli operatori sanitari richiedono, per facilitare l'accesso alle informazioni, la possibilità di utilizzo del linguaggio naturale e, ai fini dell'efficacia della ricerca, l'applicazione di sistemi di mappatura verso il linguaggio controllato (*thesauri*).

In USA già a partire dal 1975, presso la *National Library of Medicine* (NLM), si possono trovare tracce di progetti di mappatura dal linguaggio naturale al linguaggio controllato nelle basi dati biomediche (1), ma verosimilmente tale cammino è iniziato prima. Su questo argomento si può fare un po' di "storia" anche in Italia.

Una pietra miliare in tale settore è costituita dall'attivazione dei servizi di informazione ai pazienti. In campo oncologico la Biblioteca per i Pazienti di Aviano rappresenta il progetto pilota – già attivata nel 1998 e inaugurata con un convegno nel 1999 (2). Anche presso l'Istituto Regina Elena di Roma è stata di recente attivata la Biblioteca del Paziente inaugurata anch'essa con un convegno nel 2005 (3).

Vi sono ormai in Italia, a differenza di 8-9 anni fa, parecchie esperienze di servizi di informazione e supporto ai pazienti. Fra di essi vi sono molti elementi di similitudine e alcune diversità. Li unisce l'orientamento al paziente e alcuni principi cardine su cui sono basati. Un censimento di tali realtà è in corso all'interno del progetto SICOP (Sistema Informativo per la Comunicazione Oncologica ai Pazienti), finanziato dal ministero della Salute nel 2006 e coordinato dall'Istituto Regina Elena.

Le biblioteche degli IRCCS (Istituti di Ricovero e Cura a Carattere Scientifico) specializzati in oncologia (Figura 1) hanno fra l'altro prodotto una Biblioteca Digitale accessibile sul sito www.azaleaweb.it* che è divenuto supporto operativo delle esperienze di biblioteche fisiche per i pazienti diffuse in varie istituzioni.

In questa ultima decade è quindi accaduto che uno sparuto gruppo di bibliotecari, in collaborazione con psicologi, medici e altri professionisti della salute, abbiano realizzato, anche nel nostro Paese, ciò che in America e in altri Paesi europei era, già da tempo, una realtà.

Di fronte alle domande di maggiore informazione di pazienti, familiari e cittadini, ai bibliotecari e documentalisti è sembrato doveroso fornire una risposta organizzata e professionale per almeno 3 ragioni:

* Il sito alla data di pubblicazione di questo documento risulta non più attivo (*nota dei curatori*).

- a) la deontologia professionale *Every Reader his Book e Every Book its Reader* (cioè “ad ogni lettore il suo libro e a ogni libro il suo lettore”), come insegna un padre della biblioteconomia: S. R. Ranganathan;
- b) il contesto di lavoro: un Istituto di Ricovero e Cura a carattere Scientifico, vale a dire il luogo della ricerca translazionale, cioè della ricerca che si fa pratica clinica e viceversa;
- c) le esperienze internazionali già sviluppate da tempo: CancerBacup (UK) i cui opuscoli sono tradotti in Italia da AIMaC, Cancer.gov (USA), Medlineplus (USA), CISMef (FR), Caphis (USA), ecc.



Figura 1. Le Biblioteche degli IRCCS partecipanti al Progetto Azalea

Il modo più naturale per fare proprio il linguaggio del paziente, passa attraverso il suo ascolto. Infatti, uno dei tre punti cardine delle biblioteche per i pazienti in Italia è la presenza di un intermediario che deve essere un professionista dell’ascolto e dell’accoglienza oltre che un buon informatore (4). Gli altri due punti sono: uno spazio dedicato e la presenza di materiale adeguato.

La banca dati dell’informazione ai pazienti in oncologia: prodotto e servizio

Una banca dati della letteratura raccolta dalla Biblioteca per i Pazienti (primo prototipo di Azalea) era un già una realtà al CRO di Aviano nel 1999 (5); è diventato un progetto di Ricerca

Finalizzata degli IRCCS oncologici nel 2001 ed è stato poi finanziato nel 2003 come progetto multicentrico collaborativo da Alleanza Contro il Cancro (6).

Il MESH tradotto in italiano dall'Istituto Superiore di Sanità è stato sperimentato per la prima volta nella banca dati prodotta dalla Biblioteca per i Pazienti del CRO di Aviano e pubblicata su CD-ROM (Figura 2) nel 2002-2003, grazie ad un piccolo finanziamento ottenuto con il "Premio 5 Stelle". I 900 record raccolti nella banca dati hanno costituito il primo nucleo di Azalea.



Figura 2. Visualizzazione della pagina dei contribuiti alla banca dati oncologica della Biblioteca del CRO di Aviano (2003)

Già in questo primo prototipo la ricerca per soggetti MeSH in italiano e inglese era stata potenziata con l'aggiunta di tabelle di sinonimi. Così, se un utente ricercava "seno", il sistema, grazie all'indicizzazione e ai sinonimi, recuperava documenti contenenti la parola "mammella".

La traduzione in italiano del *thesaurus* MESH era quindi veramente una necessità per costruire una banca dati di materiale informativo per i pazienti.

Azalea: la biblioteca digitale per i pazienti in oncologia

Azalea, la biblioteca digitale di materiale informativo sul cancro per pazienti, familiari e cittadini, supportata da Alleanza Contro il Cancro (7), rende disponibili più di 3.300 record in totale, 1.892 record relativi al materiale informativo (614 a testo completo, 229 siti web, 89 sperimentazioni cliniche, 822 documenti ottenibili tramite le biblioteche che lo possiedono); 557 record di associazioni di volontariato e di organizzazioni che operano in ambito oncologico. In realtà le associazioni sono 1.425 in totale, tenuto conto anche delle sedi periferiche. Azalea conta circa 19.000 visitatori unici al mese.

Azalea, premiata al “Forum PA 2004” fra i “Buoni Esempi della Pubblica Amministrazione”, sta attraversando una delicata fase: si sta, infatti, espandendo anche a tematiche diverse dall’oncologia, in transizione verso il sistema Bibliosan, il portale delle Biblioteche degli Enti di ricerca biomedica dell’SSN.

Nella Figura 3 si riporta una pagina web di Azalea che presenta alcune sue caratteristiche di ricerca.



Figura 3. Pagina web con alcune caratteristiche di Azalea

I MeSH in Azalea

Perché Azalea ha utilizzato il MeSH insieme ad altri standard?

L'utilizzo degli standard è essenziale per garantire una adeguata strutturazione dei dati, l'integrazione tra diverse tipologie di informazione e l'interoperabilità tra diversi sistemi a livello internazionale definita dalla Direttiva dell'Unione Europea (91/250/EEC) come: *the ability to exchange information and mutually to use the information which has been exchanged* (8).

L'applicazione di metadati, e in particolare di metadati standard, consente, secondo una comunicazione emanata dall'Unione Europea: "eEurope 2002: Criteri di qualità per i siti web contenenti informazioni di carattere medico"(9), di garantire l'accessibilità ai documenti quale elemento essenziale di qualità. Il metadato relativo al contenuto semantico è sicuramente quello più rilevante laddove si debbano ricercare informazioni sulla salute. Il MeSH è ormai divenuto uno standard *de facto*, tradotto in moltissime lingue e applicato non solo nella catalogazione delle biblioteche biomediche e nella indicizzazione di basi e banche dati, ma anche nella organizzazione semantica di siti web di qualità, quali CISMEF, HONSELECT, WHO, OMNI, MEDCIRCLE, ecc.

Azalea può definirsi un servizio di informazione integrata per i pazienti con l'uso di standard. Gli standard adottati da Azalea sono:

1. standard descrittivi di vario materiale informativo come ad esempio per le monografie, i seriali, le risorse elettroniche: *International Standard Bibliographic Description /ISBD* (m), (s), (er),(cm), ecc.;
2. standard nazionali per l'intestazione per autori, basati su comuni regole internazionali: RICA/ Regole Italiane Catalogazione Autori;
3. standard internazionali per la soggettazione come il *Medical Subject Headings* (MESH) che si avvale oggi dell'ampliamento prodotto dell'*Unified Medical Language System* (UMLS), il più importante sistema terminologico e di conoscenza per la biomedicina, prodotto dalla *National Library of Medicine* che mette in relazione oltre 100 terminologie standard, per mezzo di un metatesauro e di una rete semantica e che applica sistemi di analisi del linguaggio naturale;
4. classificazioni standard Internazionali come la *National Library of Medicine Classification*, la *Library of Congress Classification* (LCC);
5. Dublin Core-ISO norma 15836:2003 *minimum data set*, composto da 15 campi, per l'integrazione di informazioni provenienti da diverse tipologie di documenti.

Tutti questi standard e altri (es. la classificazione Dewey e CDU e i *thesauri* di molte basi dati) sono stati prodotti in ambito bibliotecario per la gestione dell'informazione e della conoscenza.

La comunità internazionale bibliotecaria ha, infatti, prodotto tali standard per poter gestire efficacemente la selezione e lo scambio di milioni di informazioni derivate da libri, riviste, letteratura grigia ecc. su vari supporti (elettronico, cartaceo, audiovisivo, ecc.) problema divenuto ancora più drammatico con l'avvento di Internet. Azalea è stata la prima risorsa informativa per pazienti che ha utilizzato il *thesaurus* MESH in lingua italiana grazie alla collaborazione dell'Istituto Superiore di Sanità che ha reso lo disponibile nella versione italiana.

In Azalea il MeSH fa da *trait d'union* tra i diversi tipi di documenti; infatti, sia gli opuscoli per i pazienti, sia i siti web, sia le associazioni, sia i protocolli, sono indicizzati con i termini MeSH in lingua inglese e italiana. Ciò permette di recuperare, inserendo un unico termine di ricerca, tutta l'informazione su una data patologia o su un dato argomento.

Per facilitare la ricerca di informazioni il *team* di Azalea sta potenziando il *thesaurus* MeSH con l'inserimento di sinonimi, sia in lingua inglese che in lingua italiana, che rinviano automaticamente alle voci controllate (descrittori MeSH) (Figura 4).

Poiché si tratta di una banca dati di materiale esclusivamente oncologico, per rendere ancora più facile la ricerca, si è creata una mappatura automatica anche dai termini semplici degli organi (mammella, prostata, ecc.) verso i termini MeSH relativi alle patologie oncologiche (es. neoplasia della mammella). Naturalmente nell'ipotesi di allargamento di Azalea ad altre patologie (AZALEAPLUS) si dovranno rivedere queste relazioni terminologiche.

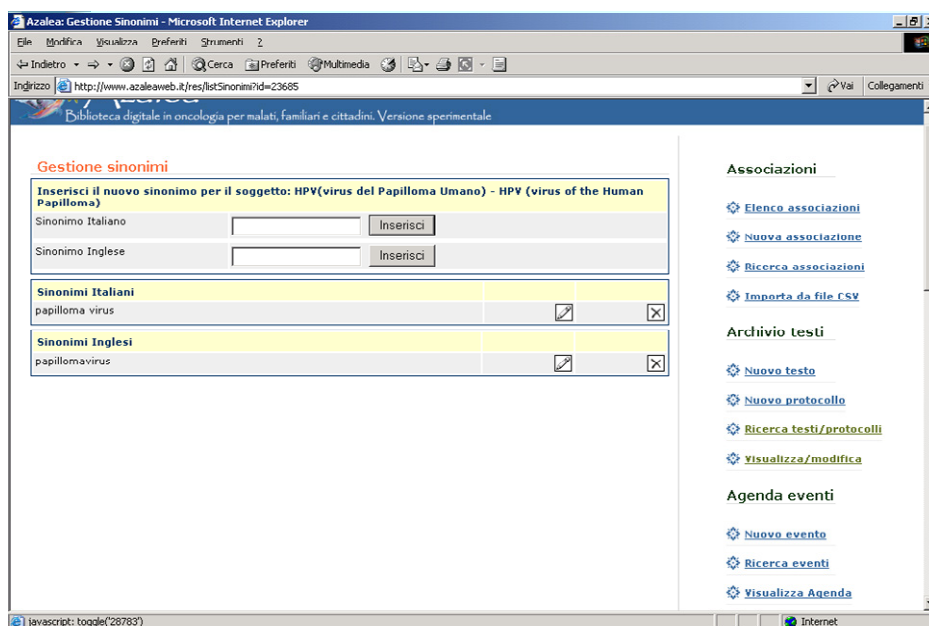


Figura 4. Gestione dei sinonimi nella parte gestionale di Azalea

L'accesso con l'utilizzo della terminologia MeSH è possibile con diverse alternative, sia digitando nella finestra di ricerca il termine MeSH o il suo sinonimo, sia selezionando i termini in un alfabetiere, sia con un menu dei profili più ricercati (prevenzione, fattori ambientali, diritti del malato, ecc.), sia attraverso la selezione di parti del corpo umano (vedi Figura 3). In quest'ultimo caso si è operata "manualmente" una sorta di esplosione dei termini MeSH per cui purtroppo non è ancora presente in Azalea un modulo automatico come in Pubmed.

