

***Corso di formazione per operatori delle biblioteche del Friuli Venezia Giulia
Villa Manin di Passariano, Codroipo (UD), 20 maggio 2019***



I FERRI DEL MESTIERE: tecniche per l'informazione

**Corso di base teorico pratico di indicizzazione per soggetto
(Parte 1)**

Silvia Bonfietti (Università degli studi di Udine)

Programma

- L'indicizzazione semantica
- L'analisi concettuale
- I linguaggi di indicizzazione semantica
- L'indicizzazione alfabetica per soggetto
- Il Thesaurus del Nuovo soggettario
 - L'analisi categoriale
 - Il controllo morfologico
 - Le relazioni
 - I nomi propri
 - Le fonti

Indicizzazione semantica

L'indicizzazione semantica ha lo scopo di rispondere a richieste informative e ricerche su un argomento, un soggetto, una materia, formulate prescindendo dalla conoscenza di specifici autori e titoli

Indicizzazione semantica / 2

Fasi dell'indicizzazione semantica:

1) Analisi concettuale

esame del documento al fine di stabilirne il soggetto e individuare i concetti principali presenti nel soggetto stesso

2) Traduzione nel linguaggio di indicizzazione

formulazione dei concetti selezionati e delle relazioni logiche individuate nei termini e nelle forme proprie del linguaggio di indicizzazione prescelto

Analisi concettuale

Lo scopo dell'analisi concettuale è quello di individuare e definire esattamente il soggetto di un documento.

Sull'individuazione del soggetto influiscono le capacità intellettuali e la cultura dell'indicizzatore. Tuttavia questa attività non deve ridursi ad un atto di creatività individuale, ma deve essere condotta secondo una procedura sistematica e normalizzata, tale da produrre risultati sostanzialmente uguali o, comunque, condivisibili, anche se eseguita in contesti diversi.

Analisi concettuale/2

L'analisi concettuale è descritta nella prima parte della *Norma UNI ISO 5963-1985: Documentation -- Methods for examining documents, determining their subjects, and selecting indexing terms*

La norma stabilisce le procedure raccomandate per:

- l'analisi dei documenti
- la determinazione del loro contenuto
- la scelta dei termini di indicizzazione

Analisi concettuale/3

L'analisi del documento prevede una attenta lettura del testo, nei suoi punti più importanti:

- il titolo e sottotitolo
- il riassunto analitico, se presente
- il sommario o indice
- l'introduzione, le frasi iniziali dei capitoli e dei paragrafi e la conclusione
- illustrazioni, diagrammi, tavole con relative didascalie
- parole o gruppi di parole sottolineate/evidenziate

E' importante il confronto fra più fonti!

(nessuna fonte da sola deve essere considerata sufficiente per l'esame del testo)

Analisi concettuale/4

Il prodotto finale dell'analisi concettuale del documento è l'**enunciato di soggetto**, frase in forma di sintagma nominale con la quale viene espresso il tema del documento.

Cosa non è contenuto nel tema di base (e quindi nel soggetto o nella classe che lo esprime)

- Il riassunto del documento
- Il percorso o i risultati di un'indagine
- Le riflessioni dell'autore

Analisi concettuale/5

Esempio1:

Euro sì : morire per Maastricht / Enrico Letta. - Roma : Laterza, 1997. - 92 p. ; 17 cm (Il nocciolo ; 35). - ISBN 8842052485

Euro no : non morire per Maastricht / Lucio Caracciolo. - Roma : Laterza, 1997. - 102 p. ; 17 cm (Il nocciolo ; 36). - ISBN 8842052493

Tema di base: La moneta unica nei Paesi dell'Unione europea

Moneta – Unificazione – Paesi dell'Unione europea

Analisi concettuale/6

Esempio2:

Non è da tutti : l'indicibile fortuna di nascere donna / Luisa Muraro
Roma : Carocci, 2011. - 126 p. ; 22 cm. (Sfere ; 57). - ISBN
9788843057825

Questioni di genere : tra vecchi e nuovi pregiudizi e nuove o presunte
libertà / a cura di Margarete Durst, Sonia Sabelli. - Pisa : ETS, 2013.
211 p. ; 22 cm (Scienze dell'educazione ; 161). - ISBN 9788846737663

Tema di base: Il ruolo della donna nella società

Donne – Ruolo sociale

Analisi concettuale/7

L'enunciato di soggetto (e la corrispondente traduzione in un linguaggio di indicizzazione) deve possedere le seguenti caratteristiche:

- Eshaustività:

L'analisi concettuale dovrebbe portare all'individuazione di tutti i concetti e delle relazioni che definiscono esattamente il soggetto del documento

- Specificità:

Precisione con cui un particolare concetto identificato nel soggetto del documento è espresso nell'enunciato di soggetto

Analisi concettuale/8

Esempio3:



Cuccioli e primi mesi. - Milano : Corriere della Sera : Oggi, 2018.
(Il gatto : capirlo, conoscerlo, amarlo ; 4)

Tema di base: Cuccioli di gatto

Gatti
Cuccioli di gatto

Analisi concettuale/9

Esempio4:



La fabbrica dei profumi : Seveso 40 anni fa /
Daniele Biacchessi. - Milano : Jaca book, 2016. - 130
p. ; 23 cm (Ecologica). - ISBN 9788816413603

Tema di base: Effetti dannosi dell'industria chimica
a Seveso nel 1976

Industria - Effetti dannosi - Seveso – 1976

Industria chimica - Effetti dannosi - Seveso

**Industria chimica - Effetti dannosi - Seveso -
1976**

Valutazione dei sistemi di indicizzazione

Esaustività e specificità influenzano due misuratori per la valutazione dell'efficienza di un sistema di recupero dell'informazione:

- il richiamo
- la precisione

Valutazione dei sistemi di indicizzazione/2

Richiamo

Il richiamo indica la percentuale di documenti rilevanti effettivamente ottenuti, rispetto al totale (di quelli rilevanti presenti nel catalogo), in base a una richiesta su un determinato soggetto o classe.

