

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Facoltà di Psicologia e Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Specialistica in Teoria e Tecnologia della Comunicazione



IL TESAURO NON È UN DINOSAURO

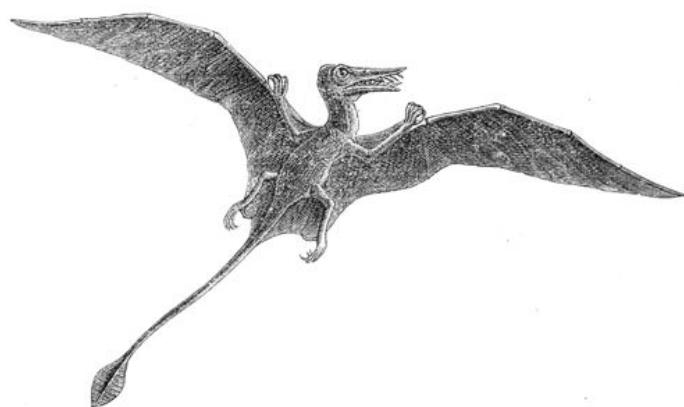
DI

DARIO RODIGHIERO
mat. 571070

Relatore: Gabriella Pasi

Correlatore: Carlo Ferigato

Controrelatore: Luca Bernardinello



Anno Accademico 2009/2010

Indice

Introduzione	1
I Presupposti	5
Linguaggi di indicizzazione controllati	5
Vocabolari controllati	7
Thesauri	8
Proprietà e definizioni di relazioni e grafi	12
Semantica	13
Sintassi	14
Pragmatica	16
Biblioteconomia e information retrieval	17
Biblioteconomia e ontologie	19
Il Contesto	21
Panorama attuale	21
Esempio nell'uso dei thesauri	22
SIRS	23
Ponto	25
Il Metodo	27
Analisi del problema	27
Thesaurus digitale e cartaceo	29
Nuovo soggettario	31
I Costituenti	33
Termine	33
Tipografia	34
Supporti ad alta definizione	35
Scelta del carattere	36
Lista	43
Adattare il testo alla visualizzazione	44

Lo Spazio	49
Schermo	49
Impaginare la lista	50
Dimensione del carattere	52
Daltonismo e accessibilità	52
Pagine	57
Le Caratteristiche del Termine	63
Preferiti e non preferiti	63
Selezione	64
Dettaglio	65
L'Ordinamento	69
Relazioni	69
Relazioni gerarchiche	71
Relazioni gerarchiche a più livelli	71
Relazioni gerarchiche a doppia direzione	71
Termini	76
Ordine alfabetico	76
Ordine alfabetico permutato	77
Ordine sistematico	77
Ordine cronologico	78
Ordine quantitativo	78
Dinamicità degli ordinamenti	78
Il Sistema Multi-Thesaurus	85
Uno, due, tre thesauri	86
Navigare le relazioni	87
Lavorare con le relazioni	88
Lo Spazio Personale	93
Muoversi nel percorso di navigazione	93
Segnalibri	94
La Vita di un Termine	95
Creazione	96
Caratterizzazione	96
Cancellazione	97
Ricerca	97
Conclusioni	101
Ringraziamenti	103

IL TESAURO
NON È UN DINOSAUR

Introduzione

La tesi ha l'obiettivo di progettare un'interfaccia per un sistema di gestione di thesauri e vocabolari controllati più in generale. Si vuole in particolar modo realizzare un lavoro innovativo, ovvero un'interfaccia grafica che ponga in discussione gli attuali paradigmi applicativi nel mondo della biblioteconomia e dell'information retrieval.

In Italia la biblioteconomia è una nicchia affascinante che ospita un mondo di professionisti. Per un periodo di tre anni ho avuto il piacere di fare parte del mondo delle biblioteche e la tesi vuole essere la conclusione di un'esperienza lavorativa, ma anche un apprezzamento alla comunità che mi ha consentito di crescere professionalmente.

Il Joint Research Centre è un organo di ricerca che fa parte della Commissione Europea. Il suo principale compito è quello di fornire supporto tecnico alle politiche della Commissione. Il JRC ha sei sedi sparse in Europa, a Ispra, località italiana in provincia di Varese, si trova la più grande nella quale trovano occupazione quasi tremila persone.

All'interno del JRC di Ispra esiste una biblioteca, la Central Library del JRC. Oltre a svolgere i compiti di una normale biblioteca all'interno del centro, essa ha il compito di integrare e di organizzare la conoscenza delle varie sedi sparse per l'Europa [3]. Il progetto al quale ho lavorato, il Scientific Information Retrieval System, è il risultato di uno sforzo mirato a fornire nuovi servizi bibliotecari agli utenti [18].

In particolare SIRS è stato creato sull'idea di integrare l'accesso a differenti risorse bibliografiche per mezzo di un'unica interfaccia, in modo da permettere alle persone che lavorano all'interno del Centro di avere una visione globale dello stato dell'arte della ricerca scientifica sia all'interno che all'esterno dell'istituto. Infatti le risorse comprendono non solo le pubblicazioni prodotte dal JRC, ma anche gli articoli sia cartacei che elettronici della ricerca scientifica nel mondo.

SIRS ha inoltre dato vita a un progetto parallelo chiamato Ponto [24]. Dal momento in cui ci si è resi conto della necessità di impiegare un sistema atto alla gestione simultanea di thesauri e data l'assenza di un prodotto simile sul mercato, si è scelto di creare *intra muros* un software *open-source* capace di gestire diversi thesauri contemporaneamente e le

rispettive relazioni extra-thesururus.

Su questa idea si è stretta una collaborazione con gli uffici della FAO a Roma, il British Geological Survey con sede a Nottingham e la Biblioteca Nazionale Centrale di Firenze. La collaborazione ha portato alla realizzazione di una struttura dati basata su standard internazionali [9]. All'interno di SIRS questa struttura è stata popolata e sono stati costruiti su di essa *web service* per estrarre i dati destinati a popolare l'interfaccia utente.