Sarebbe interessante se l'Istituto Superiore di Sanità sviluppasse un prototipo per la gestione dell'esplosione dei termini di cui potrebbero usufruire i vari sistemi informativi che applicano il *thesaurus* MeSH sulla pagina principale

Come si è visto il linguaggio MeSH è diventato una terminologia standard d'elezione per l'indicizzazione sia delle risorse informative per gli operatori sanitari, ma anche di quelle rivolte al pubblico.

Non solo il *thesaurus* MeSH permette di accedere in maniera esaustiva e precisa all'informazione, ma è anche il tramite per la navigazione in sistemi di conoscenza biomedica che ormai devono essere condivisi tra gli operatori e i pazienti. La NLM ha, infatti, prodotto, come già accennato, l'UMLS che tramite il *metathesaurus*, la rete semantica e sistemi di analisi del linguaggio naturale permette di fondare una vera e propria ontologia medica. Il concetto di ontologia proviene dalla filosofia e identifica la possibilità di utilizzare l'informazione e le conoscenze per navigare nella letteratura. Questa costituisce un vero e proprio ambiente di ricerca dove è anche possibile effettuare scoperte. Non solo gli operatori sanitari, ma anche i pazienti possono divenire in prima persona attori e promotori della ricerca scientifica, come è già avvenuto nel caso dei genitori di Lorenzo i quali, pur non essendo medici, attraverso l'utilizzo del Medline, sono riusciti a scoprire le dinamiche e la cura della grave patologia del figlio, l'adrenoleucodistrofia, salvando così i bambini affetti da questa malattia genetica. L'uso degli standard da parte dei sistemi informativi per pazienti potrebbe consentire, nel futuro, la costruzione di un unico *gateway* capace di rendere disponibili e diffondere tutte le informazioni prodotte in qualsiasi lingua a livello internazionale.

Bibliografia

1. Hints for computer searching of natural language. *Medlars Technical Bulletin* 1975;74:8.
2. Truccolo I, Ciolfi L, Annunziata M A, Piani B, Turrin O, Venturelli M. Quando l'utente è un paziente. *Biblioteche Oggi* 1998;16(7):26-30.
3. Cognetti G, Cecere L, Dormi K, Solero F, Tuveri M, D'Orsogna F. *Azalea: un sistema di conoscenza in oncologia a disposizione di pazienti e cittadini. Atti del Convegno Azalea: un sistema di conoscenza in oncologia a disposizione di pazienti e cittadini. Roma, 16 giugno 2005.* Disponibile all'indirizzo: <http://eprints.rclis.org/archive/00007198/>; ultima consultazione 31/10/06.
4. Biblioteca per i Pazienti - Biblioteca Scientifica e Area di Psicologia del CRO (Ed.). Raccomandazioni per la gestione dei servizi di informazione, in materia di salute, rivolti alla comunità e collocati all'interno di istituzioni sanitarie. Aviano: Centro di Riferimento Oncologico Aviano, 2006. Disponibile all'indirizzo: <http://eprints.rclis.org/archive/00006897/>; ultima consultazione 31/10/06.
5. Truccolo I, Cognetti G, Annunziata M A, Piani B. Le informazioni ai pazienti: una base di dati delle risorse informative esistenti in Italia. In: Alberani V, De Castro P (Ed.). *La letteratura grigia: politica e pratica. 3° Convegno nazionale. Atti. Roma, 25-26 novembre 1999.* Roma: Istituto Superiore di Sanità; 2000. (ISTISAN Congressi 67). p. 98-103.
6. Bogliolo A, Russell-Edu W, Alloro G, Ballarini R, Bocchini G, Cecere L, Codonesu A, Cognetti G, Montella M, Ricci R, Seroppi P, Trocino A, Truffelli R, Cognetti F, Truccolo I. More information, more choice: an Italian database for oncology patients. *Ann Oncol* 2005;16(12):1962-67.
7. Truccolo I, Cognetti G, Russell-Edu W, Bogliolo A. Patients and cancer in the Internet age: guided pathways to diseases. [Power-point presentation] In: *10th Conference of the European Association for Health Information and Libraries. Cluj-Napoca, Romania, September 11-16, 2006.* Disponibile all'indirizzo: <http://www.eahilconfcluj.ro/documents/2c/truccolo.ppt>; ultima consultazione 31/10/06.
8. Council Directive of 14 May 1991 on the legal protection of computer programs (91/250/EEC). Disponibile all'indirizzo: Http://ec.europa.eu/internal_market/copyright/docs/docs/1991-250_en.pdf; ultima consultazione 31/10/06.
9. Commissione delle Comunità Europee. Comunicazione COM (2002) 667 final. *eEurope 2002: Criteri di qualità per i siti web contenenti informazioni di carattere medico.* Disponibile all'indirizzo: http://europa.eu.int/information_society/eeurope/ehealth/doc/communication_acte_it_fin.pdf; ultima consultazione 31/10/06.

L'UTILIZZO DEI MESH IN ITALIANO NELL'EVOLUZIONE DEL PROGETTO ELISIR

Paolo Gardois (a), Sandro Garoffolo (a), Diego Arduino (b), Michele Mannucci (b), Renata Riva (b), Marcello Valente (b)

(a) Dipartimento di Scienze Pediatriche e dell'Adolescenza, Università di Torino

(b) Gruppo per l'OPAC e il Portale documentario d'Ateneo, Università di Torino

Storia del progetto ELISIR

Le risorse presenti sul web per ogni disciplina sono molte, in costante evoluzione e di variabile qualità. Sono intensivamente utilizzate dai ricercatori e dagli studenti delle facoltà universitarie, ma spesso sono di difficile reperimento. Si produce così un dispendio evitabile di tempo ed energie da parte di queste categorie di utenti. Inoltre il livello qualitativo della ricerca e della didattica prodotte potrebbe essere più elevato in assenza di questi problemi.

Navigando nel *mare magnum* del web mediante motori di ricerca generalisti (es. Google, Yahoo) si reperiscono spesso informazioni di variabile qualità scientifica, anche utilizzando i repertori per argomento generalisti (*directory*) disponibili sul web (Google Directory, Yahoo Directory, ecc.).

I criteri di selezione non risultano chiari: alcune risorse presentano *award* di approvazione, altre ne sono del tutto sprovviste; alcune dipendono da istituzioni note e rinomate, altre sono il frutto di iniziative individuali, non certificate o valutate.

I repertori disciplinari (o *virtual reference desk*) sono molto diffusi ma spesso di dimensioni relativamente piccole e scarsamente aggiornati. Inoltre, sono rari i casi di repertori accademici multidisciplinari, perlopiù realizzati da istituzioni estere, che raramente prendono in considerazione risorse interessanti per l'utenza italiana.

In risposta a tale situazione e alle necessità degli utenti, ELISIR nasce nel 2002 su iniziativa della Biblioteca del Dipartimento di Scienze Pediatriche e dell'Adolescenza come piattaforma per la catalogazione delle risorse Internet su base disciplinare. Il progetto, nato dall'esigenza di fornire a pediatri, ginecologi, ostetriche e infermieri (nel ruolo di studenti, specializzandi, dottorandi, ricercatori, docenti, ecc.) e al territorio risorse web d'interesse.

Il progetto, dopo una fase iniziale indirizzata espressamente alla sola utenza biomedica, si accinge ora ad essere esteso a tutti gli ambiti disciplinari afferenti all'Ateneo torinese.

La parte teorica e i primi risultati della sperimentazione pianificata con gli utenti sono stati presentati all'8th *European Association for Health Information and Libraries Conference*, settembre 2002 (1).

Le biblioteche aderenti ad ELISIR sono:

- Biblioteca Dipartimento di Scienze Pediatriche e dell'Adolescenza dell'Università di Torino <http://www.lib.unito.it/>;
- Biblioteca "Rosmini" dei Corsi di Laurea Professioni Sanitarie, ASO S. Giovanni Battista di Torino (da febbraio 2005 – prima come Biblioteca "Minerva medica") <http://www.lsmolinette.unito.it/biblioteca/biblioteca.htm>;
- Centro di Documentazione Regionale Salute (DORS) di Grugliasco (TO) (da novembre 2003) <http://www.dors.it/>.

Il motore di ricerca di ELISIR si trova all'interno del sito del Laboratorio Informazione Biomedica (LIB) a cura della Biblioteca Unificata Materno-Infantile (<http://www.lib.unito.it/database/elisir.html>).

Attualmente nel catalogo sono presenti 349 risorse selezionate e catalogate dai bibliotecari. Lo standard di catalogazione utilizzato è il Dublin Core (<http://www.dublincore.org>).

Le risorse di ELISIR sono in sintesi:

- Lingua: 107 italiana - 252 inglese - 17 francese - 24 spagnola;
- Destinatari delle risorse: 231 medici - 270 altre professioni sanitarie - 203 cittadini;

I criteri essenziali per includere una risorsa nel catalogo ELISIR sono:

- scopo e argomento della risorsa chiari;
- risorsa pertinente per almeno uno dei nostri gruppi di utenti (medici, altri professionisti della sanità e cittadini);
- risorsa accessibile in Internet senza nessun costo, oppure sezione gratuita di una risorsa più grande che non è accessibile gratuitamente;
- lingua italiana come prima scelta nel caso di una risorsa multilingue, e quindi, in ordine, inglese, francese, spagnolo;
- sponsor e fonti di finanziamento indicati chiaramente;
- risorsa con garanzia di un corretto trattamento dei dati degli utenti collegati.

Le tipologie di risorse prese in considerazione sono:

- libri (anche multimediali) disponibili online;
- siti di associazioni professionali e di volontariato o per pazienti;
- database bibliografici online, subject gateways e motori di ricerca medici;
- cataloghi di biblioteche biomediche, liste (PEB, *Freemedicaljournals*, ecc.) e gateway di periodici online (EBSCO, OVID, ecc.);
- siti per la formazione, formazione a distanza ed educazione permanente;
- università e istituzioni sanitarie, ospedali, centri di cura e centri di ricerca;
- Evidence Based Medicine (EBM): centri, linee guida, protocolli;
- casi clinici e immagini;
- *thesauri*, glossari, enciclopedie e *quick reference*, risorse bibliografiche delle aree disciplinari prescelte (istruzioni agli autori; BFS, *Bibliographic Formatting Software*; ecc.).

Le aree disciplinari sono:

- pediatria;
- ginecologia e ostetricia;
- *nursing* e altre professioni sanitarie;
- informazione ai pazienti e promozione della salute.

MeSH in ELISIR

Inizialmente la soggettazione è stata svolta utilizzando solamente i MeSH in inglese, servendosi come riferimento del sito: <http://www.nlm.nih.gov/mesh/> e in particolare del MeSH *Browser* <http://www.nlm.nih.gov/mesh/MBrowser.html>, direttamente utilizzabile a partire dall'interfaccia di catalogazione (Figura 1).

