

Thesaurus italiano di bioetica: Un esempio di lessico specializzato nel settore bioetico

**Maurella Della Seta*, Federica Napolitani°,
Scilla Pizzarelli*, Elisabetta Poltronieri°**

*Settore Documentazione, °Settore Attività Editoriali
Servizio Informatico, Documentazione, Biblioteca ed Attività Editoriali
Istituto Superiore di Sanità
Viale Regina Elena, 299
00161 Roma

Abstract

The article describes origin and characteristics of the Italian Bioethics Thesaurus (TIB), as an example of a useful lexicon specialized in the field of Bioethics. Conceived as a supporting instrument for SIBIL (Italian Online Bioethics Information System), which is a database produced by the Documentation Service of the Italian National Institute of Health (Istituto Superiore di Sanità, ISS), TIB is currently the only example in Italy of a thesaurus entirely dedicated to the subjects of Ethics, Life Sciences and Public Health. The article also analyses some morphological, lexical and syntactic characteristics of TIB, along with the different semantic relations between controlled terms.

1 Introduzione

Il thesaurus, la cui stessa denominazione in lingua latina è indice delle sue antiche origini, costituisce attualmente uno degli esempi di lessico specializzato di maggior supporto ai sistemi automatizzati di recupero dell'informazione.

Se è vero, infatti, che la domanda di informazione specializzata (ad esempio in ambito biomedico) non è più solo appannaggio di una ristretta cerchia di esperti (ricercatori o documentalisti), bensì è divenuta oggetto di interesse da parte del grande pubblico (pazienti, cittadini, utenti di diversa estrazione e formazione), è altrettanto vero che, in parallelo, sono notevolmente incrementate la rilevanza e l'utilità pratica di questo importante strumento lessicografico.

Strumento che ha subito col tempo un non indifferente adattamento a questi nuovi scenari, e soprattutto alle importanti innovazioni e rapide trasformazioni imposte dalla Information Communication Technology (ICT) in diversi settori. L'impostazione grafica della versione cartacea dei thesauri di un tempo ha lasciato il posto a una veste editoriale spesso più leggera e maneggevole e soprattutto a versioni elettroniche, tramite l'utilizzo di applicazioni software *ad hoc* e di interfacce amichevoli. Tale passaggio dal prodotto cartaceo alla disponibilità in linea ha fatto sì che il thesaurus divenisse con gli anni uno degli strumenti più utili a supporto della terminografia assistita da computer.

In questo nuovo contesto si inserisce la realizzazione del Thesaurus Italiano di Bioetica (TIB).

2 Il Thesaurus Italiano di Bioetica (TIB)

Si tratta di un thesaurus monolingue in lingua italiana, costruito in piena adesione ai criteri stabiliti nello standard UNI ISO 2788/1993.¹

Il TIB è un vocabolario strutturato, ordinato alfabeticamente, composto da oltre 1000 descrittori correlati semanticamente e controllati rispetto al linguaggio naturale al fine di garantire la corrispondenza biunivoca tra un dato concetto e il termine che lo rappresenta. L'effetto è stato quello di ottenere un linguaggio comune a documentalisti e utenti, in grado, allo stesso tempo, di offrire "ancoraggi" al linguaggio libero.

Le funzioni del TIB costituiscono il cardine del Sistema Informativo per la Bioetica in Linea (SIBIL), un portale realizzato dal Settore Documentazione dell'Istituto Superiore di Sanità² per fornire supporto agli studi sulle implicazioni etiche della ricerca scientifica.

Il thesaurus rappresenta, da un lato, lo strumento di indicizzazione utilizzato nella base dati documentaria del SIBIL allo scopo di sintetizzare e tradurre in termini di significato univoco i concetti trattati nei documenti e, dall'altro, costituisce il canale di ricerca preferenziale della documentazione indicizzata, attraverso l'impiego di termini combinati con gli operatori logici booleani (AND, OR, NOT).

Il processo di sviluppo del TIB si è svolto a partire da un *corpus* lessicale primigenio di circa 800 descrittori raccolti nel Bioethics Thesaurus prodotto dal Kennedy Institute of Ethics della Georgetown University di Washington³ sul quale si sono col tempo innestati vari interventi di ampliamento e modifica, ad esempio: l'integrazione di terminologia direttamente estratta dalla letteratura indicizzata, gli aggiornamenti o le modifiche suggerite da parte di esperti di documentazione sulla base di fonti lessicografiche scientifiche (dizionari specializzati, nomenclature), l'apporto di specialisti del settore, la verifica della rispondenza dei termini alla realtà del dibattito bioetico in Italia, e così via.