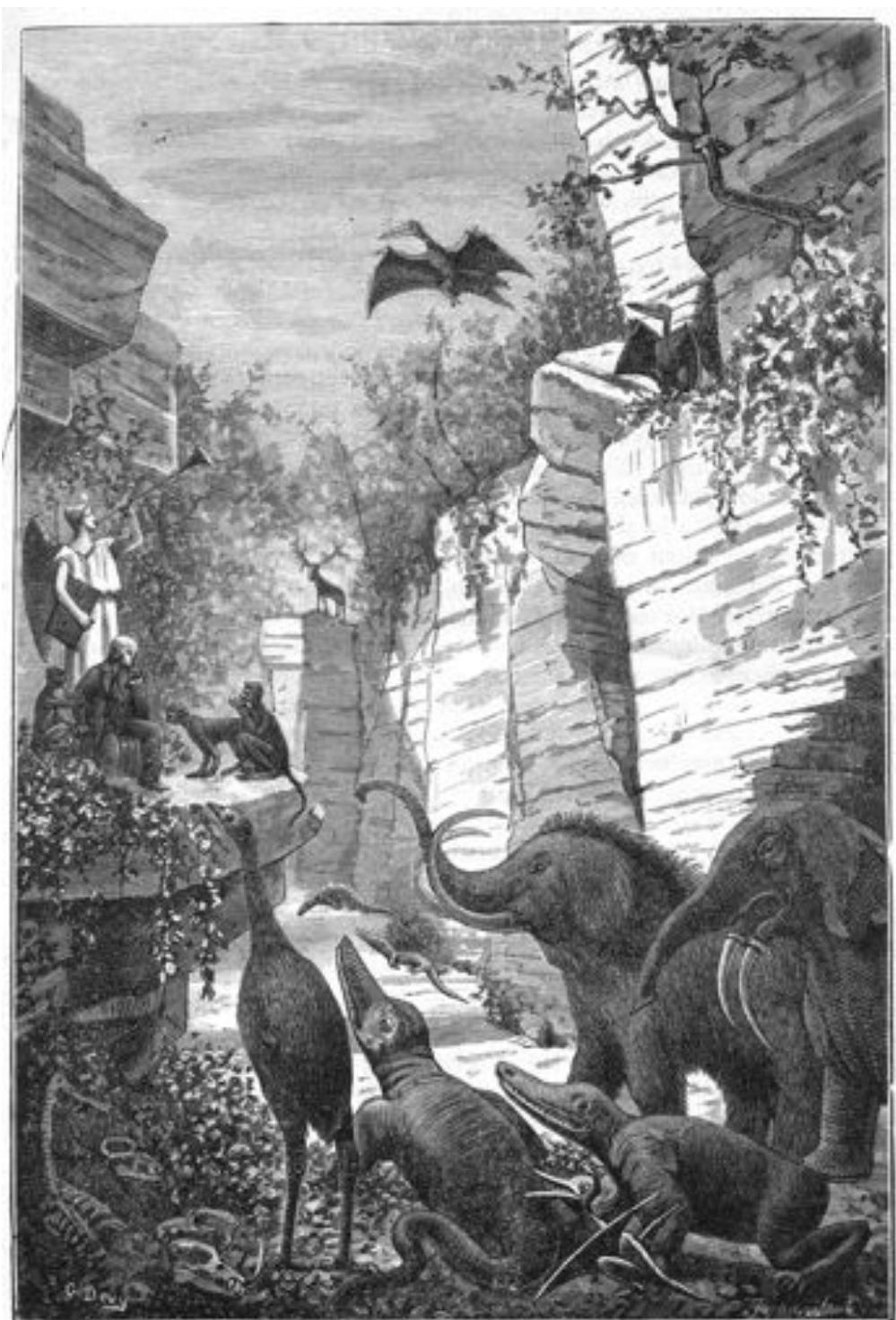
Questi due progetti sono la premessa di cui questa tesi sarà la naturale prosecuzione e conclusione di questa esperienza lavorativa che coincide con la fine del ciclo di studi in Teoria e Tecnica della Comunicazione, corso specialistico interdisciplinare organizzato dai dipartimenti di informatica e psicologia dell'Università degli Studi di Milano-Bicocca. La tesi si colloca così in un quadro che comprende sia l'attività lavorativa che quella accademica, utilizzando un linguaggio che travalica in più di un'occasione il confine tra i due mondi, quello di un'organizzazione governativa internazionale e quello dell'università.

L'argomento della tesi è la progettazione di un'interfaccia utente che eredita l'esperienza di SIRS e la riversa in Ponto. Si tratta di intraprendere il processo di design ricucendo la spaccatura creata tra i volumi dei vocabolari cartacei e le attuali interfacce utente basate su paradigmi troppo schematici e poco intuitivi. Durante la dissertazione si farà dunque riferimento all'interfaccia come all'interfaccia utente di Ponto, in quanto strumento finalizzato alla manipolazione di thesauri. In quest'ottica inoltre il lavoro di tesi potrà essere visto anche come documentazione *open-source*, cioè come un manuale per passare alla fase d'implementazione dell'interfaccia di Ponto.

Ai fini di una lettura corretta della tesi, si tiene a precisare che ciò che si vuole scrivere è un esercizio di progettazione e non d'implementazione. A questo punto si rende anche necessario dare una descrizione di progettazione, che si ispira al lavoro di Löwgren e Stolterman [32]. Dal loro punto di vista il *design* è un processo che comprende tre fasi: *Vision*, *Operative Image*, *Specification*. Queste fasi, che arrivano da un approccio umanistico alla progettazione di un'interfaccia, si differenziano tra loro in termini di specificità: con il progredire della progettazione e delle fasi il design diventa sempre più complesso e specifico. Ad un certo punto della fase *Specification*, il progetto diventa maturo per sviluppare il software. Questo punto sopraggiunge nell'ultima fase di progettazione e rappresenta il limite oltre il quale la tesi non si spingerà nel design di un'applicazione destinata a gestire i thesauri.