Dal 2005 si è cominciato ad usare i MeSH in italiano disponibili sul sito dell'Istituto Superiore di Sanità (ISS), in aggiunta ai MeSH in inglese. Con i MeSH in italiano si intende ampliare il numero di utenti potenziali di ELISIR. Per quanto riguarda l'implementazione dei MeSH in italiano nel database, si tratta di un campo ripetibile della tabella principale.

| | |
|--|--|
| IDENTIFIER (compulsory) | <input type="text" value="http://www.chw.edu.au/prof/handbook/index.htm"/> |
| TITLE (compulsory) | <input type="text" value="The New Children's Hospital Handbook : Royal Alexe"/> |
| ALTERNATIVE TITLE (one per row) | <input type="text"/> |
| CREATOR Verifica o inserisci autore | <input type="text" value="Royal Alexandra Hospital for Children : Westmead N.S.W. Australia"/> <input type="button" value="AGGIUNGI"/> |
| SUBJECT Verifica o inserisci MeSH Verifica o inserisci MeSH in italiano | Mesh <input type="text" value="Pediatrics"/> <input type="button" value="..."/> <input type="text" value="Reference Books, Medical"/> <input type="button" value="..."/> <input type="text" value="Textbooks"/> <input type="button" value="..."/> <input type="button" value="AGGIUNGI"/> |
| | Mesh in italiano <input type="text" value="Pediatria"/> <input type="button" value="..."/> <input type="button" value="AGGIUNGI"/> |
| | DDC <input type="text"/> |
| | DESCRIPTION |

Figura 1. Interfaccia di catalogazione in ELISIR

Ad ogni risorsa catalogata si possono attribuire più soggetti, quindi la relazione tra le tabelle degli URI (*Uniform Resource Identifier*) e le tabelle dei MeSH è di molti a molti. I MeSH in italiano, nell'interfaccia di catalogazione in ELISIR, sono presentati in Figura 2.

| | | | |
|--|--------------------------|---------------------------------------|----------------------|
| MESH | <input type="text"/> | RICERCA | <input type="text"/> |
| <input type="button" value="INSERISCI"/> | | <input type="button" value="Search"/> | |
| Cerca nei MeSH(Link esterno) Chiudi la finestra | | | |
| Modifica e eliminazione | | | |
| Chirurgia | MODIFICA | ELIMINA | |
| Clinica medica | MODIFICA | ELIMINA | |
| Computer | MODIFICA | ELIMINA | |
| Controllo delle infezioni | MODIFICA | ELIMINA | |
| Controllo di qualità | MODIFICA | ELIMINA | |
| Diabete mellito | MODIFICA | ELIMINA | |
| Dieta | MODIFICA | ELIMINA | |
| Disordini genetici | MODIFICA | ELIMINA | |
| Disrafia spinale | MODIFICA | ELIMINA | |
| Editoria | MODIFICA | ELIMINA | |
| Educazione a distanza | MODIFICA | ELIMINA | |
| Educazione del paziente | MODIFICA | ELIMINA | |
| Educazione del paziente | MODIFICA | ELIMINA | |
| Educazione sanitaria | MODIFICA | ELIMINA | |
| Emoglobinopatie | MODIFICA | ELIMINA | |
| Completato | | | |

Figura 2. I MeSH in italiano nell'interfaccia di catalogazione in ELISIR

ELISIR: politica di soggettazione

Ad ogni risorsa ordinaria sono assegnati da 3 a 5 MeSH, sia in italiano sia in inglese. Per grandi portali e siti molto complessi (es. WHO, ecc.) si può arrivare a 10 MeSH. La politica di soggettazione tiene conto di:

- MeSH inglesi estratti da PubMed (<http://www.nlm.nih.gov/mesh/2002/MBrowser.html>);
- MeSH italiani estratti dal sito dell'ISS (<http://www.iss.it/site/mesh/>), a partire dal 2005.

I MeSH in italiano nell'interfaccia di ricerca in ELISIR sono presentati in Figura 3. In Figura 4 si presentano i risultati per la ricerca del termine "Pediatria".

Ricerca di base

- Inserire uno o più termini da ricercare nel campo "Cerca".
- Possono essere inseriti operatori logici per combinare le parole: AND (and, +), OR (or, |), NOT (not, -)
- Si può ricercare una frase esatta inserendo più parole di seguito senza operatori e racchiudendo tutta la frase tra virgolette ("search this only").
- Per ulteriori esempi e suggerimenti, vedere il [file di aiuto](#).
- Per effettuare selezioni multiple nei campi **Utenti** e **Lingua**, selezionare più voci tenendo premuto il tasto CONTROL.

Cerca [\[aiuto\]](#)

Dove [\[aiuto\]](#)

Maiuscole/minuscole [\[aiuto\]](#) Abilitate Disabilitate

Lingua [\[aiuto\]](#)

Data creazione [solo dove disponibile]

Data ultima modifica

Data ultima revisione catalogatore

Utenti [\[aiuto\]](#)

Tipo

Formato

Figura 3. I MeSH in italiano nell'interfaccia di ricerca in ELISIR

11 Risultati trovati per Ricerca di pediatria in MESH in italiano
Disabilità maiuscole/minuscole

| NOME SITO (vai alla scheda descrittiva) | INDIRIZZO (vai al sito) | RESTRIZIONI D'ACCESSO |
|--|---|---|
| 1 Canadian Paediatric Society = Société canadienne de pédiatrie | http://www.cps.ca/index.htm | |
| 2 Centro per la Valutazione dell'Efficacia dell'Assistenza Sanitaria | http://www.ceveas.it/ceveas/index.jsp | |
| 3 FIMP : Federazione Italiana Medici Pediatri | http://www.fimp.org/ | |
| 4 Istituto Giannina Gaslini - Ospedale pediatrico IRCCS | http://www.gaslini.org/ | |
| 5 MammaePapa.it | http://www.mammaepapa.it/home/index.asp | Richiede un'iscrizione gratuita per poter accedere ad alcune sezioni del sito: gruppi di discussione e richiesta di informazioni personali o servizi. |
| 6 News Release Archive of the American Academy of Pediatrics | http://www.aap.org/advocacy/archives/archive.htm | |

Figura 4. Risultati in ELISIR per la ricerca con MeSH in italiano per "Pediatria"

Proposte per una nuova versione di ELISIR

È allo studio la possibilità di inserimento dei MeSH in XML sia in inglese (NLM) sia in italiano (ISS).

I vantaggi principali sono costituiti dalla disponibilità di un *authority file* completo, dalla velocizzazione delle procedure di catalogazione e dall'eliminazione delle possibilità di errori in fase di inserimento dati.

Tra le prospettive future dell'evoluzione dell'interfaccia di ELISIR si segnalano le seguenti:

- segnalare gli aggiornamenti;
- salvare da parte degli utenti registrati le ricerche e impostare *alert* automatici sulle tematiche preferite;
- permettere agli utenti di segnalare nuove risorse;
- gestire strumenti di *community* come *forum*, ecc.

Bibliografia

1. Gardois P, Grillo G, Bassi C, Lingua C, Piga A, DaglioM. *ELISIR: Planning with users: a catalogue of Internet resources for clinicians, nurses and citizen*. 8th EAHIL Conference. Colonia (Germania), September 2002. Disponibile all'indirizzo http://www.zbmed.de/fileadmin/pdf_dateien/EAHIL_2002/gardois-proc.pdf; ultima consultazione 22/12/06.

IMPIEGO DEI MESH NEL PROCESSO EDITORIALE: DALLA SUBMISSION ALLA PUBBLICAZIONE

Paola De Castro, Federica Napolitani, Elisabetta Poltronieri
Settore Attività Editoriali – Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali
Istituto Superiore di Sanità, Roma
paola.decastro@iss.it

Introduzione

Il presente lavoro analizza l'impiego dei termini MeSH (*Medical Subject Headings*), sia nella versione inglese che italiana, nei diversi stadi del processo editoriale di una pubblicazione scientifica: dalla stesura di un manoscritto, alla sua presentazione in Redazione – attraverso le diverse fasi di lavorazione e revisione – fino alla sua pubblicazione definitiva e al successivo “deposito” in basi di dati bibliografiche o archivi digitali aperti.

In particolare, viene illustrato il ruolo svolto dal Settore Attività Editoriali dell'Istituto Superiore di Sanità (ISS) nel processo editoriale di pubblicazione delle riviste editate dall'ente, e nella promozione di un orientamento uniforme nell'assegnazione dei termini di indicizzazione della letteratura scientifica, anche in previsione di obiettivi di interoperabilità dell'archivio digitale istituzionale ISS, che verranno attuati con l'applicazione dei metadati per l'identificazione univoca delle risorse documentarie. Viene anche illustrata l'attività di registrazione delle pubblicazioni prodotte dai ricercatori ISS destinate alle serie istituzionali e non.

Sono presentati infine i risultati di uno studio finalizzato a rilevare l'utilizzo dei MeSH negli articoli di periodici sui quali la comunità di ricerca dell'Istituto ha maggiormente pubblicato nell'anno 2005.

Termini MeSH e processo editoriale

Il “processo editoriale” può essere inteso, in senso generale, come quel ciclo di lavorazione che porta alla pubblicazione di un documento e coinvolge diverse figure professionali. L'espressione ha dunque una valenza piuttosto ampia e può interessare diverse tipologie di prodotto editoriale: monografie (libri, atti di congressi, rapporti tecnici, tesi) o serie periodiche (riviste, notiziari, bollettini, serie di rapporti) e anche diverse tipologie di contributi da pubblicare all'interno di questi prodotti (articoli originali, articoli di rassegna, brevi note, editoriali, lettere, recensioni, segnalazioni).

Per gli scopi di questo contributo si prenderà in considerazione solo una ben precisa tipologia di documento: quella dell'articolo scientifico pubblicato su rivista indicizzata. Si intenderà quindi come processo editoriale, il tradizionale ciclo di pubblicazione degli articoli scientifici. Processo che si esplica attraverso un susseguirsi di fasi ben determinate, ciascuna delle quali risulta essere strettamente connessa sia a quella che immediatamente la precede sia a quella che la segue secondo un preciso e costante ordine di successione.

Dalla stesura originale, alla sua presentazione in Redazione, dal vaglio dei revisori, alle diverse correzioni redazionali fino alla stampa definitiva, il manoscritto attraversa numerose fasi, si

sofferma su molte scrivanie, viene vagliato da persone con specifiche competenze professionali (autori, editori, revisori, redattori, indicizzatori) e analizzato da differenti punti di vista.

Le parole chiave: un filo conduttore del processo editoriale

Uno degli elementi ricorrenti e rintracciabili in ciascuna fase del ciclo editoriale e che costituisce, per così dire, una sorta di “filo conduttore”, è rappresentato da quelle che vengono impropriamente definite come “parole chiave”, intese come termini vagamente significativi dei contenuti riportati in un testo. In realtà, più che di parole chiave (*key words* in inglese) sarebbe decisamente più appropriato parlare di “descrittori” ovvero di termini appartenenti a un linguaggio predefinito, controllato e non ambiguo, che identificano in modo univoco l'informazione contenuta in un articolo. Tuttavia, poiché il termine “parole chiave” è quello *de facto* utilizzato nell'editoria scientifica, lo si utilizzerà anche in questo lavoro, seppure con l'accezione di cui sopra.

Le parole chiave vengono individuate e formulate *ab initio* dall'autore del manoscritto, vagliate dai revisori, eventualmente corrette dai redattori, stampate dagli editori, utilizzate o riassegnate dagli indicizzatori, depositate negli archivi digitali istituzionali e ricercate dagli utenti. A quest'ultima fase del ciclo editoriale, che si “estende” al di là della pubblicazione del manoscritto ed è costituita dal recupero dell'informazione (*information retrieval*), sono stati dedicati altri contributi pubblicati in questo stesso Rapporto.

Il percorso compiuto dalle parole chiave all'interno del ciclo editoriale sarà di seguito analizzato con lo scopo di individuare quale sia il ruolo dei termini MeSH e soprattutto quale la loro rilevanza.

Autori e MeSH: l'individuazione dei descrittori

Sono gli autori di un articolo ad essere chiamati per primi ad individuare le parole chiave e a riportarle nel proprio manoscritto prima ancora che questo venga presentato per la pubblicazione.

Alla formulazione delle parole chiave non sempre gli autori dedicano l'attenzione dovuta. Eppure, il fatto che gli editori tradizionalmente vi concedano ampio spazio e particolare evidenza all'interno della prima pagina stampata di ogni singolo articolo (con accorgimenti tipografici quali spaziatura, variazione di carattere o di corpo, ecc.) dovrebbe indurre anche l'autore più superficiale ad interrogarsi sulla loro funzione, utilizzo e rilevanza.

Spesso tuttavia non è così. I ricercatori, a volte carenti di competenze in campo documentalistico e bibliotecario, difficilmente si rendono conto dell'importanza della scelta di adeguate parole chiave e si rifanno alle “Istruzioni agli Autori” formulate dagli editori delle singole riviste che raramente forniscono precise indicazioni (per esempio sulla opportunità di ricorrere, ove possibile, ai termini MeSH). La mancata percezione dell'importanza della scelta di una parola chiave appropriata è spesso alla base del mancato utilizzo dei termini MeSH da parte degli autori.

Revisori e MeSH: un'occasione mancata

Nel processo di *peer-review* (revisione dei manoscritti da parte di “pari”, cioè di qualificati esperti della materia), raramente i revisori si soffermano a valutare la congruità delle parole chiave di un manoscritto, né d'altronde le guide che le Redazioni stilano allo scopo di coadiuvare il loro operato ne fanno in genere menzione. Se tra i punti ritenuti in qualche modo

qualificanti ai fini dell'accettazione del manoscritto fosse inserito anche un riferimento alla congruità e significatività delle parole chiave (e magari l'eventuale ricorso ai termini MeSH), al revisore non sfuggirebbe una valutazione anche di questo non irrilevante aspetto.