3 Caratteristiche morfologiche, lessicali e sintattiche dei termini

Si forniscono qui di seguito alcune caratteristiche riguardanti la tipologia e la struttura dei descrittori del TIB.

3.1 Varianti morfologiche

3.1.1 Forma singolare vs forma plurale

Il TIB rispetta il criterio classico di distinzione tra termini numerabili espressi al plurale e non numerabili resi al singolare, anche se in alcuni casi la scelta tra le due forme è stata ope-

¹ La norma riporta le raccomandazioni cui attenersi nella costruzione di thesauri monolingue.

² Il sistema SIBIL www.iss.it/sibi è stato varato nell'ambito di un progetto approvato dall'Istituto Superiore di Sanità nel 2001 e si è incentrato sulla produzione di una base di dati che registra le pubblicazioni di bioetica edita in Italia a partire dal 1995.

³ L'ultimo aggiornamento del Bioethics Thesaurus risale al 2001; a partire da questa data i termini in esso raccolti sono confluiti nel thesaurus MeSH (Medical Subject Headings) prodotto dalla National Library of Medicine di Bethesda.

rata interpretando le consuetudini e le preferenze della comunità degli utenti del thesaurus. Assumono la forma plurale le classi di persone riferite a fasce di età (ad es. *Bambini, Anziani*), professione (ad es. *Giudici, Medici*), status sociale (ad es. *Carcerati, Stranieri*), confessione religiosa (ad es. *Ebrei, Protestanti*), i nomi di entità concrete quali documenti, prodotti e strutture sanitarie (ad es. *Cartelle cliniche, Farmaci, Ospedali*) e i nomi relativi a insiemi di attività e classi di concetti (ad es. *Cure palliative, Discipline umanistiche*). Il singolare è utilizzato per i concetti astratti (ad es. *Dolore, Potere*), i nomi di discipline (ad es. *Agricoltura, Oncologia*), i nomi di attività (ad es. *Allattamento, Farmacovigilanza*), di istituti giuridici (ad es. *Adozione*), di fenomeni di rilevanza sociale (ad es. *Suicidio, Uccisione*), di categorie di individui umani (ad es. *Donna, Padre*), di specie (ad es. *Animale*) di organismi (ad es. *Embrione, Neonato*) e di correnti filosofiche o di pensiero (ad es. *Agnosticismo, Comportamentismo*). Il TIB registra inoltre alcuni descrittori espressi nelle due forme, singolare e plurale, nei casi in cui un concetto si debba intendere come astratto e più generico rispetto alle sue varianti concrete (ad es. *Biotecnologia*, ma *Biotecnologie animali; Feto*, ma *Feti abortiti*).

3.1.2 Nomi propri

I nomi propri utilizzati nell'indicizzazione dei documenti della base di dati SIBIL sono raccolti in un Soggettario che costituisce un complesso di oltre 300 voci accessorie ai termini del TIB. Tali nomi sono impiegati a supporto dei descrittori, al fine di offrire ulteriori elementi di approdo all'informazione. Essi comprendono soggetti biografici, nomi di luogo e altre voci rappresentate da titoli di documenti contenenti principi di bioetica e denominazioni di enti e istituzioni note per aver dettato indirizzi e orientamenti della disciplina (ad es. *Hume, David (1711-1766), America Latina, Dichiarazione Universale dei Diritti dell'Uomo, Comitato Nazionale per la Bioetica*).

3.2 Varianti lessicali

3.2.1 Termini di prestito vs traduzione italiana

Nel TIB sono presenti i seguenti casi di prestito dalla lingua inglese:

- descrittori in inglese adottati come preferenziali rispetto alla forma in italiano. I descrittori sono accolti in inglese, in assenza di validi corrispondenti in italiano o in quanto maggiormente attestati rispetto alla forma in italiano nella letteratura censita (ad es. *Counselling, Doping, Single*);
- termini in inglese combinati con parole in italiano a formare un unico descrittore. Si tratta di forme consolidate nel linguaggio settoriale utilizzato per esprimere concetti di rilevanza bioetica (ad es. *Compliance del paziente, Counselling genetico, Screening genetico, Screening di massa, Test obbligatorio, Triplo test*);
- termini in lingua inglese che rinviano a descrittori in italiano. L'introduzione di tali termini quali punti di accesso all'informazione indica una apertura del TIB alle forme in uso corrente nel linguaggio naturale (Computer → *Elaboratori elettronici*, Welfare → *Bene comune*);
- termini in lingua inglese che vantano una discreta diffusione d'uso tra gli operatori del settore, per i quali, tuttavia, si è ritenuto conveniente conservare il rinvio dalla forma in ita-

liano (ad es. *DNA fingerprinting*, accolto in luogo di *Identificazione mediante confronto dei DNA*; *Managed care* accolto in luogo di *Assistenza sanitaria manageriale*).