Per agevolare la lettura, la tesi è suddivisa in brevi capitoli qui descritti:

- *I Presupposti* introducono le basi teoriche sulle quali la tesi poggia.
- *Il Contesto* descrive la situazione attuale delle interfacce per la gestione di thesauri.
- *Il Metodo* spiega come si è affrontato il percorso di progettazione.
- *I Costituenti* sono gli elementi fondamentali che andranno a popolare l'interfaccia. Termini e liste saranno valutati dal punto di vista tipografico.
- *Lo Spazio* individua il supporto sul cui verranno visualizzati i costituenti.
- *Le Caratteristiche del Termine* mostrano i dettagli di un termine all'interno dell'interfaccia.
- *L'Ordinamento* è alla base della visualizzazione dei termini, permette all'utente di accedere ai termini in modo diverso.
- *Il Sistema Multi-Thesaurus* dimostra come sia possibile gestire le relazioni di uno o più thesauri contemporaneamente.
- *Lo Spazio Personale* permette all'utente di accedere a una serie di funzionalità atte a rendere proprio l'ambiente di lavoro.
- *La Vita di un Termine* offre alcuni spunti su come gli utenti andranno a lavorare sui termini.



I Presupposti

Qui introduciamo il contesto della tesi. La riflessione che nasce risulta fondamentale dal momento che verranno presentate le basi teoriche che delimiteranno lo svolgimento dell'argomentazione. I limiti serviranno da guida a un testo di carattere multidisciplinare, il cui fuoco si muoverà tra *informatica*, *interaction design* e *biblioteconomia*.

Se la relazione tra informatica e interaction design appare naturale quanto quella tra informatica e biblioteconomia, non si può dire la stessa cosa per le discipline di interaction design e biblioteconomia. A tal proposito non esistono grandi sperimentazioni nell'ambito, seppur nelle biblioteche si sia sempre dato rilievo alla consultazione dei documenti. Sia l'indicizzatore che l'utente nel passato hanno potuto lavorare con strumenti che sono sempre stati caratterizzati da cura e qualità nel dettaglio. Purtroppo, nel momento esatto in cui il cartaceo tramuta in digitale, questi valori si perdono nel passaggio ai nuovi media [35], punto dove vogliamo lavorare nella tesi. L'intento è di riprendere il filo storico spezzato alla dismissione degli strumenti cartacei, di cui rimangono eccellenti esempi tuttora, così da ripensare un'interfaccia utente che riversi l'esperienza del passato nei moderni strumenti informatici che abbiamo a disposizione.

Nel percorso che porterà a una rivisitazione dell'interfaccia utente per consultare e modellare i thesauri, un buon punto d'inizio è descrivere i fondamenti teorici.

Andremo qui a introdurre i *linguaggi di indicizzazione controllati* usati nelle biblioteche per indicizzare i documenti, descriveremo i *thesauri* ai quali la progettazione dell'interfaccia sarà particolarmente dedicata, analizzeremo i rapporti della biblioteconomia rispettivamente con l'information retrieval e le ontologie.

Linguaggi di indicizzazione controllati

L'avvento del *World Wide Web* ha creato una grande quantità di informazione condivisa, anche se spesso orientarvisi risulta difficoltoso. Gli

strumenti più diffusi per accedere all'informazione che vi risiede sono i motori di ricerca. Essi funzionano in questa maniera: un utente formula una richiesta formata da una serie di parole chiave e il motore di ricerca restituisce una lista di documenti, ordinati sulla stima di una loro rilevanza rispetto alla richiesta stessa. Senza addentrarci nel complicato mondo degli algoritmi dei motori di ricerca, questo modo di cercare informazione pone un grave problema. Se l'informazione cercata è posizionata all'inizio della lista in cui si presentano i risultati la richiesta avrà avuto successo, ma se così non fosse la dimensione del numero di documenti presenti nella risposta renderebbe difficoltosa una successiva ricerca sistematica tra i documenti.

Ben prima dell'avvento dei motori di ricerca, la biblioteconomia aveva trovato una risposta nei *linguaggi di indicizzazione controllati*, strumenti che permettono di migliorare l'indicizzazione dei documenti e il successivo processo di estrazione dell'informazione [54]. Sono linguaggi di organizzazione composti da *soggetti* la cui lunghezza varia a seconda delle esigenze, da una singola parola fino a descrizioni più complesse, e la cui specificità cambia in base al numero di documenti all'interno di una collezione: più documenti dovranno essere organizzati, più il lessico diventerà specifico.

L'indicizzare si basa sulla natura stessa dei linguaggi di indicizzazione: si tratta di creare un'associazione diretta tra documento e soggetto. I linguaggi possono diventare strumenti così curati e completi da trasformarsi in informazione a loro volta: attraverso la costruzione e il mantenimento di un gruppo di soggetti prende forma un *vocabolario controllato* che permette il reperimento diretto e indiretto di documenti, ma allo stesso tempo si crea una struttura semantica che può integrare l'*informazione* già resa disponibile in forma di documenti. Nello scambievole ruolo di strumento per il reperimento e di informazione stessa ha luogo la compenetrazione tra biblioteconomia e information retrieval. È infatti qui che i linguaggi a soggetto, appartenenti al mondo della biblioteconomia, diventano preziosi strumenti per l'information retrieval nell'indicizzazione automatica e semi-automatica di documenti e il successivo reperimento di informazioni [34].

I linguaggi di indicizzazione controllati si dividono in *linguaggi controllati basati su parole* e *organizzati alfabeticamente* e *linguaggi controllati organizzati sistematicamente* [54]. Essi si distinguono per caratteristiche, ma allo stesso tempo sono simili per esigenze. Per esempio i linguaggi controllati basati su soggetti e organizzati alfabeticamente, come le *Library of Congress Subject Headings*, sono caratterizzati da una visualizzazione dei soggetti in ordine alfabetico, e dall'uso di *authority list* e thesauri (vedi p.8). Invece i linguaggi controllati organizzati sistema-

ticamente come la *Dewey Decimal Classification*, fanno uso di indici per classi e di schemi di classificazione per l'ordinamento delle *notazioni*.