Sarebbe quindi auspicabile che le riviste *peer-reviewed* inseriscano nelle proprie "Guide ai revisori" un accenno alla verifica delle parole chiave e nelle "Istruzioni agli autori" il suggerimento (se non l'obbligo) di ricorrere (ove possibile) all'utilizzo dei termini MeSH.

Editori e MeSH: una doppia responsabilità

Le riviste scientifiche di norma pubblicano "Istruzioni" o "Guide", redatte dall'editore, le quali hanno lo scopo pratico di fornire all'autore informazioni utili per la stesura del manoscritto, offrendo consigli e norme pratiche che gli consentano di redigere il testo in modo chiaro, conciso e opportunamente strutturato (si veda ad es. la cosiddetta strutturazione IMRAD: *Introduction, Materials and methods, Results And Discussion*) e soprattutto in modo conforme alle particolari regole in uso in quella specifica rivista (impostazione della bibliografia secondo un determinato stile, presentazione di tabelle e figure e così via). Le "Istruzioni agli Autori" possono essere più o meno dettagliate, riassunte in una singola pagina che usualmente coincide con la pagina finale di ciascun fascicolo, oppure redatte su più pagine e suddivise in varie sezioni, qualora la rivista pubblichi diverse tipologie di contributi (lettere, editoriali, monografie, *clinical trial*, ecc.). Nel caso degli *Annali dell'Istituto Superiore di Sanità* (rivista scientifica *peer-reviewed*), esse sono riportate in un'unica pagina pubblicata alla fine di ciascun fascicolo.

Le "Istruzioni agli Autori" delle maggiori riviste biomediche sono disponibili online, sia presso i siti delle singole riviste, sia in siti appositamente dedicati (si veda *Instructions to Authors in Health Sciences* che raggruppa Istruzioni di più di 3.500 riviste: <http://mulford.mco.edu/instr>).

Da un rapido esame condotto su rinomate riviste biomediche internazionali, i cosiddetti "*big journals*" tra cui *Nature*, *Science*, *Lancet*, *BMJ*, *Annals of Internal Medicine*, *The New England Journal of Medicine*, è emerso che scarsa attenzione viene assegnata alle parole chiave nelle "Istruzioni agli Autori" e che l'uso della terminologia MeSH non viene espressamente suggerito, pur essendo ovviamente tali riviste indicizzate su PubMed. Negli *Annali ISS* si raccomanda, già da qualche anno, di assegnare termini MeSH ai propri contributi, ove possibile.

Nelle "Istruzioni agli Autori" degli *Annali*, già da qualche anno, è stato aggiunto un paragrafo in cui si consiglia l'uso dei termini MeSH nell'assegnazione delle parole chiave del contributo, ove possibile.

Il redattore ha sicuramente tra i suoi compiti quello di verificare con attenzione e scrupolosità che le parole chiave siano appropriate rispetto ai contenuti dell'articolo, ma non sempre si agisce in tal senso. Ove le "Istruzioni agli Autori" prevedano l'uso dei termini MeSH, dovrà anche verificarne il concreto e corretto utilizzo da parte degli autori. Anche se ciò non fosse previsto, il redattore dovrebbe comunque valutare l'opportunità di farvi ricorso e, in ogni caso, nell'*editing* del manoscritto non dovrebbe mai sorvolare sul testo, e tanto meno sulla parte che riguarda le parole chiave.

Vantaggi nell'uso dei MeSH

Da questo breve resoconto delle diverse fasi del processo editoriale, considerato in relazione all'identificazione delle parole chiave, emergono chiaramente i vantaggi che l'utilizzo dei MeSH potrebbe comportare a diversi livelli.

L'utilizzo di termini MeSH *ab initio* (ovvero la loro formulazione da parte dell'autore stesso nella stesura del manoscritto) creerebbe una perfetta linearità in tutto il percorso editoriale, evitando successive correzioni (e dunque di inserimento di eventuali errori) da parte dei revisori, dei redattori e degli indicizzatori. In particolare, definirebbe con chiarezza il settore di pertinenza del lavoro (che l'autore conosce meglio di chiunque altro), faciliterebbe notevolmente il lavoro dell'indicizzatore (che potrebbe riutilizzare le parole chiave senza essere chiamato ad operare interpretazioni o selezioni di termini e concetti) e permetterebbe un recupero più immediato dei contenuti concettuali del lavoro da parte degli utenti; in breve, l'utilizzo dei termini MeSH potrebbe essere inteso quale sinonimo di univocità e qualità.

Impiego dei MeSH negli articoli dei ricercatori ISS

Il Settore Attività Editoriali dell'ISS costituisce un punto di incontro privilegiato tra fasi, competenze ed esperienze della letteratura scientifica prodotta dai ricercatori dell'ISS che vengono puntualmente registrati in una base di dati bibliografica interna.

I termini da indicizzare, in italiano e inglese, inseriti nella base di dati sono sia quelli "pre-pubblicazione" ovvero relativi alla registrazione preliminare dei lavori (assegnati dagli autori stessi), sia quelli estratti dai lavori pubblicati, che potrebbero non coincidere con i primi, ma aver subito variazioni durante le diverse fasi di lavorazione del manoscritto.

Nella registrazione delle parole chiave nella base di dati, il sistema prevede che il termine da inserire possa essere automaticamente ricercato nel corpus dei descrittori MeSH originali e anche di quelli tradotti in italiano; nel caso in cui il termine coincida con un descrittore del *thesaurus* esso viene identificato dal sistema come "MeSH" (Figura 1).

The screenshot shows a Microsoft Access window titled "Microsoft Access - [Articoli]". The main form is titled "Lavoro scientifico" and contains the following data:

- Chiave:** 18570
- Richiedente:** [M]PI
- Autorizzato:** []
- Publicato:** []
- Codice:** DOI 10.1016/j.vetpar.2005.05.038
- Tipo:** AR STD
- Anno:** 2005
- Titolo:** Evaluation of the infectivity of Trichinella papuae and Trichinella zimbabwensis for equatorial freshwater fishes
- Volume:** 132
- Fascicolo:** 1-2
- Pagine:** 113-114
- Totale Pagine:** []
- Paese:** NL
- Lingua:** ENG
- Per, Mono...:** Veterinary parasitology
- Congr, Corso:** []
- Buttons:** FULL TEXT, Tipo fascicolo, Spoglio, Annulla

Below the form, there is a tabbed interface with tabs for "Autori", "Impact Factor", "Keywords", "Abstract", "Contenitori", "Altro Supporto", "Note", "Autorizzazione", and "Citazione". The "Keywords" tab is active, showing a table of MeSH terms:

| Lingua | Descrizione | Tesauro |
|--------|-------------------|---------|
| ITA | Infezione | MESH |
| ITA | Epidemiologia | MESH |
| ENG | Trichinella | MESH |
| ENG | Freshwater fishes | |

At the bottom right of the form, there are buttons for "Duplica", "Salva", and "Elimina". The status bar at the bottom indicates "Records: 1 di 1 (Filtrati)".

Figura 1. Esempio di record bibliografico, con parole chiave, dell'archivio delle pubblicazioni ISS

Al fine di offrire la più ampia base possibile di chiavi di ricerca, non è stata effettuata una normalizzazione dei termini non corrispondenti al MeSH. Questa scelta ha dunque comportato la coesistenza sia di descrittori del *thesaurus* MeSH che di parole chiave di altra derivazione trascritte fedelmente dai lavori indicizzati.

Per orientare i ricercatori ISS verso un impiego sistematico dei descrittori MeSH, anche in considerazione del fatto che la maggior parte della letteratura prodotta rientra nelle aree tematiche di pertinenza MeSH, è stata realizzata una applicazione informatica che consente agli utenti la selezione automatica dei termini da assegnare ai vari lavori in produzione. Ciò avviene tramite la compilazione di un modulo di registrazione online che, per facilitare l'impiego dei MeSH, permette agli autori di ricercare in modo automatico le parole chiave nel *thesaurus* MeSH online (inglese e italiano). I termini non identificati come MeSH vengono invece digitati a cura degli autori. È ovviamente prevista la possibilità di un insieme di parole chiave di cui alcune MeSH e altre non MeSH che potrebbero essere comunque inserite ad integrazione (Figura 2).

4 - Parole Chiave

| Termine MeSH | Lingua (It/Eng) | Parola chiave | Cancella |
|--------------|-----------------|----------------------------|----------|
| D017209 | ENG | Apoptosis | ✘ |
| D017209 | ITA | Apoptosi | ✘ |
| | ITA | Separazione di enantiomeri | ✘ |

Aggiungi parola chiave MeSH:

Termine MeSH : [Ricerca MeSH](#)
 Codice MeSH :

Aggiungi parola chiave diversa da MeSH:

Lingua: Termine:

Figura 2. Schermata di inserimento delle parole chiave (MeSH e non) nel modulo di registrazione online delle pubblicazioni prodotte dai ricercatori ISS

Indagine sull'archivio bibliografico ISS

Metodologia

Ai fini di quantificare la presenza dei descrittori MeSH tra tutte le parole chiave dei lavori pubblicati dai ricercatori ISS, sono state innanzitutto individuate le riviste sulle quali questi hanno maggiormente pubblicato durante il 2005. La pubblicazione sulle riviste edite dall'ISS è stata oggetto di una valutazione a parte.

La ricerca è stata effettuata nell'archivio bibliografico ISS che contiene le registrazioni di tutti i lavori pubblicati dai ricercatori dell'ISS, in media circa 1.600 per anno. L'archivio contiene attualmente circa 15.000 record tra articoli di periodico, monografie, capitoli di monografie, atti di congresso, rapporti tecnici ecc.

In particolare, il campione dei lavori censiti ha interessato 12 riviste, 2 pubblicate in Italia e 10 all'estero, selezionate in ordine decrescente di numero di articoli di autori ISS ivi ospitati, da un massimo di 17 articoli per rivista ad un minimo di 6, per un totale di 107 lavori.

Tutte le riviste censite sulla base di questo criterio rientrano nelle aree disciplinari riferite alla biomedicina; pertanto è giustificata l'aspettativa di rilevare articoli indicizzati con descrittori MeSH.

Risultati

Di ciascuna delle 12 riviste selezionate in base ai criteri descritti, sono stati rilevati il numero totale degli articoli pubblicati dai ricercatori ISS, il numero degli articoli con parole chiave (MeSH e non), il numero degli articoli contenenti termini MeSH e le relative percentuali.

Dai dati riportati nella Tabella 1 emerge che, in generale, gli articoli contenenti parole chiave sono il 77,6% di quelli analizzati. Quelli contenenti termini MeSH costituiscono il 66,4%, sono cioè pari a 71 su 107 articoli censiti. In particolare, le più alte percentuali di articoli che attestano la presenza di MeSH sono quelle riscontrate in 3 riviste pubblicate all'estero (*Vet Parasitol*, *J Immunol* e *Vaccine*), rispettivamente con il 100% e l'85,7% (a pari merito), mentre il valore più basso (10%) è stato riscontrato per l'unica rivista edita in Italia (*Igiene e Sanità Pubblica*).

Tabella 1. Risultati dell'analisi effettuata sugli articoli pubblicati dai ricercatori ISS e inseriti nella base di dati bibliografica (anno 2005)

| Titolo rivista | Totale articoli ISS | Articoli con parole chiave* | % articoli con parole chiave* su totale articoli | Articoli con MeSH | % articoli con MeSH su totale articoli |
|-----------------------------|---------------------|-----------------------------|--|-------------------|--|
| Microchem J | 17 | 17 | 100,0 | 12 | 70,6 |
| J Clin Micr | 13 | 4 | 30,8 | 4 | 30,8 |
| Biologi Italiani | 11 | 9 | 81,8 | 9 | 81,8 |
| Ig Sanità Pubbl | 10 | 1 | 10,0 | 1 | 10,0 |
| Vet Parasitol | 9 | 9 | 100,0 | 9 | 100,0 |
| J Infec Dis | 8 | 7 | 87,5 | 6 | 75,0 |
| J Immunol | 7 | 6 | 85,7 | 6 | 85,7 |
| J Med Virol | 7 | 7 | 100,0 | 5 | 71,4 |
| Vaccine | 7 | 7 | 100,0 | 6 | 85,7 |
| Antimicrob Agents Chemother | 6 | 4 | 66,7 | 4 | 66,7 |
| AIDS | 6 | 6 | 100,0 | 4 | 66,7 |
| J Pharm Biomed Anal | 6 | 6 | 100,0 | 5 | 83,3 |
| Totale | 107 | 83 | 77,6 | 71 | 66,4 |

* Mesh e non

Tra le serie di pubblicazioni istituzionali, sono stati selezionati due periodici: a) gli *Annali dell'Istituto Superiore di Sanità* (rivista scientifica trimestrale *peer-reviewed*) e b) il *Notiziario dell'Istituto Superiore di Sanità* (bollettino mensile che contiene informazioni correnti su attività istituzionali).