3.2.2 Termini di uso comune vs lessico tecnico

L'oscillazione tra forme di uso corrente e terminologia specialistica esprime la volontà di fondere nel TIB il registro più colloquiale della lingua con il linguaggio specialistico noto agli utenti della materia, anche se non sempre si è operata una scelta omogenea. Coesistono infatti termini come *Aborto* (ritenuto più familiare per l'utente rispetto a *Interruzione della gravidanza*) e *Imperizia medica* (tecnicismo privilegiato in luogo del più popolare *Malasannità*). In alcuni casi si è dovuto necessariamente adottare la voce specificamente tecnica, in assenza di termini equivalenti per esprimere un concetto altamente specialistico e per evitare di accogliere una perifrasi in luogo di un singolo descrittore. E' il caso di *Anencéfalia* e *Determinismo genetico*. Il termine prettamente scientifico è stato accolto anche in relazione a patologie, come per *Sclerosi laterale amiotrofica* (nota comunemente come *Malattia di Lou Gehrig*).

3.3 Varianti sintattiche dei descrittori

3.3.1 Pre-coordinazione vs post-coordinazione

L'apparato terminologico del TIB si compone di descrittori costituiti da singoli morfemi e di descrittori derivati dalla associazione di più morfemi (sintagmi), caratterizzati spesso da una forte componente sintattica. Nell'economia del thesaurus questa compresenza fa sì che ai vantaggi della pre-coordinazione (specificità dei concetti condensati in locuzioni complesse e alleggerimento del numero complessivo dei descrittori) si sommino quelli derivanti dalla scelta di descrittori post-coordinati (modularità e quindi flessibilità del thesaurus in fase di indicizzazione e di ricerca). La struttura composita dei termini adottati nel TIB riproduce nel complesso un'ampia gamma di soluzioni sintattiche, espressive della ricchezza concettuale che caratterizza il dominio semantico della bioetica.

Su un totale di 1065 descrittori presenti nel TIB, quelli costituiti da un singolo termine (di valenza generica come nel caso di *Antropologia*, *Biosicurezza*, *Diritto* o di portata molto settoriale come per *Beneficialità*, *Consequenzialismo*, *Delibazione*) sono 396.

I descrittori costituiti da sintagmi (dal numero minimo di elementi: sostantivo più aggettivo, alla massima estensione sintattica: locuzioni sintatticamente molto articolate) sono 669. In particolare, si evidenziano:

- 1) sostantivo + aggettivo (ad es. *Malattie iatrogene*, *Etica ambientale*);
- 2) sostantivo + complemento di specificazione in cui il sostantivo esprime un aspetto e il complemento l'intero (ad es. *Anomalie della differenziazione sessuale*, *Malattie del rene*);
- 3) sostantivo + complemento di specificazione in cui il sostantivo esprime proprietà e il complemento il possessore (es. *Diritti dei pazienti*, *Tessere dei donatori*);
- 4) sostantivo + complemento di specificazione in cui il sostantivo esprime, di volta in volta, un approccio di studio, un comportamento, un'entità normativa (ad es. *Storia della medicina*, *Ruolo del medico*, *Statuto dell'embrione*) e il complemento indica il soggetto di riferimento;

5) sostantivo + complemento di specificazione che esprime l'azione implicita nel sostantivo (ad es. *Selezione dei pazienti, Crioconservazione degli embrioni*);

6) locuzioni articolate che riassumono concetti complessi e molto specifici (ad es. *Alternative alla sperimentazione animale, Atteggimento verso la morte, Metodi di regolazione naturale della fertilità, Effetti posticipati dell'esposizione prenatale*).