Nella letteratura biblioteconomica siamo abituati a parlare di *soggettari* e *sistemi di classificazione*. I soggettari appartengono ai linguaggi controllati basati su termini e organizzati alfabeticamente come le già citate *Library of Congress Subject Headings*. In contrapposizione abbiamo i sistemi di classificazione che corrispondono invece ai linguaggi controllati, formati da classi e organizzati sistematicamente, il cui più famoso esempio è la *Dewey Decimal Classification*.

Sebbene questa divisione sia ben distinta a livello teorico, nei fatti è molto più difficile classificare un linguaggio controllato: caratteristiche dei soggettari vengono usate nei sistemi di classificazione e viceversa. Per esempio nei soggettari capita che vengano applicate divisioni per categoria. Nei sistemi di classificazione invece quasi sempre viene impiegato un indice alfabetico, anche qui siamo di fronte a una forma di contaminazione.

Vocabolari controllati

L'evoluzione e l'ampliamento strutturale di biblioteche, archivi e collezioni ha portato allo studio e alla realizzazione di strumenti in grado di offrire un valido aiuto all'ordinamento del materiale, i *vocabolari controllati* nascono in questo contesto [27]. Essi derivano dal linguaggio naturale, ma differiscono da quest'ultimo in quanto sono semplificati e circoscritti da una serie di regole.

I vocabolari controllati obbligano all'uso di termini predefiniti. In contrasto col linguaggio naturale dove si usa una vasta moltitudine di parole, i linguaggi controllati basano la costruzione del linguaggio su insiemi definiti di termini semplici o composti che vengono scelti per definire un concetto. Questi termini verranno poi impiegati da indicizzatori e utenti nelle diverse fasi dell'organizzazione della conoscenza.

Il vocabolario controllato è una lista di termini che aiuta l'indicizzatore nel suo lavoro. Nell'atto di descrivere il contenuto di un libro, l'indicizzatore si avvale di questo strumento sintattico per costruire una descrizione che successivamente verrà impiegata dall'utente nella ricerca di documenti. L'impiego di un vocabolario controllato consente non solo all'indicizzatore di evitare l'uso di termini obsoleti, ma anche a fondare l'organizzazione su un linguaggio condiviso con i colleghi. Essendo il lavoro delle biblioteche un impegno quotidiano, gli indicizzatori avranno inoltre il ruolo di aggiornare costantemente la lista di soggetti in modo da mantenerla aggiornata rispetto al linguaggio cor-

rente. Si può immaginare che il linguaggio parlato, il materiale di una biblioteca e il relativo sistema di indicizzazione siano delle entità che lavorano in simbiosi col passare degli anni, un avvenimento legato alla vita di uno di essi si diffonderà naturalmente negli altri.

Thesauri

I vocabolari controllati impiegati nei linguaggi di indicizzazione possono raggiungere forme complesse, i *thesauri* ne sono un esempio. Jean Aitchinson definisce il thesaurus come un “vocabulary of controlled indexing language, formally organized so that *a priori* relationships between concepts are made explicit” [1]. Ispirati da questa definizione, affermiamo che i thesauri sono dei vocabolari controllati a supporto dell’indicizzazione, caratterizzati da un’organizzazione formale data da relazioni esplicite nella struttura stessa dello strumento (la semantica relazionale è spiegata più avanti a p. 13).

I thesauri sono usati sia nel processo di indicizzazione che nel reperimento di informazioni, ma il loro contributo può essere specifico a seconda delle esigenze:

- Indicizzazione e reperimento, quando il thesaurus è utilizzato sia in fase di indicizzazione che in fase di reperimento di documenti.
- Solo indicizzazione, il thesaurus qui è usato solo in fase di indicizzazione mentre il reperimento dei documenti avverrà con altre tecniche.
- Solo reperimento, il thesaurus può essere utilizzato per migliorare e completare la formulazioni delle *query*.

Oltre ad essere importante per l’information retrieval, il thesaurus è uno strumento fondamentale anche in biblioteconomia sebbene il suo utilizzo sia leggermente differente. Infatti, mentre in information retrieval il thesaurus è usato in maniera automatica e semi-automatica, in biblioteconomia esso viene impiegato in modalità manuale o semi-manuale. Inoltre è importante sottolineare che in biblioteconomia il thesaurus assume la massima potenzialità nel momento in cui indicizzazione e reperimento coincidono in una forma di intermediazione tra linguaggio utente e linguaggio di indicizzazione. In questo caso, l’attività di indicizzazione effettuata in un primo momento sarà in un secondo momento utile all’utente in fase di reperimento.

Relazioni

Il thesaurus è caratterizzato da tre tipi di relazioni tra termini, altrimenti dette relazioni *standard* [10]. Queste relazioni possono essere estese a relazioni più specifiche o affiancate da relazioni *non standard* [1]. Nei thesauri le relazioni standard sono raggruppate in tre categorie:

- Relazioni di equivalenza
- Relazioni gerarchiche
- Relazioni associative

Relazioni di equivalenza

Le relazioni di equivalenza servono a mettere in relazione tra loro termini con lo stesso significato o con significato simile. Essendo i vocabolari controllati nati per semplificare e snellire il linguaggio naturale, questo tipo di relazione permette di distinguere termini *preferiti* e termini *non-preferiti*, escludendo questi ultimi nella fase di indicizzazione. L'obiettivo è di indirizzare indicizzatore e utente all'uso dei termini preferiti e di evitare l'uso dei termini non-preferiti.