Dai dati relativi a entrambe le pubblicazioni, anch'essi ricavati dall'archivio bibliografico interno e verificati sulle copie stampate dei singoli fascicoli, è emerso che il 77% di articoli pubblicati sugli *Annali* utilizza termini MeSH, il 23% utilizza parole chiave non-MeSH. Per il *Notiziario* si registrano dati leggermente differenti: la percentuale di termini MeSH scende al 43% e quella delle parole chiave non-MeSH sale al 57%. Tale leggera differenza nelle percentuali è facilmente riconducibile alla diversa natura e finalità delle due testate.

L'indagine condotta sul campione dell'archivio bibliografico ISS ha fornito altri dati rilevanti: la media di termini MeSH utilizzati per articolo è di 1,44; mentre il rapporto MeSH vs non-MeSH è pari al 20% in riviste italiane e 89% in riviste straniere. Quest'ultimo dato induce a credere che esista una maggiore attenzione (degli autori, revisori, redattori) verso l'uso di una terminologia standardizzata in articoli pubblicati su riviste straniere piuttosto che su quelle italiane.

Riflessioni conclusive

Dal breve resoconto sulle diverse fasi del processo editoriale e dai risultati emersi dallo studio sull'impiego dei MeSH nella produzione editoriale dei ricercatori ISS, è possibile ricavare alcune considerazioni finali e un auspicio.

Nonostante gli evidenti vantaggi dell'utilizzo dei termini MeSH (contribuire a migliorare la qualità dei prodotti realizzati, a facilitare il recupero di informazioni pertinenti e ad agevolare l'intero processo editoriale), questi di fatto sono ancora troppo poco diffusi e scarsa è in generale la consapevolezza della loro rilevanza.

Questo lavoro ha dimostrato come l'“educazione” di tutti gli attori del processo editoriale e lo sfruttamento intelligente delle moderne applicazioni informatiche possano notevolmente contribuire in tal senso.

In definitiva, se standardizzare (tecniche, sistemi e linguaggio) è ormai obiettivo di primaria importanza in tanti settori della società odierna per una migliore comunicazione e un più rapido e corretto trasferimento delle informazioni, l'impiego dei termini MeSH nella produzione editoriale biomedica diventa sempre più auspicabile e opportuno.

CONOSCENZA E WEB SEMANTICO NELL'AREA SALUTE, SANITÀ E PREVENZIONE

Paola Capitani
Gruppo di lavoro Semantica e terminologia nei portali
paolacapitani@libero.it

Premessa

L'Europa è sempre più caratterizzata da ricerche e operatori impegnati in un panorama multilingue e globalizzato: da qui la necessità di lessici comuni per muoversi nel vasto e differenziato panorama della conoscenza e della comunicazione. Internet e le nuove tecnologie sono al servizio di una comunità eterogenea e multiculturale, diversificata, che richiede strumenti di controllo terminologico, di intermediazione culturale, siti web in grado di far interagire terminologie diverse tra loro.

Ecco quindi il ricorso a web semantici, gruppi in linea, comunità in rete, che usano gli stessi concetti e gli stessi codici linguistici per facilitare l'accesso e la ricerca, ma soprattutto per collaborare ad una condivisione concettuale e ad un'integrazione tra popoli e culture e in questo senso ben si colloca la traduzione italiana dei *Medical Subject Headings*, realizzata dal Settore Documentazione dell'Istituto Superiore di Sanità.

Nel 1999 il Gruppo Interdisciplinare per la Salute e la Terminologia (GIST) nasceva su iniziativa di alcuni operatori dell'informazione e della conoscenza, a seguito del Seminario di Perugia (22-23 aprile 1999), per creare sinergie nelle problematiche comuni di ricerca nella terminologia in ambito biomedico e sociosanitario.

Il GIST era nato per garantire l'accessibilità universale alle informazioni, quale condizione fondamentale per lo sviluppo della ricerca scientifica, la prevenzione, la cura, l'educazione alla salute dei cittadini e la verifica della qualità delle prestazioni erogate, favorire la cultura della promozione della salute, utilizzando al meglio le nuove tecnologie della comunicazione (Internet, media, ecc.).

L'obiettivo primario era favorire una migliore comunicazione tra tutti i soggetti a vario titolo coinvolti nella promozione della salute: decisori e gestori delle politiche sanitarie, operatori sociali e sanitari, associazioni di volontariato, scuola, mondo del lavoro, pazienti e cittadini. Occorre, infatti, orientare e indirizzare l'utenza verso l'informazione scientifica migliorando l'accessibilità intra e interregionale dei servizi e delle strutture che erogano le prestazioni.

Tra le prime attività che il Gruppo si è impegnato a fare è stata quella di verificare la terminologia esistente nel Piano Sanitario Nazionale, per utilizzare una fonte riconosciuta e utilizzata sul territorio da tutti gli operatori, e verificarla con gli strumenti terminologici prodotti e utilizzati dai diversi enti. È risultata, dalla prima analisi, una difformità terminologica tra gli accessi utilizzati nel PSN, una differente interpretazione a livello degli strumenti di controllo terminologico e una disomogenea traduzione dei termini inglesi di riferimento. Molto spesso si parte da fonti esistenti in lingua straniera adattate al contesto italiano. Seguono alcuni esempi significativi:

- patologia e malattia; abitudini e comportamenti; terapia e cura; tumore-cancro-carcinoma-neoplasia; sistema e apparato; alimentazione e nutrizione;
- diverse traduzioni in lingua italiana per *hazard* e *risk*;

- l’annoso problema fra i termini educazione alla salute e educazione sanitaria.

Nel frattempo era stata organizzata a Pisa la Conferenza nazionale “I nuovi orizzonti dell’educazione sanitaria e della promozione della salute: politiche e reti per la salute” (Pisa 8-10 giugno 2000), che aveva puntualizzato alcuni interessi nel settore, in particolare nell’area relativa alla terminologia e al suo uso tra gli addetti ai lavori e gli utenti.

Purtroppo dopo circa due anni di attività e numerose riunioni il gruppo si è sciolto; tutto quanto era stato effettuato è andato quindi disperso. Esperienze differenti, proposte, risorse, progetti, sono stati vanificati e il lavoro di due anni perso per strada, come spesso accade nelle situazioni progettuali e di ricerca che costantemente vediamo intorno a noi.

Knowledge management e web semantico

Il Gruppo web semantico (www.indire.it/websemantico), nato nel 2000, ha continuato il lavoro intrapreso dal GIST, creando ulteriori collegamenti con università e centri di ricerca impegnati nel settore della conoscenza e della terminologia per razionalizzare risorse ed esperienze e facilitare l’accesso agli utenti, soprattutto se non esperti nel settore. Nel Gruppo sono anche presenti alcuni nominativi del precedente GIST: per cui non tutte le esperienze sono andate perse, ma reinvestite nei nuovi rapporti e progetti.

Il Gruppo è formato da esperti nei diversi settori trattati, costituitosi volontariamente nel gennaio 2000, svolge attività di ricerca nell’ambito della terminologia e della traduzione nei siti web. Tra gli obiettivi prioritari: lo studio dei vocabolari controllati, la gestione della conoscenza, la traduzione univoca della terminologia, la formazione degli addetti ai lavori e la facilità del recupero per gli utenti, specialmente quelli non specializzati.

Obiettivi e destinatari

Gestione della terminologia e della traduzione per il recupero delle informazioni in Internet, riducendo i problemi di comprensione linguistica e limitando inesatte traduzioni. Fra i partecipanti al gruppo un particolare contributo è venuto in tal senso da parte di rappresentanti del Progetto Minerva finalizzato alle linee guida dei siti web, oltre alla partecipazione dell’Università di Bologna, Scuola superiore traduttori e interpreti, che ha supportato attivamente i progetti svolti fino ad ora e le traduzioni necessarie.

Ricercatori, docenti, bibliotecari, esperti dell’informazione interessati al recupero delle informazioni e ai problemi inerenti la traduzione e la terminologia attraverso l’uso del web semantico.

Aree tematiche

Di seguito si elencano le aree tematiche e fra parentesi i siti web degli enti/istituti di riferimento per la tematica e il totale dei termini disponibili nei glossari sul *forum* del sito o dei siti indicati:

- biblioteche scolastiche (www.indire.it: 100 termini);
- diritto (www.ITTIG/CNR.it);
- *e-learning* (sito del gruppo europeo www.filnet.org: 200 termini);
- economia (www.liuc.it);

- educazione/formazione (www.aifonline.it, www.unifi.it, www.unipr.it, www.indire.it/bibl, www.istitutodeglinnocenti.it, www.personae.it, www.firenzetecnologia.it, www.taccone.net: 2000 termini);
- informazione/comunicazione (www.repubblica.it);
- *library and information science* (www.e-lis.org: 2000 termini);
- moda (www.polimoda.com);
- sanità (www.iss.it/site/Mesh);
- traduzione/terminologia (www.sslmit.unibo.it, www.unicei.com, www.liuc.it: 200 termini).

Inoltre è attivo un collegamento con la Biblioteca Nazionale Centrale di Firenze, dipartimento che si occupa dell'aggiornamento del Soggettario della BNI, revisione terminologia e traduzioni usate.

Sono allo studio moduli base per gli indicizzatori dei servizi bibliotecari (aggiornamento/formazione in presenza e a distanza), su: standard di trattamento (terminologia/traduzione); criteri di indicizzazione; web semantico.

Metodologie

I lavori del Gruppo si svolgono per posta elettronica, attraverso il forum del sito e con momenti di confronto in presenza (seminari tematici e tavole rotonde annuali).

Nel forum del sito i contributi delle precedenti tavole rotonde e dei lavori in corso oltre ai contributi e documenti per confronti ed esperienze.

Progetti in corso

I Progetti in corso sono:

- revisione di un glossario bilingue (inglese/italiano) tratto dagli standard vigenti ANSI, BS, ISO, UNI per una condivisione di una terminologia base condivisa nel settore dei *thesauri* e della terminologia controllata (con Piero Cavalieri, www.liuc.it)
- glossario bilingue di termini relativi al settore formazione (con il Gruppo Qualità di AIF, www.aifonline.it)
- glossario bilingue di termini relativi all'educazione, formazione ed *e-learning* comprendendo anche il KM e la Qualità (con il gruppo europeo Filterproject, www.filternetwork.org)
- promozione di un gruppo di lavoro su terminologia, qualità e gestione della conoscenza (con la Biblioteca Nazionale Centrale)

Non appena definiti tempi e modalità dei progetti saranno comunicati nel forum del sito: www.indire.t/wesemantico, veicolo di comunicazione e di accesso ancora scarsamente utilizzato, anche se continuamente consultato dai ricercatori. In tal senso occorre perseguire un incentivo agli utenti dei siti web affinché non siano solo fruitori, ma anche, e soprattutto, degli effettivi utenti. Grazie alle esperienze, ai contributi, ai commenti un sito può effettivamente chiamarsi tale, altrimenti è solo una vetrina, una televisione, che non svolge a pieno il suo portato comunicativo e formativo.

Conclusioni

Oggi i progetti terminologici sono moltissimi e per ovviare a doppioni e razionalizzare costi e tempi, ma soprattutto per aiutare l'utente a usare al meglio l'informazione è obbligatorio riuscire a mettere in comune tutte le risorse esistenti facilitando l'approccio e dando veramente un contributo a quella così detta "società dell'informazione e della comunicazione" che rischia sempre più di creare paurosi pantani e infidi meandri a quanti si avventurano senza bussola e senza conoscenza. Soprattutto occorre seguire tre obiettivi basilari:

1. formazione e aggiornamento degli operatori;
2. formazione degli utenti;
3. continua revisione di tecniche e procedure per tenere il passo con le esigenze di ricerche.