Es. Se si recuperano tutti i documenti rilevanti presenti nel catalogo, il richiamo è del 100%

Valutazione dei sistemi di indicizzazione/3

Precisione

La precisione indica la percentuale di documenti effettivamente rilevanti, rispetto al totale dei documenti ottenuti, in base a una richiesta su un determinato soggetto o classe

Es. Se si recuperano solo documenti rilevanti, la precisione è del 100%

Valutazione dei sistemi di indicizzazione/4

Esempio:

In una raccolta ho 24 documenti che parlano della Batteria (nel senso di strumento musicale)

Risultato della ricerca a catalogo inserendo Batterie nella ricerca per soggetto:

- 9 documenti recuperati dei quali
- 4 effettivamente rilevanti

Richiamo: $8/24 = 33\%$

Precisione: $4/9 = 44\%$

Linguaggio naturale vs Linguaggio documentario

Il **linguaggio naturale** è un linguaggio libero:

- i termini che lo compongono non sono stati sottoposti a regole prestabilite atte a controllare la loro forma ed il loro significato
- è un linguaggio poco costoso

Il **linguaggio documentario** è un linguaggio controllato

- i termini che lo compongono sono stati sottoposti a delle regole prestabilite atte a controllare la loro forma ed il loro significato
- è un linguaggio costoso

Linguaggi di indicizzazione semantica

I linguaggi di indicizzazione semantica sono linguaggi documentari che vengono utilizzati per la rappresentazione coerente, formalizzata e sintetica del contenuto concettuale dei documenti, funzionale alla segnalazione e al reperimento dei documenti stessi

Linguaggi di indicizzazione semantica / 2

Nell'**indicizzazione derivata**:

- si usano come indici i termini del documento stesso (parole del titolo, o dell'intero testo, o dell'abstract o parole chiave estratte automaticamente)
- la fase di input è automatizzabile perciò molto economica, mentre la fase di ricerca non garantisce richiamo e precisione

Nell'**indicizzazione assegnata**:

- i termini utilizzati vengono sottoposti a delle regole prestabilite atte a controllare la loro forma ed il loro significato
- La fase di input è più impegnativa, ma è più economica in fase di output (cioè di ricerca)

Linguaggio di indicizzazione semantica / 3

I linguaggi di indicizzazione semantica sono costituiti da:

- Un insieme di termini ammessi per descrivere il contenuto dei documenti (vocabolario)
- Un insieme di regole che stabiliscono come i termini devono essere usati (sintassi), e in quale relazione si trovino con ogni altro termine (semantica)

Linguaggi di indicizzazione semantica /4

Si distinguono in base alla sintassi utilizzata in:

- **Linguaggi postcoordinati**

I termini d'indice attribuiti a un documento non sono legati tra loro e quindi sono ricercabili singolarmente o associandoli al bisogno nel momento della ricerca

- **Linguaggi precoordinati**

I termini d'indice scelti sono presentati secondo relazioni e in un ordine stabiliti. Il soggetto che ne risulta dà un'immagine complessiva del contenuto del documento

Linguaggi di indicizzazione semantica/5

I linguaggi di indicizzazione preordinati si distinguono in base al struttura utilizzata in:

- Enumerativi: elencano tutti i termini o combinazioni di termini che devono essere usati per descrivere il soggetto di un documento e non consentono altre combinazioni se non quelle già previste
- Sintetici: elencano i termini che devono essere usati, fornendo una serie di regole per la loro combinazione

Linguaggi di indicizzazione semantica / 6

I linguaggi enumerativi sono:

- più rigidi
- più difficili da aggiornare
- meno espressivi (possono esprimere solo ciò che elencano)

Esempi: CDD, Soggettario, LCSH

I linguaggi sintetici sono:

- più flessibili
- più facili da aggiornare
- più espressivi

Esempi: Classificazioni a faccette, Nuovo Soggettario

Linguaggi di indicizzazione semantica / 7

I linguaggi di indicizzazione preordinati si distinguono in base al vocabolario utilizzato in:

- **Indicizzazione alfabetica (o verbale)**

l'indice rappresenta il soggetto del documento con espressioni della lingua naturale in un linguaggio controllato (soggettazione)

- **Indicizzazione sistematica (o classificata)**

l'indice rappresenta il soggetto del documento inserendolo nel contesto delle classi di significati secondo cui è trattato, con notazioni convenzionali previste da un sistema di classificazione strutturato (classificazione)

Linguaggi di indicizzazione semantica / 8

Indicizzazione alfabetica vs sistematica

- L'indicizzazione verbale alfabetica per soggetto è molto semplice e intuitiva, ma ha la caratteristica di disperdere concetti affini e di essere legata alla conoscenza specifica di una lingua