Le relazioni di equivalenza vengono impiegate in casi di sinonimia (es. *Cadetti di marina* e *Allievi ufficiali di marina*), quasi sinonimia (es. *Abiti* e *Vestiti*) e varianti lessicali (es. *Abbreviature* e *Abbreviazioni*). La relazione *Use for* (UF) indica che il termine corrente è il termine preferito, ma suggerisce altri termini non utilizzati per indicizzare che sono utili per descrivere il contesto del termine preferito. La relazione *Use* (USE) indica invece che il termine corrente non è utilizzato per indicizzare, invitando a usare il termine preferito corrispondente.

Per esempio il termine *Nani da giardino* ha come relazione UF *Nanetti da giardino*, che indica il termine non preferito. Viceversa il termine *Nanetti da giardino* ha come relazione USE *Nani da giardino*, che indica il termine preferito da usare.

Relazioni associate a <i>Nani da giardino</i>	
UF	Nanetti da giardino
TT	Oggetti
BT	Statue
RT	Giardini

Relazioni gerarchiche

La relazione che distingue maggiormente il thesaurus dai vocabolari controllati è la relazione gerarchica, esemplificata nelle relazioni di

Broader Term (BT) e *Narrower Term* (NT). Tutti i termini preferiti di un thesaurus sono organizzati attraverso relazioni di tipo padre-figlio, attraverso cui il thesaurus assume la caratteristica forma ramificata che lo contraddistingue.

Un termine preferito, che non sia radice o foglia dell'albero, generalmente è caratterizzato da un padre e diversi figli. Tuttavia esistono delle eccezioni per cui un termine può avere più di un legame BT. Quando i termini di un thesaurus hanno al massimo una relazione BT parliamo di thesauri *monogerarchici*. Se invece il thesaurus ammette relazioni BT multiple, allora parliamo di thesauri *poligerarchici* [39].

Per esempio il termine *Nani da giardino* ha come unica relazione gerarchica BT *Statue*, in questo caso il primo termine specifica il significato del secondo. Simmetricamente il termine *Statue*, mostra tra le sue relazioni gerarchiche NT *Nani da giardino*.

Relazioni associate a <i>Statue</i>	
TT	Oggetti
BT	Sculture
NT	Nani da giardino
	Statue equestri
	Statue votive
	Statuette
RT	Statue egizie

Come accade nel Nuovo soggettario, in alcuni thesauri è possibile trovare la relazione Top Term che si indica con TT. Normalmente viene introdotta quando c'è la necessità di creare un collegamento da un termine al relativo termine padre, come se si trattasse di un super BT che permette di risalire al nodo padre. Nel caso specifico del Nuovo soggettario le relazioni TT corrispondono alle faccette e aiutano nell'organizzazione strutturale dei termini del thesaurus

Relazioni associative

Le relazioni associative attivano passaggi tra sezioni diverse del thesaurus. La relazione, che in inglese prende il nome di *Related Term*, permette a chi costruisce il thesaurus di creare relazione tra termini che condividono argomenti comuni ma che non appartengono alla stessa alberatura. L'uso che ne deriva è molto importante sia in fase di indirizzazione, perché permette di confrontare soggetti diversi per scegliere il più adatto, che in fase di reperimento, lasciando all'utente la libertà di spaziare su temi relativi allo stesso argomento.

Per esempio il termine *Nani da giardino* ha tra le sue relazioni **RT** *Giardini*. Essendo la relazione simmetrica, il termine *Giardini* avrà come relazione associativa *Nani da giardino*.

Relazioni associate a <i>Giardini</i>	
TT	Spazio
BT	[Aree ricreative]
NT	Giardini all'inglese Giardini all'italiana
RT	Cancellate Città-giardino Giardinaggio Grotte artificiali Mobili da giardino Monumenti nazionali Nani da giardino Padiglioni Panchine Parchi giochi Pergolati Piante ornamentali Viali Zone verdi

Relazioni estese

Le *relazioni standard* qui elencate sono linee guida e pertanto possono essere adattate alla specificità delle situazioni. Tra le decisioni che un costruttore di thesaurus può attuare, vi è la possibilità di estendere una relazione a relazioni più specifiche così da declinare le relazioni introdotte estendendo il significato di base.

La relazione di **BT** viene normalmente declinata in tre modi [39]:

- Generica
- Parte-tutto
- Esemplificativa

La relazione generica definisce un collegamento tra genere e specie. Un esempio può essere dato da *Alcaloidi* e *Narcotina* dove la prima è padre della seconda in un legame generico. La relazione parte-tutto

estende un termine come parte della classe. Per esempio *Motore* e *Pistone*, essendo il secondo termine parte del primo. La relazione esemplificativa esprime un'istanza. *Città* e *Milano* sono un esempio di questo tipo di relazione, dove il secondo termine è un'istanza del primo.

Proprietà e definizioni di relazioni e grafi

Le relazioni e le proprietà delle relazioni appartenenti ai thesauri si possono tradurre in linguaggio logico.

Definiamo \mathcal{I} come l'insieme dei termini di un thesaurus. Una struttura relazionale su \mathcal{I} può essere definita così: $\langle \mathcal{I}, \mathcal{R} \rangle$ dove gli elementi di \mathcal{R} sono relazioni binarie $\text{NT}, \text{RT}, \dots, \text{USE} \subseteq \mathcal{I} \times \mathcal{I}$. Quando due termini indice i e j sono correlati attraverso una relazione, per esempio NT , in \mathcal{R} , scriviamo $\langle i, j \rangle \in \text{NT}$ e diciamo che i è NT rispetto a j .

Dopo aver dato la definizione di relazione, introduciamo quelle particolari proprietà di cui godono le relazioni all'interno di un thesaurus.