La comunicazione e la conoscenza sono in continua evoluzione in quanto seguono l'informazione e la scienza e promuovono e documentano la ricerca. Ma non possiamo avere conoscenza senza una adeguata terminologia che sia soprattutto un veicolo di accesso e di diffusione e non solo, come spesso accade, uno strumento per gli addetti ai lavori. Qui sta il vero valore aggiunto che la conoscenza porta sul panorama internazionale e sull'eccesso distribuito, dove l'utente è o dovrebbe essere il destinatario privilegiato dell'informazione e della conoscenza.

Bibliografia

- Azzariti F, Mazzon P. *Il valore della conoscenza: teoria e pratica del knowledge management prossimo e venturo*. Milano: Etas; 2005.
- Batini F, Fontana A. *Comunità di apprendimento*. Rapallo: Editrice Zona; 2004.
- Bonani GP. *La sfida del capitale intellettuale. Principi e strumenti di knowledge management per organizzazioni intelligenti*. Milano: Franco Angeli; 2002.
- Calvani A, Rotta M. *Fare formazione in rete*. Trento: Erickson; 2002.
- Capitani P. *Il knowledge management. Strumento di orientamento e formazione per la scuola, l'università, la ricerca, il pubblico impiego*. Milano: Franco Angeli, Collana AIF; 2006.
- Capitani P. Terminologia controllata. *Biblioteche Oggi* 2005;5:41-49.
- Capitani P. Gestione della conoscenza e formazione. In: Valente A, Luzi D (Ed.). *Partecipare la scienza*. Roma: Biblink editori; 2004.
- Capitani P. E-learning e terminologia. *Biblioteche Oggi* 2004;7:63-4.
- Capitani P. *Il servizio informativo: iter di gestione e proposte di formazione*. Milano: Editrice Bibliografica; 2003.
- Capitani P. Gruppo semantica e terminologia nei portali - Gruppo interdisciplinare e multilingue. In: *Atti del Convegno Biblioteche digitali per la ricerca e la didattica: esperienze e prospettive. Università di Parma, 22 novembre 2003*.
- Capitani P. Educare all'informazione: l'esigenza di creare un costume documentario a partire dalla scuola. *Biblioteche Oggi* 2001;1:24-9.
- Capitani P (Ed.). Il controllo terminologico delle risorse elettroniche in rete. Tavola rotonda, Firenze 27 gennaio 2000. Disponibile all'indirizzo: <http://eprints.unifi.it/archive/00000330/01/F0009-Capitani.pdf>; ultima consultazione 22/12/06.
- Capitani P. Tavole rotonde del gruppo web semantico, Liuc Papers n. 183, Serie Materiali bibliografici 6, suppl. a dicembre 2005. Disponibile all'indirizzo www.biblio.liuc.it/liucpap/pdf/183.pdf; ultima consultazione 22/12/06.

Progetto Minerva. Manuale per la qualità dei siti web pubblici culturali. Roma: Ministero per i beni e le attività culturali; 2005.

Rotondi M. *Formazione outdoor: apprendere dall'esperienza. Teorie, modelli, tecniche, best practices.* Milano: FrancoAngeli; 2004.

Ruta Cataldo D. *Organizzare il knowledge management.* Padova: EGEA; 2002.

Sardelli A. *Dalla certificazione di qualità alla qualità totale.* Milano: Editrice Bibliografica; 2002.

AGGIORNAMENTO DELLE MODALITÀ DI RICERCA IN PUBMED

Scilla Pizzarelli, Rosaria Rosanna Cammarano
Settore Documentazione– Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali
Istituto Superiore di Sanità, Roma
scilla.pizzarelli@iss.it, rosaria.cammarano@iss.it

Introduzione

PubMed è l'interfaccia web del MEDLINE, l'archivio più noto e consultato del sistema MEDLARS (*MEDical Literature Analysis Retrieval System*), un insieme di basi di dati bibliografici e banche dati fattuali di interesse biomedico in senso lato, prodotto nel 1971 dalla *National Library of Medicine* (NLM) di Bethesda (Stati Uniti).

Fino a pochi anni or sono il MEDLINE, e con esso tutti gli altri archivi del sistema, era accessibile soltanto agli utenti provvisti di *password*, fornita dal Settore Documentazione dell'Istituto Superiore di Sanità nella sua veste di centro nazionale di riferimento per il MEDLARS, previo pagamento di un canone di abbonamento annuale.

Da quando il complesso di queste risorse è diventato consultabile gratuitamente sul *world wide web* (1), alla fine degli anni '90, il suo bacino di utenza, composto prevalentemente da enti e istituzioni del Servizio Sanitario Nazionale, si è progressivamente allargato fino a comprendere una schiera sempre più nutrita di biblioteche. Queste ultime, infatti, con l'ingresso delle tecnologie digitali, hanno dovuto riconfigurare il servizio di *reference*, basato non più esclusivamente sulle fonti informative in formato cartaceo, ma soprattutto sulle risorse elettroniche, locali e remote.

Ed è proprio alle biblioteche che si rivolge, in particolare, il presente lavoro, nel quale vengono illustrate le nuove capacità di ricerca del PubMed.

Descrizione del PubMed

Prodotto dal *National Center for Biotechnology Information* (NCBI) presso l'NLM, il PubMed è stato lanciato ufficialmente nel 1997, quando il governo degli Stati Uniti ha annunciato di voler distribuire gratuitamente su Internet il MEDLINE, che rappresenta il nucleo originario intorno al quale si è sviluppato il PubMed (2).

È accessibile all'indirizzo <http://www.pubmed.gov> tramite il sistema Entrez, implementato sempre dall'NCBI, che comprende una serie di basi di dati di genetica e biologia molecolare (3). Come ogni altro archivio informativo in formato elettronico, è strutturato in record, ognuno corrispondente ad una citazione bibliografica, suddivisi in campi utilizzabili come punti di accesso per la ricerca, in ciascuno dei quali è registrata una diversa categoria di informazioni (Figura 1).

PubMed contiene quasi 17 milioni di riferimenti bibliografici provenienti dallo spoglio e indicizzazione di circa 5000 riviste scientifiche, edite negli Stati Uniti e in altri 70 paesi, a partire dal 1950, è aggiornato quotidianamente e abbraccia tutti i settori delle discipline mediche, infermieristiche, odontoiatriche, veterinarie, di organizzazione sanitaria e scienze precliniche.

```

PMID- 17058533
OWN - NLM
STAT- MEDLINE
DA - 20061023
DCOM- 20061102
PUBM- Print
IS - 0350-6134 (Print)
VI - 30
IP - 3
DP - 2006 Sep
TI - Influence of passive smoking on basic anthropometric characteristics and
    respiratory function in young athletes.
PG - 615-9
AB - The primary objective of this study is to investigate the maintenance
    difference in basic anthropometric characteristics and to outline the
    dynamics of respiratory function change in youngsters athletes exposed to
    passive smoking (PS) and athletes not exposed to passive smoking in their
    families (NPS). High and weight were determined as basis anthropometric
    characteristics. Measured parameters for respiratory function were vital
    capacity (VC), forced expiratory volume in the first second (FEV1),
    maximum expiratory flow (PEF), forced expiratory flow at 50% forced vital
    capacity (MEF 50) and forced expiratory flow at 25% forced vital capacity
    (MEF 25). Significant statistical differences in separate spirometric
    variable were found in three variables (FEV1, MEF50, and MEF25) for group
    older youngsters. Analysis of variance showed statistical differences
    between athletes unexposed to passive smoking (NPS) and athletes exposed
    to passive smoking (PS) in even four spirometric variables (VC, FEV1,
    MEF50 and MEF25).
AD - Department of Clinical Microbiology, University Hospital Split, Split,
    Croatia. ivanagoicbar@net.hr
FAU - Goic-Barisic, Ivana
AU - Goic-Barisic I
FAU - Bradaric, Anteo
AU - Bradaric A
FAU - Erceg, Marko
AU - Erceg M
FAU - Barisic, Igor
AU - Barisic I
FAU - Foretic, Nikola
AU - Foretic N
FAU - Pavlov, Neven
AU - Pavlov N
FAU - Tocilj, Jadranka
AU - Tocilj J
LA - eng
PT - Journal Article
PL - Croatia
TA - Coll Antropol
JT - Collegium antropologicum.
JID - 8003354
SB - IM
MH - Adolescent
MH - Analysis of Variance
MH - *Anthropometry
MH - Child
MH - Humans
MH - Lung/*drug effects
MH - Lung Volume Measurements
MH - *Sports
MH - Tobacco Smoke Pollution/*adverse effects
EDAT- 2006/10/25 09:00
MHDA- 2006/11/03 09:00
PST - ppublish
SO - Coll Antropol. 2006 Sep;30(3):615-9.

```

Figura 1. Esempio di record bibliografico in formato MEDLINE, aggiornato al 07/12/06

La sua componente principale è costituita dal MEDLINE, che raccoglie oltre 14 milioni di citazioni bibliografiche contenenti i termini di indicizzazione, ovverosia i descrittori estratti dal *thesaurus* MeSH (*Medical Subject Headings*), che descrivono il contenuto concettuale degli articoli (4).

Negli ultimi anni è stato riversato nel PubMed anche il contenuto dell'OLDMEDLINE, un complesso di oltre 1.700.000 riferimenti bibliografici alla letteratura pubblicata dal 1950 al 1965, anch'essi dotati di descrittori MeSH, grazie ad un processo di conversione delle *keyword* (parole chiave) originarie nei termini equivalenti del *thesaurus* MeSH, di cui è prevista la conclusione nel corso del 2007.

Oltre alle citazioni indicizzate per il MEDLINE, contraddistinte dall'etichetta [PubMed – indexed for MEDLINE], e ai riferimenti provenienti dall'OLDMEDLINE, etichettati [PubMed – OLDMEDLINE for Pre1966], il PubMed comprende diversi tipi di record privi di completa indicizzazione e, pertanto, sprovvisti dei termini MESH. Si tratta delle citazioni di articoli di recente immissione, ancora in corso di lavorazione [PubMed - *in process*], citazioni di articoli trasmesse al PubMed direttamente dagli editori [PubMed - *as supplied by publisher*] ed, infine, citazioni di articoli che non rientrano nel campo di copertura del MEDLINE o pubblicati su fascicoli di riviste apparsi prima della data di inserimento delle riviste stesse nel PubMed [PubMed] (5).

Oltre a contenere le diverse tipologie di citazioni bibliografiche fin qui esaminate, il PubMed fornisce l'accesso alle basi di dati Entrez e ad un cospicuo numero di risorse correlate, consentendo, inoltre, di usufruire di una serie di servizi accessori che facilitano e completano la ricerca, alcuni dei quali sono oggetto del presente lavoro.

Le nuove modalità della ricerca per autore

Com'è noto agli utenti abituali del PubMed, la ricerca per autore si effettua inserendo nella casella di ricerca (*query box*) il cognome e la/e iniziale/i del nome dell'autore, racchiusi tra virgolette, seguiti dal qualificatore di ricerca (*search tag*) [au] (*author*). Se, ad esempio, si vogliono recuperare le pubblicazioni di Olof Werner bisogna digitare: "Werner O" [au] (6).

L'utilizzo del qualificatore di ricerca consente di limitare l'interrogazione al campo degli autori, escludendo dall'indagine tutti gli altri campi che compongono il record bibliografico.

L'impiego delle virgolette, invece, è necessario per dare istruzioni al sistema di non applicare il troncamento automatico dopo l'iniziale del nome, troncamento che includerebbe nella ricerca tutte le possibili varianti del nome (nomi composti, come Werner OA, Werner OB, ecc., o nomi contenenti specificazioni quali jr., sr., ecc., come, ad esempio, Werner Ojr).

La ricerca per autore effettuata attraverso le iniziali del nome ha sempre lasciato aperto il problema dei casi di omonimia, peraltro assai frequente, imponendo all'utente il ricorso a stratagemmi vari per identificare l'autore di interesse (quali, per esempio, incrociare i risultati ottenuti con l'affiliazione professionale dell'autore oppure con uno o più soggetti che descrivono il suo ambito di attività).

Oggi questo problema è stato parzialmente risolto con l'introduzione del qualificatore [fau] (*full author name*), che consente di impostare la ricerca digitando il nome in forma estesa anziché abbreviata. Così, tornando al nostro caso, se si inserisce nella casella di ricerca: Werner Olof [fau], si ha la certezza di recuperare solo dati pertinenti, eliminando le pubblicazioni di autori dal cognome identico e con le stesse iniziali del nome (Figura 2).