Esempio1:

Nuovo soggettario:

Cereali
Grano
Mais



CDD (14a ridotta):

633.1 Cereali
Include avena, grano, mais

Linguaggi di indicizzazione semantica/9

Indicizzazione sistematica vs alfabetica

- L'indicizzazione sistematica si avvale di una notazione specifica che rende più immediate le relazioni gerarchiche del soggetto, ma disperde i documenti che trattano dello stesso soggetto nell'ambito di discipline differenti

Esempio2:

Nuovo soggettario:

Matrimonio-Concezione cristiana →

Matrimonio-Diritto romano →

CDD (23a integrale):

241.63 Teologia morale e cristiana. Etica delle relazioni familiari

346.37016 Diritto privato. Matrimonio. Italia antica

Indicizzazione alfabetica per soggetto



SOGGETTO = ogni concetto o combinazione di concetti che rappresenta un tema all'interno di un documento (Norma ISO 5963/1985)

L'indicizzazione alfabetica risponde a due esigenze fondamentali:

- ricerca secondo il contenuto
- la cui espressione sia ordinata alfabeticamente

Indicizzazione alfabetica per soggetto / 2

Fasi del processo di indicizzazione per soggetto:

- analisi concettuale  *per determinare il soggetto di un documento*
 - controllo terminologico (piano lessicale-semantico)
 - costruzione delle stringhe di soggetto (piano logico-sintattico)
-  *per tradurre il soggetto di un documento nel linguaggio di indicizzazione*

Indicizzazione alfabetica per soggetto/3

Principi dell'indicizzazione per soggetto:

- **specificità**
- **esaustività**
- **uniformità** = un soggetto deve essere rappresentato nel linguaggio di indicizzazione da un solo termine o sequenza di termini
- **univocità** = ciascun termine o sequenza di termini deve rappresentare un solo soggetto

Indicizzazione alfabetica per soggetto/4

Cenni storici:

- Sec. 18.: ordinamento alfabetico utilizzato per indici per soggetto in appendice a bibliografie, o classificazioni e presenza di elementi di soggetto nel catalogo per autori (es. biografie)
- Sec. 19.: nasce il catalogo dizionario e a fine secolo il catalogo alfabetico per soggetti
- Sec. 20.: vengono pubblicati e diffusi i soggettari e vengono formulati principi e metodi dell'indicizzazione semantica per soggetto

Indicizzazione alfabetica per soggetto/5

Cenni storici in Italia:

- 1899** > la Biblioteca Angelica di Roma inizia il suo catalogo per soggetto
- 1910** > il BOMS (Bollettino delle opere moderne straniere acquistate dalle biblioteche pubbliche governative del regno d'Italia) inaugura un indice alfabetico per soggetto
- 1925** > la BNCF adotta il catalogo per soggetto e inaugura l'indice alfabetico per soggetto nel *Bollettino delle pubblicazioni ricevute per diritto di stampa*
- 1930** > la Biblioteca Vaticana redige le sue norme di soggettazione ispirate alla tradizione americana
- 1956** > la BNCF pubblica il *Soggettario di Firenze*
- 1977-1999** > la BNCF pubblica diverse *Liste di aggiornamento* delle voci di soggetto
- 1996** > il Gruppo GRIS (AIB) pubblica Guida all'indicizzazione per soggetto

Indicizzazione alfabetica per soggetto/6

Evoluzione teorica internazionale: norme ISO

ISO 5963/1985 Methods for examining documents, determining their subjects, and selecting indexing terms

ISO 5964/1985 Guidelines for the establishment and development of multilingual thesauri

ISO 2788/1986 Guidelines for the establishment and development of monolingual thesauri

ISO 25964/2011-2013 Thesauri and interoperability with other vocabularies
Part 1: Thesauri for information retrieval
Part 2: Interoperability with other vocabularies

Indicizzazione alfabetica per soggetto / 7

Evoluzione teorica: documenti IFLA

- 1991** > Guidelines for subject authority and reference entries (GSARE)
- 1998** > Functional requirements for bibliographic record (FRBR)
- 1999** > Principles underlying subject heading languages
- 2005** > Guidelines for multilingual thesauri
- 2009** > Statement of international cataloguing principles (ICP)
- 2010** > Functional requirements for subject authority data (FRSAD)
- 2017** > Ifla Library reference model (LRM)

Subject Analysis and Access Section <https://www.ifla.org/subject-analysis-and-access>

Indicizzazione alfabetica per soggetto/8

Tappe del rinnovamento del Soggettario di Firenze

2000 > Inizia il Progetto sul rinnovamento del *Soggettario* promosso dalla BNCF

2002 > Pubblicazione di: *Per un nuovo Soggettario* (Studio di fattibilità commissionato dalla BNCF alla IFNET di Firenze)

2004 > Iniziano i lavori sul Thesaurus del *Nuovo soggettario*

2006 > Pubblicazione del Prototipo del Thesaurus e della "Guida" del *Nuovo soggettario*

2007- > Implementazione aggiornamento del Thesaurus e impiego in BNI

Nuovo Soggettario



Linguaggio di indicizzazione per soggetto utilizzato in Italia, realizzato a cura della Biblioteca nazionale centrale di Firenze

Nuovo Soggettario/2

Caratteristiche principali del Nuovo soggettario:

- È un linguaggio preordinato di tipo analitico-sintetico
- organizza concetti e termini su due piani complementari:
 - semantica (il vocabolario in forma di thesaurus)
 - sintassi (le stringhe di soggetto)

Nuovo Soggettario/3

Indicizzazione preordinata



l'indicizzatore 'coordina' (mette in relazione) i termini del vocabolario al fine di costruire le stringhe di soggetto



regole sintattiche (implicite o esplicite)
per la costruzione delle stringhe

Nuovo Soggettario/4

Modello analitico-sintetico

Analisi

Possibilità di analizzare qualsiasi soggetto nei suoi elementi costitutivi, i singoli concetti e i termini che li esprimono, attraverso:

- l'analisi del significato
- l'analisi categoriale (per strutturare e controllare il vocabolario)
- l'analisi dei ruoli (per costruire le stringhe di soggetto)

Nuovo Soggettario / 5

Modello analitico-sintetico

Sintesi

Possibilità di organizzare concetti e termini su due piani distinti e complementari:

- quello della semantica
- quello della sintassi

N.B. Viceversa nei linguaggi enumerativi le stringhe di soggetto sono solamente quelle specificate in una lista di autorità

Nuovo Soggettario / 6

Modello analitico-sintetico

Vantaggi

- Semplicità: il vocabolario è ristretto ai soli termini (e non alle combinazioni fra termini) e le regole sintattiche sono più semplici perché basate su criteri di applicabilità generale
- Precisione: è facilitata la costruzione delle stringhe 'su misura', coesistenti con il contenuto del soggetto che debbono rappresentare
- Predittività: essendo costruito su criteri logici è più idoneo a esprimere i soggetti in modo coerente e omogeneo, quindi prevedibile



Nuovo Soggettario /7

Corredo sintattico – applicativo

- Note sintattiche
- Manuale



Aggiornamenti
semestrali



OPAC
Archivio dei soggetti
Stringhe di soggetto

Thesaurus

Il thesaurus è un vocabolario di un linguaggio di indicizzazione controllato, organizzato formalmente in modo da rendere esplicite le relazioni a priori tra concetti" (ISO 2788)

The screenshot shows the website for the 'Nuovo soggettario THESAURUS' from the Biblioteca Nazionale Centrale Firenze. The page features a navigation menu with 'Criteri', 'Aiuto alla ricerca', 'Sigle e simboli', 'Fonti', 'Novità', and 'Download'. On the left, there are links for 'Ricerca per termine del Nuovo soggettario', 'Soggettario(1956) Liste di aggiornamento (1956-1985)', and 'Sistema Nuovo soggettario' with sub-links for 'home page' and 'manuale applicativo'. The main search area is titled 'ricerca per termine' and includes radio buttons for search criteria: 'comincia per' (selected), 'termine esatto', and 'contiene'. Below this is a search input field and an 'esegui' button. A prompt 'digita un termine o parte di esso nella casella' is followed by a keyboard layout of letters A-Z. At the bottom, it states 'Prima edizione: gennaio 2007' and 'Ultimo aggiornamento: marzo 2017'.

Thesaurus/2

Regole fondamentali per la costruzione di un thesaurus:

- un concetto è sempre rappresentato da un solo termine (principio di uniformità)
- ad ogni termine corrisponde un solo concetto (principio della univocità)
- le relazioni presentate nel vocabolario sono sempre relazioni a priori, ossia valide in tutti i contesti in cui il concetto può comparire
- le relazioni tra i termini sono ricondotte a tipi di relazione definiti in modo rigoroso (USE/UF, BT/NT, RT)

Thesaurus/3

Principio di uniformità —→ controllo della sinonimia

Il **controllo della sinonimia** si realizza in campo semantico attraverso la scelta di un termine preferito e l'approntamento di rinvii dai sinonimi e quasi-sinonimi



Thesaurus/4

Principio di univocità —> controllo della polisemia

Il **controllo della polisemia** avviene precisando, nel caso di termini omonimi o polisemici, il preciso significato col quale il termine è utilizzato all'interno del linguaggio d'indicizzazione

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z	
termini trovati: 39	
Riso (LCSH: Oryza)	Agenti: Organismi
Riso <Cereali> (LCSH: Rice)	Cose: Materia
Riso <Comportamento> (LCSH: Laughter)	Azioni: Attività
Riso <Psicologia>	Azioni: Attività
Riso brillante	Cose: Materia
Risolubilità del contratto per inadempimento	Azioni: Attività

Thesaurus/5

Formalizzazione del vocabolario

riduzione della complessità semantica del linguaggio naturale attraverso operazioni di analisi e controllo del significato dei termini e delle loro relazioni attraverso:

- riduzione della molteplicità dei termini impiegabili per esprimere un concetto
- riduzione dei significati associabili ad un termine
- riduzione delle relazioni che si possono instaurare fra termini

Thesaurus/6

Provenienza dei termini del Thesaurus

- termini provenienti da *Soggettario* e *Aggiornamenti* (pubblicati)
- termini introdotti dalla BNI dal 1999 al 2005 (mai pubblicati)
- termini introdotti per le relazioni o importanti per la struttura
- termini suggeriti dagli enti che collaborano
<https://thes.bncf.firenze.sbn.it/enti.htm>

Thesaurus/7

Fasi della 'formalizzazione' per ogni termine:

- analisi del significato
- analisi categoriale
- controllo morfologico
- definizione delle relazioni e della gerarchia

Thesaurus: analisi del significato

Principio della definizione tipica

il significato formalizzato di un termine in un vocabolario controllato deve essere il più vicino possibile al significato tipico del termine nel linguaggio naturale, ossia al significato di base (verifica nei dizionari)

Thesaurus: analisi categoriale

- Individuazione di una serie di tipi concettuali fondamentali
- utilizzo di questi tipi concettuali per la costruzione di una macrostruttura semantica del vocabolario
- per ogni singolo termine del vocabolario, formulazione di un giudizio di appartenenza categoriale basato sull'analisi del suo significato di base

Thesaurus: analisi categoriale/2

Macrostruttura in cui viene organizzato il thesaurus

Categorie	Faccette
Agenti	Organismi Organizzazioni Persone e gruppi
Azioni	Attività Discipline Processi
Cose	Forme Materia Oggetti Spazio Strumenti <u>Strutture</u>
Tempo	

Thesaurus: analisi categoriale/3

Funzioni dell'analisi categoriale:

- raggruppare i termini del linguaggio in unità più piccole rispetto all'universo terminologico di partenza, in modo da rendere più agevoli tutte le operazioni successive
- individuare le classi fondamentali nelle quali è articolato il vocabolario su basi strettamente definitorie
- controllare la coerenza delle gerarchie

Thesaurus: analisi categoriale/4

Agenti: entità che possono essere agenti di azioni, singolarmente o collettivamente

- **Organismi:** comprende gli organismi viventi, sia animali che vegetali; le loro parti; i raggruppamenti di organismi esclusi quelli umani
- **Organizzazioni:** comprende le organizzazioni o parti di organizzazioni
- **Persone e gruppi:** persone e gruppi che non danno luogo ad alcuna organizzazione

Thesaurus: analisi categoriale/5

Azioni: tipologie dell'agire

- **Attività:** azioni dotate di un agente diretto chiaramente individuabile
- **Discipline:** azioni che presentano come caratteristiche principali: una forte componente di ricerca nelle loro attività; una finalità di accumulo e diffusione di conoscenze; una struttura organizzativa più o meno formalizzata che supporta quest'attività (università, pubblicazioni, ...)
- **Processi:** azioni prive di un agente diretto, bensì provocate da una causa, naturale o sociale oppure da azioni o movimenti collettivi

Thesaurus: analisi categoriale/6

Cose: entità concrete, tangibili (materia, oggetti, strutture), ma anche astratte (strumenti e forme)

- **Materia:** comprende la materia intesa in senso astratto, come anche i materiali: elementi, sostanze, prodotti (materiali che hanno subito una qualche trasformazione ma non sono ancora oggetti)
- **Oggetti:** cose, almeno potenzialmente mobili, che abbiano forma definita, tangibile, finita
- **Strutture:** manufatti che occupano uno spazio in maniera stabile, cioè che non sono mobili, anche se l'occupazione può essere temporanea

Thesaurus: analisi categoriale/7

Cose [segue]:

- **Strumenti:** cose, elementi, situazioni e strumenti astratti, che costituiscono il mezzo per ottenere qualcosa; non comprende tutti i concetti che hanno funzioni strumentali, ma solo gli strumenti astratti o quelli in cui la componente oggettuale è secondaria
- **Forme:** comprende l'aspetto esteriore, la rappresentazione, la manifestazione di entità oggettuali, astratte o frutto della creatività e della fantasia, comprese le forme della comunicazione (es. letterarie, musicali, ecc.)
- **Spazio:** comprende i concetti di estensione, ma non le entità che occupano tale estensione

Thesaurus: analisi categoriale/8

Tempo: contiene termini per i quali uno svolgimento cronologico è l'elemento principale o esclusivo; comprende i concetti puramente cronologici e quelli nei quali il tempo, rimanendo caratteristica fondamentale, si sovrappone a quella di avvenimento o attività

Thesaurus: controllo morfologico

- **Numero**
- **Scomposizione**
- **Disambiguazione**

Riferimenti:

> ISO 2788-1986

> *Soggettario* (= in caso dubbio, ove la norma non aiuti, si è preferito tenere conto della tradizione mantenendo la forma degli strumenti preesistenti)

Thesaurus: controllo morfologico Numero

Il Thesaurus contiene:

- sia termini singolari
- che termini plurali

Motivazioni:

- per non perdere per quanto possibile il patrimonio terminologico del Soggettario
- per l'opportunità di diversificare (in un sistema preordinato) la valenza sintattica di certi termini

Thesaurus: controllo morfologico

Numero/2

Criteria per la scelta del numero:

- principio della numerabilità

si usa il plurale per i termini che rappresentano concetti numerabili (quanti?), il singolare per i concetti non numerabili (quanto?); in caso dubbio si da precedenza al numero tipico della categoria semantica di appartenenza del termine

es.1 **Gessi** (Soggettario)

USE **Gesso** (Nuovo Soggettario)

es.2 **Madre** (Soggettario)

USE **Madri** (Nuovo Soggettario)

Thesaurus: controllo morfologico Numero/3

Categorie/Faccette		Numero tipico
Agenti	Organismi Organizzazioni Persone e gruppi	Plurale
Azioni	Attività Discipline Processi	Singolare, con numerose eccezioni
Cose	Forme Materia Oggetti Spazio Strumenti Strutture	Plurale Singolare, con molte eccezioni Plurale Plurale Plurale, con molte eccezioni Plurale
Tempo		

Thesaurus: controllo morfologico Numero/4

Criteria per la scelta del numero [segue]:

- principio dello scostamento categoriale

possono coesistere due forme, plurale e singolare dello stesso termine, purché appartengano a due faccette diverse

es.3: **Legno**

Nota d'ambito: La parte solida e compatta del tronco [...]