Sia R una relazione in \mathcal{R} , diciamo che R è:

- riflessiva quando $\forall i \in \mathcal{I} \langle i, i \rangle \in R$
- non riflessiva quando $\exists i \in \mathcal{I} \langle i, i \rangle \notin R$
- antiriflessiva quando $\forall i \in \mathcal{I} \langle i, i \rangle \notin R$
- simmetrica quando $\langle i, j \rangle \in R \Rightarrow \langle j, i \rangle \in R$
- non simmetrica quando $\langle i, j \rangle \in R \not\Rightarrow \langle j, i \rangle \in R$ o allo stesso modo $\exists \langle i, j \rangle \in R$ tale che $\langle j, i \rangle \notin R$
- antisimmetrica quando $\langle i, j \rangle \in R \Rightarrow \langle j, i \rangle \notin R$
- transitiva quando $\langle i, j \rangle, \langle j, k \rangle \in R \Rightarrow \langle i, k \rangle \in R$
- non transitiva quando $\langle i, j \rangle, \langle j, k \rangle \in R \not\Rightarrow \langle i, k \rangle \in R$ o allo stesso modo $\exists i, j, k \in \mathcal{I}$ tale che $\langle i, j \rangle, \langle j, k \rangle \in R$ e $\langle i, k \rangle \notin R$
- intransitiva quando $\langle i, j \rangle, \langle j, k \rangle \in R \Rightarrow \langle i, k \rangle \notin R$

Le relazioni del thesaurus verranno così definite:

- NT è una relazione *antiriflessiva*, *antisimmetrica* e *transitiva*
- BT è la relazione inversa a NT . $\langle i, j \rangle \in \text{BT} \Leftrightarrow \langle j, i \rangle \in \text{NT}$
- RT è normalmente interpretata come una *simiglianza*, è *riflessiva*, *simmetrica* e *non transitiva*

- USE è normalmente interpretata come una relazione *antiriflessiva*, *antisimmetrica* e *intransitiva*
- UF è l'opposto della relazione USE

Semantica

In questa sezione ci occuperemo di come lo strumento thesaurus si re-laziona con la semantica, quella disciplina che si occupa del significato e in particolare dell'interpretazione delle espressioni in un sistema lin-guistico [12]. Assunto un qualsiasi sistema di indicizzazione come lin-guaggio, allora possiamo definire l'ambito della semantica in tre precise forme [54]:

- Semantica relazionale
- Semantica referenziale
- Semantica delle categorie

Semantica relazionale

La semantica relazionale nel thesaurus si manifesta attraverso le rela-zioni USE/UF, BT, NT e RT (vedi p. 9). Ci sono due momenti nei quali si può vedere l'interazione tra il linguaggio naturale e il linguaggio formalizzato del thesaurus. Nelle fasi di creazione e manipolazione del thesaurus è l'indicizzatore che modifica la struttura. Le relazio-ni in questa fase vengono inserite, cancellate o modificate, traducendo in azione la semantica del linguaggio naturale, del contenuto dei do-cumenti e dell'interpretazione dell'indicizzatore che viene riversata nel thesaurus sotto forma di relazioni termine-termine. Viceversa la seman-tica costruita all'interno del thesaurus viene estratta nei momenti di in-dicizzazione e reperimento dell'informazione, quando indicizzatore e utente attingono informazioni per completare le proprie attività.

Semantica referenziale

La semantica referenziale indica il collegamento esistente tra un termi-ne e i documenti, definendo la spaccatura tra i significati di concetto e di termine. La parola concetto deriva dall'espressione latina *cum capio*, che significa raccolgo, prendo assieme. Infatti un concetto è un'entità che porta in sé la definizione dell'insieme di idee, cose o persone che rappresenta. Quando si usa la parola concetto, ci si vuole riferire a un

gruppo di cose che condividono proprietà comuni che le caratterizzano. Un concetto è quindi un'unità di pensiero logico che porta in sé la definizione di ciò che rappresenta. Al contrario il segno [25] [36] non porta con se nessuna definizione, la definizione è data dal suo utilizzo così come Wittgenstein sosteneva che il significato delle parole stava nel loro utilizzo, nella lingua parlata, nel dialogo e non in una definizione scritta [59]. Quindi il significato di un termine è dato dall'indicizzatore nel momento in cui usa un termine, sia nel momento dell'inserimento di un termine nel vocabolario controllato che nella scelta di un termine in fase di indicizzazione. Partendo dal presupposto che il linguaggio è comunicazione, anche l'utente – quando usa un termine per cercare del materiale – partecipa attivamente a definire il significato del termine usato. In questo contesto la semantica referenziale è presente in tutti gli usi che si fanno dei termini che coinvolgono i documenti, come per esempio nell'indicizzazione o nella ricerca di un libro.

Semantica delle categorie

Infine abbiamo la semantica delle categorie. Se apriamo una grammatica, possiamo vedere come le parole vengano suddivise in nomi, verbi, aggettivi, congiunzioni, etc. Anche nei vocabolari controllati esiste una divisione in categorie, anche se diversa. Per esempio le *Library of Congress Subject Headings* [31] suddividono i termini in argomento e posizione. Prendiamo invece la *Colon Classification* di Ranganathan [45], essa suddivide i termini in personalità, materia, energia, spazio e tempo, introducendo il concetto di *faccette*. L'uso di queste faccette non è semplicemente quello di creare degli insiemi ben distinti di termini, ma di usare i termini stessi per costruire stringhe di soggetto secondo delle regole precise. Le categorie nel nostro discorso servono a introdurre la sintassi delle stringhe di soggetto.

Sintassi

Il *New Oxford American Dictionary* [37] definisce la *sintassi* come “the arrangement of words and phrases to create well-formed sentences in a language”, ovvero si definisce la sintassi come quell’organizzazione, implicita ed esplicita, che permette la costruzione corretta di una frase di senso compiuto.

In biblioteconomia e nell’ambito dell’information retrieval il ruolo della sintassi viene applicato al soggetto. Non tutti i sistemi informativi di cui ci stiamo occupando utilizzano sintassi, anzi nell’information retrieval si parla specificatamente di analisi lessicale [20]. Tuttavia una

parte di loro, considerevole in termini di importanza, ne fa uso. In quest'ambito vengono usate delle regole sintattiche allo scopo di costruire stringhe di soggetto costituite da termini appartenenti ad un vocabolario controllato.