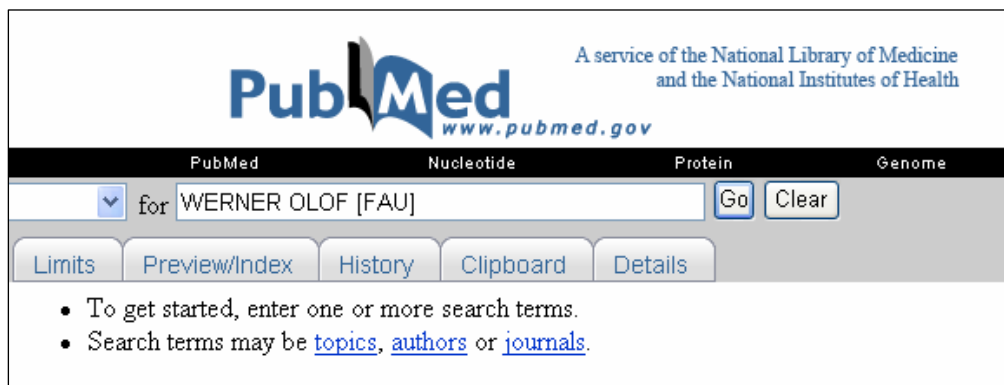


Figura 2. Esempio di ricerca per autore con il qualificatore [fau]. Aggiornato al 07/12/06

Si tratta, però, come dicevamo, di una risoluzione parziale in quanto l'impiego di questa procedura comporta due grossi inconvenienti: l'esclusione dei record inseriti prima del 2002, che contengono soltanto le iniziali dei nomi degli autori, e la perdita di tutte le citazioni di articoli in cui nell'edizione a stampa il nome dell'autore non compare per esteso.

Del tutto funzionali, e operanti sull'intero archivio, sono, invece, i qualificatori di ricerca [1au] (*first author*) e [lastau] (*last author*), anch'essi di recente istituzione, che consentono di distinguere il contributo prestato, rispettivamente, dal primo e dall'ultimo autore di un'opera scritta in collaborazione da più autori. Così, ad esempio, digitando: "Werner O" [1au] si recupereranno tutti i documenti prodotti da Werner in qualità di primo autore, viceversa, inserendo: "Werner O" [lastau] si otterranno i lavori ai quali ha contribuito in veste di ultimo autore.

La funzione "Limits"

Richiamabile dalla barra orizzontale posta sotto la maschera di interrogazione, questa funzione è stata recentemente potenziata e arricchita nei contenuti, consentendo anche all'utente meno esperto di costruire strategie di ricerca mirate, mediante l'applicazione di filtri tematici o cronologici che permettono di limitare il campo di indagine (7).

Una volta entrati nella funzione, una serie di menu a tendina, pulsanti e caselle selezionabili consentono di attivare i seguenti filtri:

1. Autore

(Search by Author)

Cliccando sul pulsante "Add Author" e inserendo il cognome di un autore si apre un menu a tendina che permette di accedere all'indice alfabetico degli autori presenti nel PubMed a partire dall'autore inserito, dal quale si può selezionare direttamente quello di proprio interesse, grazie alla nuova funzione di completamento automatico delle parole.

Se si vogliono esaminare i documenti relativi a più autori basta cliccare sul tasto "Add Another Author" che, con le stesse modalità appena indicate, consente di aggiungere ulteriori autori da combinare logicamente tra di loro con l'operatore OR o AND attraverso le opzioni "Any of these" o "All these authors" (applicata dal sistema per default).

2. *Rivista**(Search by Journal)*

La limitazione per rivista, attivabile mediante il tasto “*Add Journal*”, funziona in maniera analoga, con l’apertura di una maschera di interrogazione ad hoc che consente di selezionare il titolo desiderato dall’indice delle riviste censite dal PubMed. Anche qui viene offerta l’opportunità di specificare altre riviste, cliccando sul tasto “*Add Another Journal*”, ma, naturalmente, in questo caso l’unica combinazione logica prevista è quella effettuata con l’operatore OR.

3. *Testo completo, testo completo gratuito e abstract**(Full Text, Free Full Text and Abstract)*

Questa sezione consente di restringere la ricerca alle citazioni che contengono il collegamento al testo completo in formato elettronico, scaricabile a pagamento (*Full Text*) o a titolo gratuito (*Free Full text*), nonché di limitare l’interrogazione alle sole citazioni provviste di *abstract* (presenti nel 76% dei record inseriti a partire dal 1975).

4. *Data di pubblicazione**(Published in the last)*

È possibile limitare l’interrogazione agli articoli pubblicati in un determinato arco di tempo, selezionando un intervallo temporale prestabilito dal sistema o digitando l’intervallo di data richiesto, espresso nel formato Anno/Mese/Giorno (l’anno va inserito sempre, mentre il mese e il giorno sono facoltativi).

5. *Data di inserimento**(Added to PubMed in the last)*

Quest’opzione consente di restringere la ricerca alle citazioni inserite nel PubMed in un arco temporale predefinito o specificato dall’utente.

È bene precisare che la data di pubblicazione dell’articolo non sempre coincide con la data di inserimento del record corrispondente nell’archivio, che dipende dalla data di sottomissione del documento al PubMed da parte degli editori e dai tempi di lavorazione del record da parte della NLM.

6. *Uomo o Animale**(Humans or Animals)*

Attraverso questo filtro si può circoscrivere l’ambito di ricerca agli studi compiuti sull’uomo (clinici) o sull’animale (sperimentali).

Applicando una di queste opzioni l’interrogazione viene limitata alle sole citazioni MEDLINE. Infatti i record che non sono stati sottoposti al processo di indicizzazione, ovvero sia le citazioni “*in process*” (in lavorazione) e quelle recanti la dicitura “*as supplied by publisher*” (inviate dagli editori), saranno automaticamente escluse in quanto non contengono questa tipologia di informazioni.

7. *Genere**(Gender)*

Quest’opzione offre l’opportunità di selezionare il sesso (maschile o femminile) del gruppo oggetto dello studio, restringendo contemporaneamente l’interrogazione agli studi clinici.

Per le stesse motivazioni indicate in precedenza la ricerca viene, altresì, circoscritta alle citazioni MEDLINE.

8. *Lingua*
(*Languages*)

È possibile effettuare una limitazione per lingua di pubblicazione, selezionabile dall'apposito menu a tendina, che riporta inizialmente la lista alfabetica delle lingue di pubblicazione più richieste, e, quindi, sotto la voce "*More Languages*", l'elenco di tutte le altre lingue considerate nell'archivio (ca. 60 in totale).

Per scegliere più opzioni contemporaneamente basta cliccare con il mouse su ognuna di esse.

9. *Sottoinsiemi*
(*Subsets*)

Questo filtro permette di limitare la ricerca per sottoinsiemi di citazioni suddivisi per: gruppi di riviste (cliniche, dentistiche e infermieristiche); argomento (AIDS, bioetica, oncologia, storia della medicina, ecc.); banche dati (MEDLINE, OLDMEDLINE, e PubMed Central, archivio digitale di letteratura biomedica, contenente il testo completo gratuito degli articoli).

10. *Tipo di articolo*
(*Type of Article*)

La ricerca può essere limitata sulla base del tipo di materiale che l'articolo rappresenta, selezionabile dall'apposito menu a tendina dove sono riportati in un primo gruppo i tipi di pubblicazione maggiormente richiesti (sperimentazioni cliniche, editoriali, lettere all'editore, meta-analisi, linee guida per la pratica clinica, ecc.) e, in un secondo gruppo, tutti gli altri tipi di pubblicazione considerati nel PubMed (come bibliografie, biografie, *case reports*, ecc.).

11. *Gruppi di età*
(*Ages*)

Selezionando una o più fasce d'età del gruppo oggetto dello studio vengono automaticamente esclusi gli studi sperimentali e la ricerca viene limitata alle sole citazioni MEDLINE.

12. *Campi di ricerca*
(*Tag Terms*)

L'interrogazione può essere limitata ad un campo specifico del record bibliografico, selezionabile dall'apposito menu a tendina, dove sono elencati in ordine alfabetico tutti i campi del record ricercabili (autore, titolo, termine MeSH, ecc.).

Una volta applicati i filtri desiderati, per far eseguire l'interrogazione basta cliccare sul pulsante "*Go*", posto in fondo alla pagina.

Tutti i limiti impostati vengono mantenuti in memoria durante la stessa sessione di collegamento per cui, prima di passare all'esecuzione di una nuova ricerca, bisogna avere l'accortezza di disattivarli togliendo la spunta dalla casella che compare alla sinistra del pulsante "*Limits*" nella pagina dei risultati.

Questi filtri sono stati riportati nell'ordine in cui appaiono nella pagina dei limiti (Figura 3).

The image shows the PubMed search limits interface. At the top, it says 'Limit your search by any of the following criteria.' Below this are several filter sections, each with a 'CLEAR' button:

- Search by Author:** Includes an 'Add Author' button.
- Search by Journal:** Includes an 'Add Journal' button.
- Full Text, Free Full Text, and Abstracts:** Includes checkboxes for 'Links to full text', 'Links to free full text', and 'Abstracts'.
- Dates:** Includes dropdown menus for 'Published in the Last' and 'Added to PubMed in the Last'.
- Humans or Animals:** Includes checkboxes for 'Humans' and 'Animals'.
- Gender:** Includes checkboxes for 'Male' and 'Female'.
- Languages:** Includes checkboxes for English, French, German, Italian, Japanese, Russian, Spanish, and 'More Languages' (Afrikaans, Albanian).
- Subsets:** Includes 'Journal Groups' (Core clinical journals, Dental journals, Nursing journals) and 'Topics' (AIDS, Bioethics, Cancer, Complementary Medicine, History of Medicine).
- Type of Article:** Includes checkboxes for Clinical Trial, Editorial, Letter, Meta-Analysis, Practice Guideline, Randomized Controlled Trial, Review, and 'More Publication Types' (Addresses, Bibliography).
- Ages:** Includes checkboxes for All Infant: birth-23 months, All Child: 0-18 years, All Adult: 19+ years, Newborn: birth-1 month, Infant: 1-23 months, Preschool Child: 2-5 years, Child: 6-12 years, Adolescent: 13-18 years, Adult: 19-44 years, and Middle Aged: 45-64 years.
- Tag Terms:** Includes a dropdown for 'Default Tag' set to 'All Fields'.

At the bottom, there is a 'GO' button and a 'Clear All Limits' button.

Figura 3. La pagina dei limiti, in PubMed, aggiornata al 07/12/06

II Journals Database

Un'altra importante innovazione è stata apportata al *Journals Database*, la base di dati contenente informazioni catalografiche sulle riviste indicizzate dal PubMed e dalle altre basi di dati del sistema Entrez, accessibile dalla barra laterale della maschera di interrogazione sotto la voce "PubMed Services".

Oltre ai tradizionali canali di ricerca (titolo esteso o abbreviato della rivista, abbreviazione ISO, numero identificativo assegnato dalla NLM e ISSN), il *Journals Database* permette ora di effettuare l'interrogazione per soggetto, utilizzando come punto di accesso alle riviste il termine

controllato che descrive il loro ambito disciplinare (8). La lista completa dei descrittori assegnati alle riviste può essere consultata all'indirizzo <http://www.nlm.nih.gov/bsd/journals/subjects.html>, raggiungibile dalla *home page* della base di dati tramite il *link* ipertestuale *subject terms* (Figura 4).

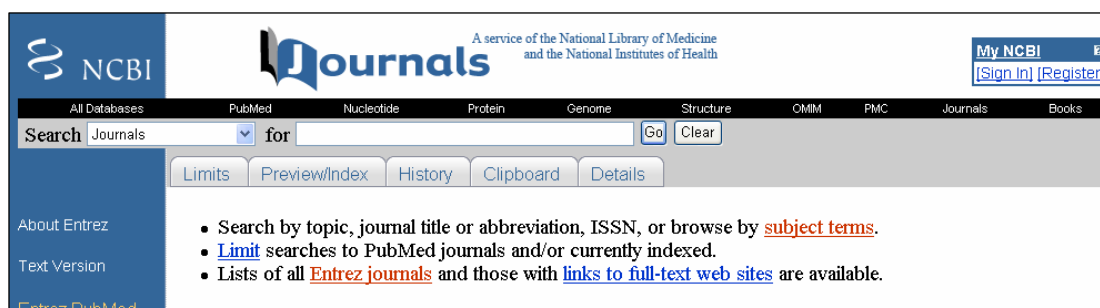


Figura 4. Home page del Journals Database, aggiornata al 07/12/06

Per effettuare l'interrogazione secondo questa nuova modalità è necessario utilizzare il qualificatore di ricerca ST (*Subject Terms*).