Legni

Nota d'ambito: Strumenti a fiato costruiti, almeno alle origini, con il legno [...]

Altri esempi: **Bibliografia** e **Bibliografie**; **Scenografia** e **Scenografie**; **Croce** e **Croci**

Thesaurus: controllo morfologico Scomposizione

Secondo la norma ISO 2788-1986 un termine composto si scompone se:

- Il focus si riferisce ad una parte o ad una proprietà o ad un materiale e la differenza ne rappresenta l'intero o il possessore

es.1: **Musei degli enti locali**

Focus

Differenza

MUSEI + ENTI LOCALI

Thesaurus: controllo morfologico Scomposizione/2

E ancora, un termine composto si scompone se:

- Il focus indica un'azione transitiva e la differenza indica il suo oggetto

es.2: **Protezione della natura**

Sfruttamento della prostituzione

- Il focus indica un'azione intransitiva e la differenza indica il suo agente

es.3: **Delinquenza minorile**

Thesaurus: controllo morfologico

Scomposizione/3

Casi particolari di scomposizione:

- Il focus rappresenta un'azione transitiva e la differenza il suo agente (anche indiretto)
es.: **Danni da calamità**
Infezione da Papillomavirus
Contaminazione da funghi
- Il focus rappresenta l'agente e la differenza l'oggetto su cui si esercita l'azione
es.: **Dirigenti pubblici**
Dirigenti sanitari
Insetti nocivi alle conifere

Thesaurus: controllo morfologico

Scomposizione/4

Casi particolari di scomposizione nel NS:

Termini in cui il focus esprime la forma dei documenti, che hanno come argomento la disciplina, l'attività, etc., espressa dalla differenza

es.: **Libri di magia**
Periodici letterari

Voci costituite da sintagmi che indicano un particolare 'tema' in contesti letterari, artistici, cinematografici ...

es.: **Abbigliamento nella letteratura**

Thesaurus: controllo morfologico

Scomposizione/5

Casi particolari di scomposizione nel NS:

Voci costituite da sintagmi che specificano un oggetto-méta, un ambito o un agente indiretto:

'Impiego in ...'; 'Concezione di...'; 'Concetto di...'; 'Influssi di ...'; 'Atteggiamento verso ...'; 'Rapporti con ...'

es.: **Impiego nelle operazioni finanziarie**

Concezione del nulla

Influssi del giudaismo

Atteggiamento verso il pericolo

Rapporti con i malati

Thesaurus: controllo morfologico

Scomposizione/6

Non si scompongono:

Termini che sono entrati nel lessico corrente o che sono attestati nella letteratura scientifica di un settore

es.: **Qualità della vita**

Termini, usualmente considerati nomi comuni, che contengono nomi propri

es.: **Arca di Noè**

Malattia di Crohn

Termini che hanno un valore metaforico

es.: **Arca dell'Alleanza**

Thesaurus: controllo morfologico

Scomposizione/7

Non si scompongono:

Termini composti da nomi, non uniti da preposizione, che assolvono contemporaneamente e reciprocamente sia la funzione di focus che di differenza

es.: **Preti operai**
Bambini attori

Termini che contengono sostantivi sincategorematici

es.: **Fiori di carta**
Uova di cioccolata

termini che non possono
essere considerati una
sottoclasse del loro focus

Thesaurus: controllo morfologico

Disambiguazione

Omografia

= fenomeno del linguaggio naturale per il quale possono coesistere termini con forma uguale ma con significato diverso

Disambiguazione

= tecnica con la quale termini omografi vengono distinti l'uno dall'altro attraverso l'aggiunta di una qualificazione

Thesaurus: controllo morfologico Disambiguazione/2

Regole di utilizzo nel NS

- Minimo ricorso alla disambiguazione (in un linguaggio preordinato la disambiguazione può rappresentare un ostacolo alla comprensione delle stringhe di soggetto)
- preferenza di un'espressione alternativa piuttosto che di un termine con qualificazione

es. **Calcio (Gioco)** (*Soggettario*)

Gioco del calcio (*NS*)

es. **Anelli (Matematica)** (*Soggettario*)

Anelli algebrici (*NS*)

Thesaurus: controllo morfologico Disambiguazione/3

Regole di utilizzo nel NS

- disambiguazione di 'tutti i termini meno uno', dove il termine senza qualificazione è:
 - il termine privo di qualificazione nel Soggettario
 - il significato dato come principale nei repertori
 - il termine con maggiore garanzia bibliografica
 - il primo inserito in ordine di tempo

es. **Crittografia**
Crittografia <Enigmistica>
Crittografia <Informatica>

Thesaurus: controllo morfologico Disambiguazione/4

Scelta e forma della qualificazione:

- termine che appartiene alla stessa categoria del termine da disambiguare
- termine che corrisponde preferibilmente al termine gerarchicamente sovraordinato rispetto al termine da disambiguare
- segue il termine ambiguo ed è data fra parentesi uncinate con l'iniziale maiuscola

es. **Divisione <Aritmetica>** (*Soggettario*)
Divisione <Operazioni aritmetiche> (*NS*)

Thesaurus: Relazioni

Relazioni del thesaurus:

- Equivalenza → USE (use)
UF (used for)
- Gerarchica → BT (broader term)
NT (narrower term)
- Associativa → RT (related term)

Thesaurus: Relazioni

Relazioni di equivalenza:

- relazione sinonimica vera e propria
es. **Bambini**
UF Fanciulli
- relazione di quasi sinonimia
es. **Assistenza psicologica**
UF Counseling
- relazione '2 a 1'
es. **Aziende**
UF+ Biblioteche aziendali

Gruppo di equivalenza = insieme dei termini equivalenti fra loro, fra i quali solo uno è il termine preferito

Thesaurus: Relazioni/2

Relazioni gerarchiche:

- relazione generica (genere/specie)
es. **Pompelmi**
BT Agrumi
- relazione partitiva (parte-tutto)
es. **Stomaco**
BT Apparato digerente
- relazione esemplificativa
es. **Natale**
BT Feste cristiane

N.B. Due termini possono essere legati da una relazione di gerarchia solo se appartengono alla stessa categoria

Thesaurus: Relazioni/3

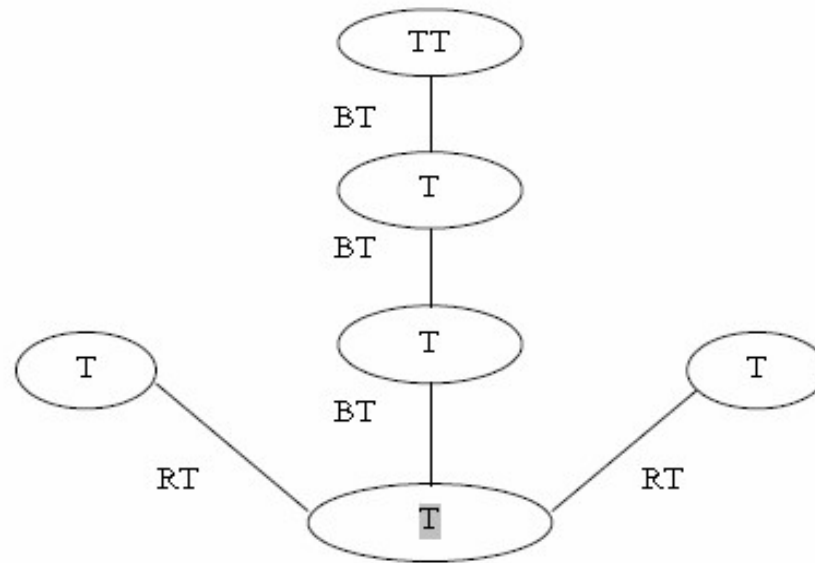
Relazione associativa:

- lega termini non legati da altra relazione se non dal fatto di rientrare uno nella definizione dell'altro
- in teoria non è reciproca, ma nel thesaurus sì, per motivi di struttura e di navigabilità
- i termini collegati non devono necessariamente appartenere ad una stessa categoria

es. **Bibliotecari**
RT Biblioteche
RT Scuole per bibliotecari

Thesaurus: Relazioni/4

Struttura ad àncora



Modello di rete "ad àncora"

Thesaurus: Relazioni/5

Poligerarchia temperata



Thesaurus: Relazioni/5

Poligerarchia temperata



Thesaurus: Relazioni/6

Poligerarchia temperata



Thesaurus: Relazioni/7

Etichette di nodo: termini o locuzioni la cui funzione è quella di indicare una caratteristica di divisione. Sono a tutti gli effetti termini di gerarchia che hanno sia BT che NT, ma non possono essere usati come termini di indicizzazione.

L'elaborazione delle etichette di nodo è stata operata a partire dai livelli più alti della gerarchia, concentrando l'attenzione sul primo livello e via via scendendo di grado, a seconda della necessità.

Thesaurus: Note

Note = informazioni aggiuntive riguardo al termine del vocabolario finalizzate a chiarire il suo significato, ove non sia sufficientemente comprensibile dal contesto fornito dalla gerarchia

- Nota di definizione
- Nota d'ambito (SN)
- Nota storica (HN)

Thesaurus: Note/2

Nota di definizione: contiene la definizione del termine

- Viene utilizzata raramente, solo per termini di ambito tecnico e settoriale o quando convivono nel thesaurus termini dal significato affine, ma non equivalenti
- Viene seguita dalla sigla del dizionario o repertorio da cui la definizione è tratta, data tra parentesi

es: **Push, Mifepristone, Prelature personali, Logica fuzzy**

Thesaurus: Note/3

Nota d'ambito: si utilizza per restringere o espandere il significato del termine, per distinguere due termini che hanno significati sovrapposti nel linguaggio naturale, oppure per fornire indicazioni sull'uso del termine

- Può contenere la definizione del termine, volta a chiarire il significato con cui il termine è acquisito nel vocabolario
- Può contenere una nota di orientamento che rinvia a termini che esprimono significati alternativi

es: **Comuni, Cinema, Bibliografia, Flora, Impianti tecnici, Corpi di polizia, Tartarughe**