3.3.2 Ordine delle parole

I descrittori sono formulati in accordo all'ordine naturale delle parole (forma diretta), sono quindi escluse le forme con inversione adottate in alcuni thesauri allo scopo di porre enfasi sul concetto focale espresso da un dato descrittore (ad es. *Alternative alla sperimentazione animale* e non *Sperimentazione animale, Alternative alla; Utilizzazione degli embrioni* e non *Embrioni, Utilizzazione degli*).

3.4 Strumenti di ausilio alla comprensione dei descrittori TIB

Il TIB è corredato di note d'ambito (*scope notes*: SN) che hanno lo scopo di selezionare tra i significati di un termine quello ritenuto valido nel contesto di riferimento (ad es. *Autonomia* SN Capacità del soggetto di prendere decisioni liberamente. Usato in bioetica per indicare il "principio di autonomia", cioè il diritto, da parte di tutti i soggetti coinvolti, di prendere decisioni in una questione bioetica).

Nel caso di termini specialistici, meno comprensibili per il largo pubblico, la nota d'ambito coincide con la definizione del termine (ad es. *Specismo* SN Teoria filosofica che giudica ingiustamente discriminatoria la distinzione etica tra la specie umana e quelle non umane).

Le note d'ambito, infine, non sono presenti nei casi di descrittori che designano concetti acquisiti (*Genitori, Madre*).

4 Relazioni semantiche tra i descrittori

I descrittori del TIB sono corredati da relazioni semantiche; 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 preferenza, la relazione gerarchica e quella associativa.

4.1 Relazione di preferenza (o 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: USE = usa e UF = usato per (used for).

Nel caso dei sinonimi, si rinvia reciprocamente da un termine ad un altro di significato equivalente (ad es. *Cancro* USE *Neoplasie; Disobbedienza civile* USE *Obiezione di coscienza*).

4.2 Relazione gerarchica

Collega ciascun termine del TIB ad uno di significato più generale o più specifico nell'ambito della medesima categoria. Tale relazione può essere del tipo genere-specie (la più

ricorrente), parte-tutto oppure esemplificativa. I simboli che esprimono tale relazione sono: BT = *Broader Term* – termine più ampio – NT = *Narrower Term* – termine più specifico – (ad es. *Ospedali* BT *Strutture sanitarie*, NT *Ospedali psichiatrici*).

4.3 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 RT = *Related Term* (termine correlato) (ad es. *Madri surrogate* RT *Fecondazione in vitro*). I nessi fra i due termini correlati possono essere di tipo causa-effetto (*Gravidanza* RT *Parto*), di agente (*Donazione d'organo* RT *Donatori*), di opposizione (*Fertilità* RT *Infertilità*), di coordinazione stretta (*Occupazione* RT *Medicina del lavoro*).

Bibliografia

A. Dizionari

- Deutsches Referenzzentrum für Ethik in den Biowissenschaften (DRZE), Centre de Documentation en Éthique des Sciences de la Vie et de la Santé (INSERM/CCNE), Informations-und Dokumentationsstelle Ethik in der Medizin, Interfakultäres Zentrum für Ethik in den Wissenschaften, Library and Information Services of the Kennedy Institute of Ethics (Ed.) (2005). *Thesaurus Ethics in the Life Sciences*. 2. ed. Bonn, DRZE. Disponibile all'indirizzo: <http://www.drze.de/BELIT/thesaurus>.
- Kennedy Institute of Ethics (2001), *Bioethics Thesaurus*. Washington, Georgetown University.

B. Altri testi

- Associazione Italiana Biblioteche (AIB), Gruppo di Ricerca sull'Indicizzazione per Soggetto (GRIS) (2001), *Guida all'indicizzazione per soggetto*. Roma, AIB.
- Della Seta, M., Dracos, A., Poltronieri, E. (2004), 'Riflessione etica e ricerca scientifica. Produzione e diffusione di risorse informative.' *Annali dell'Istituto Superiore di Sanità* 40(3), pp. 281-392.
- National Library of Medicine. *Medical Subject Headings*. (2004) Bethesda, US Medical Library of Medicine. Disponibile all'indirizzo: <http://www.nlm.nih.gov/mesh/MBrowser.html>.
- Priano, M. (1994), *La struttura del soggettario*. Milano, Editrice Bibliografica.
- Trigari, M. (1992), *Come costruire un thesaurus*. Modena, Franco Cosimo Panini Editore (Informatica e Beni Culturali, 2).
- UNI ISO 2788. (1993) *Linee guida per la costruzione e lo sviluppo di thesauri monolingue*. Milano, Ente Nazionale Italiano di Unificazione.