Così, ad esempio, se si è interessati alla consultazione delle riviste di interesse epidemiologico bisognerà digitare nella casella di ricerca: “*epidemiology* [ST]”. Alternativamente, dopo aver inserito il termine *epidemiology* si può applicare il filtro “*Subject Terms*” dalla pagina dei limiti, contenente un'altra interessante limitazione della ricerca, introdotta recentemente.

Spuntando la casella recante la dicitura “*Currently indexed in MEDLINE*”, è possibile restringere l'interrogazione alle sole riviste correntemente indicizzate nel MEDLINE, ovvero sia alle riviste che rientrano nel campo di copertura del MEDLINE. Lo stesso risultato può essere raggiunto direttamente dalla maschera di ricerca, aggiungendo alla voce *epidemiology* la stringa “*currentlyindexed*”. Così, tornando all'esempio precedente, se si vogliono recuperare i titoli delle riviste epidemiologiche indicizzate nel MEDLINE basta digitare: “*epidemiology* [ST] AND *currentlyindexed*”.

Il servizio My NCBI

Con la nuova interfaccia del PubMed sono state estese e migliorate anche le potenzialità del servizio *My NCBI*, mediante il quale l'utente può disporre di uno spazio di memoria riservato all'interno del PubMed.

Nella sua versione attuale questo strumento consente non solo di memorizzare in modo permanente le strategie di ricerca, nell'area denominata *Searches*, e riceverne gli aggiornamenti automatici al proprio indirizzo di posta elettronica, ma altresì di costruire la propria bibliografia personale salvando nelle *Collections* le citazioni bibliografiche risultanti dalle ricerche effettuate (Figura 5).

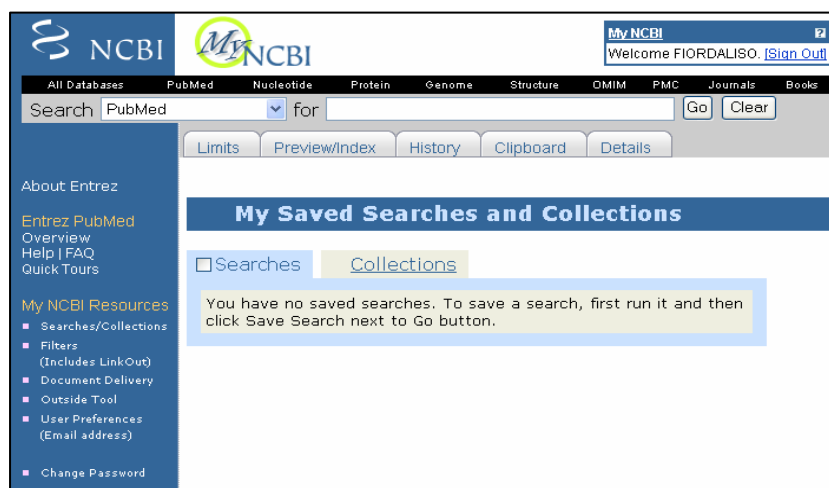


Figura 5. La home page del servizio My NCB, aggiornata al 07/12/06

Per poter usufruire di questo servizio, accessibile dalla barra laterale della schermata del PubMed, è necessario innanzitutto eseguire la registrazione, totalmente anonima e gratuita, inserendo i dati richiesti negli appositi spazi del modulo predisposto dal sistema. Dopo aver cliccato sul pulsante “Register”, l’utente avrà immediatamente a disposizione un’area di memoria personale presso il server della.nlm, in cui depositare le strategie di ricerca (fino ad un numero massimo di 100) e i risultati delle stesse, al quale potrà accedere a suo piacimento previa autenticazione.

Per memorizzare una strategia di ricerca occorre prima far eseguire l’interrogazione dal sistema, quindi cliccare sul pulsante “Save Search” presente sulla barra di ricerca della pagina dei risultati, e infine compilare il formulario intitolato ancora una volta “Save Search”, che appare sullo schermo dopo aver attivato My NCBI.

Il salvataggio delle strategie di ricerca permette di ricevere gli aggiornamenti direttamente sulla propria postazione di lavoro, via e-mail, con una periodicità e un formato di presentazione dei dati personalizzati dall’utente. A questo proposito si segnala il nuovo formato “Abstract Plus” che, oltre all’*abstract*, consente di visualizzare per ogni citazione i primi cinque record concettualmente correlati e ampliare, così, il rendimento quantitativo della ricerca.

Se si vuole effettuare l’esecuzione manuale dell’aggiornamento, dopo aver attivato il collegamento a My NCBI basta selezionare la strategia di ricerca desiderata, tramite l’apposita casella di spunta, e cliccare sul pulsante “What’s New for Selected”. Il sistema visualizzerà le citazioni inserite nel PubMed dalla data in cui è stato eseguito l’ultimo aggiornamento e manterrà traccia dell’interrogazione appena effettuata variando automaticamente la data dell’ultima esecuzione.

Il servizio My NCBI, come già accennato, può essere utilizzato oltre che per la memorizzazione delle strategie di ricerca anche per il salvataggio permanente dei risultati ottenuti, ovverosia delle citazioni bibliografiche.

Fino ad ora la memorizzazione a tempo indeterminato dei record recuperati durante un’interrogazione poteva essere effettuata soltanto archiviando questi dati in una memoria locale, come, ad esempio, il disco fisso del proprio personal computer. Da alcuni mesi, invece, è possibile usufruire del proprio account di My NCBI creando delle vere e proprie collezioni bibliografiche (fino ad un massimo di 100), contenenti ciascuna fino a 1.500 record.

La procedura da seguire per creare una collezione è suddivisa in due fasi distinte: i risultati devono essere immagazzinati preventivamente nella *Clipboard* e da qui trasferiti in *My NCBI*. Innanzitutto, dopo aver eseguito l'interrogazione, occorre selezionare dalla pagina dei risultati i record che si intendono memorizzare e, quindi, scegliere l'opzione *Clipboard* dal menu a tendina recante l'intestazione "Send to", al fine di depositare temporaneamente i dati richiesti nella *Clipboard* (che viene svuotata dopo 8 ore di inattività). Nel caso in cui non venga effettuata alcuna selezione verranno depositati automaticamente i primi 500 record ottenuti. Per trasferire i dati nel proprio account di *My NCBI*, invece, bisogna entrare nella *Clipboard*, tramite l'apposito pulsante posto nella barra orizzontale sotto la maschera di ricerca, e, selezionati i record che si vogliono memorizzare, scegliere l'opzione "*My NCBI Collections*" dal menu a tendina "Send to". Dopo essersi autenticati per attivare il collegamento a *My NCBI*, verrà visualizzata una finestra (*pop-up window*) recante l'intestazione "Save Collection", attraverso la quale si può creare una nuova collezione o aggiungere i record selezionati ad una collezione preesistente (se non si effettua alcuna selezione verranno memorizzate tutte le citazioni presenti nella *Clipboard*).

Le collezioni così create possono essere ordinate per nome (assegnato dall'utente), data dell'ultima modifica o numero di record memorizzati, mentre cliccando sulla denominazione di ogni singola collezione se ne può visualizzare il contenuto, cambiare l'intestazione o cancellare uno o più record (9).

Conclusioni

Dal momento del suo lancio ufficiale fino ad oggi il PubMed è stato sottoposto a continue revisioni miranti ad ampliarne le funzionalità e le potenzialità di ricerca, di cui abbiamo segnalato le più significative e innovative, e a farne uno strumento sempre più flessibile e maneggevole da parte del grande pubblico. Proprio per andare incontro alle esigenze di quest'ultimo è stato dotato di meccanismi tali da garantire l'esecuzione di ricerche complete ed esaustive anche da quanti non possiedono alcuna conoscenza della sua filosofia e della sua complessa architettura strutturale. Basti pensare, tanto per citarne uno, alla possibilità per l'utente inesperto di consultare l'archivio partendo dal linguaggio non specialistico, così come è abituato a fare con i comuni motori di ricerca su Internet, e ottenere una strategia di ricerca completa di descrittori controllati, grazie al processo di mappatura automatica (*automatic term mapping*), che consente di associare ai termini tratti dal linguaggio naturale i termini MeSH equivalenti.

Nell'arco di nemmeno un decennio, poi, è stato oggetto di applicazioni e innovazioni tecnologiche così incisive da modificarne radicalmente la struttura e la portata originaria.

Oltre al collegamento alle risorse correlate, interne ed esterne alla NLM, e ai servizi accessori, il PubMed offre, infatti, la possibilità di passare dalle semplici citazioni bibliografiche al testo completo, in formato elettronico, degli articoli, consentendo all'utente di concludere la ricerca di informazioni con il recupero del documento originale, contenente i dati primari (10).

Tramite il *link* ai siti degli editori scientifici che hanno aderito all'iniziativa del *LinkOut* (11), che permette di collegarsi dal PubMed al *full text* degli articoli, l'utente può visualizzare e stampare in tempo reale i documenti di proprio interesse, talvolta ancor prima della loro pubblicazione nell'edizione a stampa, alle condizioni stabilite dagli editori stessi (in genere sottoscrizione di un abbonamento, o formula *pay-per-view*).

L'accesso, invece, è totalmente gratuito per le riviste presenti in PubMed *Central* o nel caso di riviste i cui editori hanno attivato il servizio del *LinkOut* e per le quali la propria biblioteca abbia sottoscritto un contratto di licenza.

L'espansione dell'editoria in formato elettronico e l'avanzamento delle tecnologie informatiche applicate alla rete hanno così permesso al PubMed di valicare i confini propri di un database bibliografico, entro i quali è nato, per trasformarsi in uno strumento sempre più orientato verso l'integrazione delle risorse informative al servizio della collettività mondiale (12).

Bibliografia

1. Iannucci P, Trincali M. *Medline e Internet. Come utilizzare la più importante base dati bibliografica medica in Internet*. Roma: Il Pensiero Scientifico Editore; 1999.
2. National Library of Medicine. *NLM Technical Bulletin. May-June 1997, 296*. Disponibile all'indirizzo: http://www.nlm.nih.gov/pubs/techbull/mj97/mj97_web.html; ultima consultazione 07/12/06.
3. "About Entrez", disponibile all'indirizzo: <http://www.ncbi.nlm.nih.gov/Database/index.html>; ultima consultazione 07/12/06.
4. Dracos A, Della Seta M, Cammarano R. *PubMed. Guida pratica alla consultazione del Medline su Internet*. Roma: Di Renzo Editore; 1999.
5. "PubMed: MEDLINE Retrieval on the World Wide Web". Disponibile all'indirizzo: <http://www.nlm.nih.gov/pubs/factsheets/pubmed.html>; ultima consultazione 07/12/06.
6. "Search Fields Descriptions and Tags". PubMed Help. Disponibile all'indirizzo: <http://www.ncbi.nlm.nih.gov/books/bookres.fcgi/helppubmed/pubmedhelp.pdf>; ultima consultazione 07/12/06.
7. "Searching PubMed". PubMed Help. Disponibile all'indirizzo: <http://www.ncbi.nlm.nih.gov/books/bookres.fcgi/helppubmed/pubmedhelp.pdf>; ultima consultazione 07/12/06.
8. "Other Services Including the MeSH and Journals Database". PubMed Help. Disponibile all'indirizzo: <http://www.ncbi.nlm.nih.gov/books/bookres.fcgi/helppubmed/pubmedhelp.pdf>; ultima consultazione 07/12/06.
9. "My NCBI". PubMed Help. Disponibile all'indirizzo: <http://www.ncbi.nlm.nih.gov/books/bookres.fcgi/helppubmed/pubmedhelp.pdf>; ultima consultazione 07/12/06.
10. Lubiana L. PubMed. *Biblioteche oggi* 2000;6:20-5.
11. "LinkOut", disponibile all'indirizzo: <http://www.ncbi.nlm.nih.gov/entrez/linkout/>; ultima consultazione 07/12/06.
12. De Bellis N, Palazzi C. MEDLINE/PubMed nel XXI secolo. *Biblioteche oggi* 2003;3:39-49.

*La riproduzione parziale o totale dei Rapporti e Congressi ISTISAN
deve essere preventivamente autorizzata.
Le richieste possono essere inviate a: pubblicazioni@iss.it.*

*Stampato da Tipografia Facciotti srl
Vicolo Pian Due Torri 74, 00146 Roma*

Roma, dicembre 2006 (n. 4) 14° Suppl.