È importante sottolineare quanto la sintassi sia strettamente collegata alla semantica delle categorie vista nella sezione precedente. Applicando un semplice parallelismo tra linguaggio naturale e linguaggi di indicizzazione controllati, possiamo riflettere su come un'espressione del tipo "Danilo fuma" ci trasmetta la presenza di un *soggetto* e un *verb*. Questo significa che, per costruire ma anche comprendere la frase in maniera corretta, abbiamo bisogno sia di conoscere a quale categoria le parole appartengano, sia il significato della sequenza.

Nei linguaggi di indicizzazione controllati le categorie e le regole di costruzione assomigliano a quelle del linguaggio naturale. Julius O. Kaiser fu il primo a creare questa corrispondenza [26]. Nel linguaggio ideato da Kaiser, per esempio esistono due categorie – *concrete* e *process* – che possono essere combinate per creare una descrizione nella seguente forma, *concrete:process* [54]. Il linguaggio di Kaiser non è unico, altri sistemi di indicizzazione attualmente in uso come la Colon Classification [45] o il Nuovo soggettario [7], usano regole sintattiche per la costruzione di stringhe di soggetto.

È interessante quindi capire che ruolo assume la sintassi nel processo di indicizzazione. Immaginiamo di dover indicizzare un libro con un insieme di termini e decidiamo che due di loro facciano al caso nostro: *gatto* e *padrone*. In un sistema di indicizzazione senza sintassi assoceremo i due termini al documento, prima l'uno e allo stesso modo l'altro. In questo caso il libro indicizzato parla di gatti e padroni. Ipotizziamo ora di applicare una regola sintattica del tipo *soggetto:genitivo*, seguendo questa regola sarà possibile creare una definizione più precisa di ciò che il libro contiene, *padroni:gatti*, ovvero un libro il cui argomento è padroni di gatti.

Già da questo esempio è possibile comprendere la complessità e il lavoro nascosto sotto la parola sintassi. Complessità e lavoro significano grandi costi di realizzazione, ma in questo ambito esistono sistemi che si fanno forti delle potenzialità dei computer in sistemi di indicizzazione automatici e semi-automatici. Forse il più pregevole esempio nell'ambito della biblioteconomia è rappresentato PRECIS, sistema di indicizzazione ideato da Derek Austin negli anni '70 e utilizzato dalla British Library per circa un ventennio [4].

Linguaggi pre-coordinati e post-coordinati

Quando si crea un vocabolario controllato si deve fare attenzione alla scelta dei termini, dato che andranno a costituire la base del linguaggio di indicizzazione. In particolare va scelta con cura la lunghezza dei termini.

Prendiamo l'esempio di due soggetti: *Storia* e *Italia*. Nel vocabolario controllato possono essere introdotte separatamente, in questo caso *Storia* e *Italia* saranno due termini differenti che verranno inseriti nel thesaurus. In alternativa possiamo introdurre il termine *Storia d'Italia*. Le due alternative illustrano come prendono forma un linguaggio *post-coordinato* e un linguaggio *pre-coordinato*.

In un linguaggio post-coordinato i termini sono solitamente più brevi, più veloci da creare ed è necessaria una specializzazione minore per la creazione. Questo accade perché la fase di coordinazione, quando si uniscono due termini in una stringa di soggetto, non viene affrontata nel momento della creazione di un termine, bensì in fase di indicizzazione.

In un linguaggio pre-coordinato, la trasformazione di parole in termini viene fatta prima e quindi richiede una maggiore riflessione. Questo modo di lavorare è più lento, porta a una maggiore lunghezza dei termini ed implica un maggiore lavoro e una più approfondita conoscenza in fase di realizzazione [49]. La scelta tra un approccio e l'altro è data ai costruttori del thesaurus ed è una scelta molto delicata, perché andrà ad influenzare il lavoro di indicizzazione e creazione di nuovi termini.

A seguito di quanto detto è interessante vedere come si posiziona l'informazione nello spazio tra sistemi pre-coordinati e sistemi post-coordinati. La differenza risiede nella struttura del linguaggio di indicizzazione. Nei sistemi post-coordinati abbiamo un vocabolario controllato, i cui termini preferiti sono usati come parole chiave dall'utente in fase di recupero dell'informazione. Viceversa nei sistemi pre-coordinati l'informazione è composta da un vocabolario controllato e un archivio di stringhe di soggetto.

Pragmatica

La definizione di pragmatica che arriva ai giorni nostri è data dal filosofo americano Charles Morris [38]. Egli parte dal presupposto che il mondo della semiotica sia composto dalla semantica (la relazione tra segni e oggetti per i quali i segni sono utilizzabili), dalla sintassi (la relazione formale tra segni) e dalla pragmatica, che egli definisce come lo studio della relazione tra segni e interpreti.

Diversi ambiti e definizioni sono stati assegnati alla pragmatica, a noi interessa come la pragmatica del linguaggio definisca l'uso e l'applicazione dello stesso [54]. Per noi pragmatica prende il significato delle persone che interagiscono con il thesaurus attraverso il linguaggio di indicizzazione: le dinamiche a livello terminologico, nella struttura, tra le persone direttamente e indirettamente coinvolte nella costruzione del vocabolario controllato e gli utenti che accedono ai documenti attraverso il thesaurus. In pratica tutte le dinamiche che contribuiscono alla formazione e alla maturazione di un linguaggio comune.

Biblioteconomia e information retrieval

Il thesaurus nasce come supporto alla scrittura. Il più famoso esempio è sicuramente il *Roget's Thesaurus* che ancora oggi è possibile trovare in libreria [52]. Solo successivamente il thesaurus diventa strumento di aiuto all'indicizzazione, consolidando la sua forma in biblioteconomia. La disciplina informatica dell'*information retrieval* riprende questo strumento già maturo e lo riutilizza adattandolo alle proprie esigenze [27].