Thesaurus: Note/4

Nota storica: garantisce il legame tra il thesaurus e gli strumenti della nostra tradizione italiana (*Soggettario* e suoi *Aggiornamenti*)

- Si utilizza per indicare che il termine scelto come preferito ha una diversa forma o significato rispetto agli strumenti preesistenti (*Soggettario* e *Aggiornamenti*)
- Nei casi di diversa forma è accompagnata da un legame con una **variante storica**
es. **Cuoi, Fisionomia nell'arte, Diafisi, Pi greco, Comuni**
- Può contenere un'indicazione sulla frequenza d'uso del termine sostituito negli strumenti preesistenti
es. **Fichi**

Thesaurus: Nomi propri

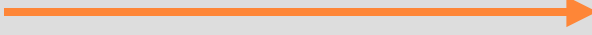
Definizione:

I nomi propri sono quelli che si riferiscono a individui, oggetti unici, realtà individuali per designarli e differenziarli da altri appartenenti alla loro stessa classe, senza che ne indichino necessariamente alcuna caratteristica dell'entità denominata

Naming principle (IFLA):

“in un linguaggio di indicizzazione la forma dei nomi di persona, dei luoghi, delle famiglie, degli enti e dei titoli di opere deve essere stabilita in accordo con i codici di catalogazione per autore e titolo usati nel medesimo catalogo”

Thesaurus: Nomi propri/2

- **Nomi con autorialità:**  Regole di catalogazione per autore
persone / famiglie / enti / titoli di opere
- **Nomi privi di autorialità:**
esseri immaginari / opere dello spettacolo / opere delle arti figurative / opere architettoniche / nomi geografici / oggetti / strumenti e forme / animali celebri / avvenimenti storici

Thesaurus: Nomi propri/3

Esempi:

nomi personali:	Francesco d'Assisi <santo>
nomi di famiglie:	Asburgo <casa>
nomi di enti:	Università degli studi di Udine
nomi di opere letterarie:	Verga, Giovanni. I Malavoglia
nomi di personaggi:	Cappuccetto rosso <personaggio>
nomi di opere artistiche:	Monumento a Giordano Bruno <Roma>
nomi di edifici:	Caffè degli specchi <Trieste>
nomi geografici:	Golfo di Trieste
nomi di oggetti:	Microsoft Excel <programma per elaboratori>
nomi di eventi:	Guerra civile spagnola <1936-1939>

Thesaurus: Nomi propri/4

- Rispetto del *naming principle* dell'IFLA
- Scelta della lingua:
 - Lingua originale per i nomi propri che per forma e contenuto si avvicinano ai nomi di competenza dell'indicizzazione per autori (uniformità)
 - Lingua italiana per i nomi propri peculiari della sola indicizzazione per soggetto o nei casi in cui la forma originale non sia accertabile o mediamente comprensibile
- Esclusione momentanea dei nomi propri dal Thesaurus, salvo alcune eccezioni, derivate dal *Soggettario*:
 - eventi storici (es. Guerra mondiale 1914-1918)
 - correnti culturali (es. Illuminismo)
 - strumenti (es. PRECIS)
 - prodotti alimentari (es. vini e formaggi)

Thesaurus: Nomi propri/5

Qualificazioni nei nomi propri:

- distinguono gli omografi (funzione valida per tutti i termini del vocabolario)
- facilitano l'identificazione dell'entità rappresentata garantendo coerenza nel suo utilizzo (funzione valida per i soli nomi propri)

Scelta della qualificazione:

- BT del termine dato al singolare
- Luogo in cui l'entità si trova
- Date per gli eventi

Thesaurus: Nomi propri/92

Riferimenti:

- **Guida**, Appendice A: Nomi propri
- **Manuale applicativo**, § D: Forma dei nomi propri (prassi della BNI)

Thesaurus: Fonti

- Strumenti catalografici attualmente impiegati per l'indicizzazione (*Soggettario e Aggiornamenti, DDC*)
- Repertori generali (dizionari della lingua italiana ed enciclopedie)
- Repertori specialistici di varie discipline
- Vocabolari controllati e thesauri nazionali ed internazionali

<https://thes.bnconfirenze.sbn.it/fonti.php?menuR=1&menuS=2>

Thesaurus: Fonti/2

Interoperabilità con le fonti:

- dizionari (De Mauro, Treccani, Zingarelli, etc.)
- Wikipedia
- Soggettario di Firenze
- altri thesauri (LCSH; AGROVOC; LIUC, etc.)
- il catalogo BNCf
- DDC

Thesaurus: Numero Dewey

Parametri per la scelta del numero Dewey:

- Numero interdisciplinare o d'insieme

es. **Corteggiamento** 306.734

- Numero che corrisponde ad un'unica sottointestazione del termine in Indice

es. **Ordine**
filosofia 117

- Numero con maggiore garanzia bibliografica BNI (nel caso di termini con più sottointestazioni in Indice)

es. **Moto**
filosofia 116
fisica 531.11
stelle 523.83

classe con
maggiore garanzia
bibliografica in BNI

Thesaurus: Numero Dewey/2

Vantaggi:

- Primo esperimento di interoperabilità fra strumenti diversi di indicizzazione
- Potenziamento di ricerca semantica
es. **Buddismo tibetano, Coppie non sposate**
- DDC come raggruppamento del patrimonio lessicale del Thesaurus