L'information retrieval nasce negli anni '60 quando inizia a occuparsi di rappresentazione, memorizzazione, organizzazione e accesso a documenti [5]. Ma se questi obiettivi sono condivisi con la biblioteconomia, dove va cercata la differenza tra le due discipline? Biblioteconomia e information retrieval differiscono innanzitutto per nascita. Mentre la prima è una scienza antica nata per ordinare oggetti fisici in uno spazio reale, la seconda è una giovane disciplina che permette l'accesso automatico a documenti digitali appartenenti a ingenti collezioni [57]. La forte connotazione delle origini si traduce in un modo diverso di perseguire lo scopo: nel caso della biblioteconomia l'obiettivo è l'utilizzo effettivo di biblioteche e archivi, mentre in information retrieval si cerca di rispondere nel modo più completo possibile a una *query* formulata da un utente. Con l'introduzione del digitale, il percorso della biblioteconomia non confluisce nell'information retrieval, le due discipline rimangono separate. Tuttavia strumenti comuni come il thesaurus, indicano che esiste un legame che potrebbe portare le due discipline alla collaborazione e alla definizione di basi condivise.

Information retrieval e biblioteconomia trovano nel thesaurus uno strumento comune. Sebbene usato in modi diversi, questo tipo di vocabolario controllato ricopre funzionalità importanti in entrambe le discipline [28]. In biblioteconomia il thesaurus è un indice che viene usato dall'indicizzatore per ordinare i documenti, associando a ogni docu-

mento una serie di termini appartenenti al thesaurus. In un momento successivo gli utenti della biblioteca potranno accedere ai documenti attraverso i termini usati nell'indicizzazione. Nell'information retrieval il thesaurus invece è usato in due modi: in un primo caso offre un aiuto all'utente nella formulazioni delle *query*, nel secondo è usato per migliorare e completare il reperimento come nel caso della *query expansion* [34].

Query expansion

Nell'ambito dell'information retrieval la tecnica di *query expansion* si applica per perfezionare le *query* degli utenti. A tale scopo si può fare uso di un thesaurus [43].

Immaginiamo un utente che esegue una *query* digitando una sequenza di *keyword*. La *query* può venire arricchita con termini aggiuntivi estratti da un thesaurus con lo scopo di perfezionare la risposta: le relazioni di equivalenza possono espandere la ricerca a termini sinonimi, le relazioni associative a documenti attigui e le relazioni gerarchiche a generalizzare una richiesta attraverso il BT o a specificarla nel caso di un NT.

La tecnica della *query expansion* va comunque applicata con riguardo. Se da un lato la risposta in termine di documenti è più ricca aumentando così il valore di *recall*, dall'altro il rischio è quello di abbassare la *precision*, rendendo la risposta talmente ampia da non permettere all'utente di trovare documenti rilevanti in tempi ragionevoli.

L'applicazione progettata

Creare e mantenere un thesaurus è un processo oneroso sia in termini di tempo che di costi, avere lo strumento adatto per lavorare può rendere più produttivi gli sforzi. L'applicazione che andremo a progettare sarà adeguata all'uso sia in biblioteconomia che in information retrieval. In particolar modo per la biblioteconomia sarà uno strumento in grado di aiutare chi gestisce la costruzione e il mantenimento di thesauri. In information retrieval l'applicazione sarà più utile in fase di lettura e controllo come nel caso di thesauri creati automaticamente da una base documentale (*pseudo-thesauri*).

L'obiettivo è progettare un'applicazione utile ad entrambe le discipline nella creazione e nel mantenimento di un thesaurus. Dato che l'information retrieval sarà più interessato alla parte di gestione, questo applicativo risulterà più mirato ad un uso di biblioteconomia che sfrutterà la complessità di tutte le sue caratteristiche.

Biblioteconomia e ontologie

In questi ultimi anni nell'ambito dell'informatica si è parlato molto di ontologie e svariati articoli sono stati presentati in letteratura [55]. Dal momento che in questi ambiti anche i thesauri vengono trattati come ontologie, vogliamo distinguere la forma del thesaurus nelle due discipline in modo da evitare qualsiasi tipo di confusione.

Il thesaurus è una struttura definita da un numero standard di relazioni che eventualmente possono essere estese e declinate in funzione della loro applicazione più o meno specifica (vedi p. 9). L'ontologia informatica si basa come il thesaurus sul modello di una tripla, collegando due elementi attraverso una relazione. Al contrario del thesaurus, l'ontologia non ha un numero standard di relazioni, anzi le relazioni declinabili all'interno di un'ontologia sono infinite.

Tale numero di relazioni è giustificato dal fatto che l'ontologia mira ad essere un modello nel quale è descritta la realtà delle cose, la rappresentazione dell'informazione. Al contrario, nel thesaurus, l'informazione non è contenuta nella struttura, ma nell'estensione stessa rappresentata dalla parte documentale. Mentre le ontologie si esauriscono in sé stesse, nei thesauri il significato è nell'estensione documentale. Le ontologie sono infatti composte da concetti che coincidono con i termini, nei thesauri i concetti sono i gruppi di sottoinsiemi possibili nell'insieme dei documenti la cui intensione è individuata nei termini come illustrato nella figura 1 (termini e concetti sono descritti a p. 13).

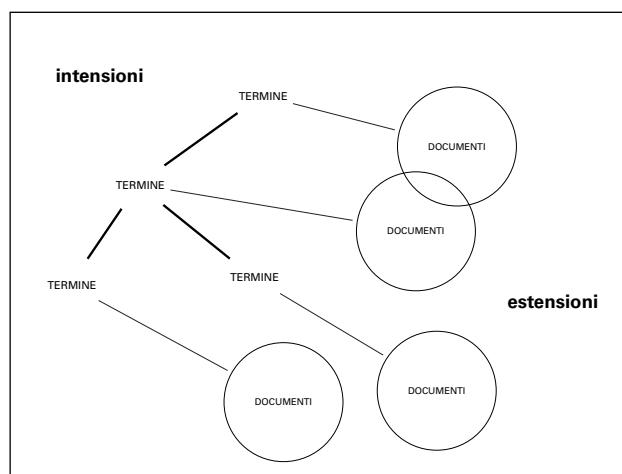
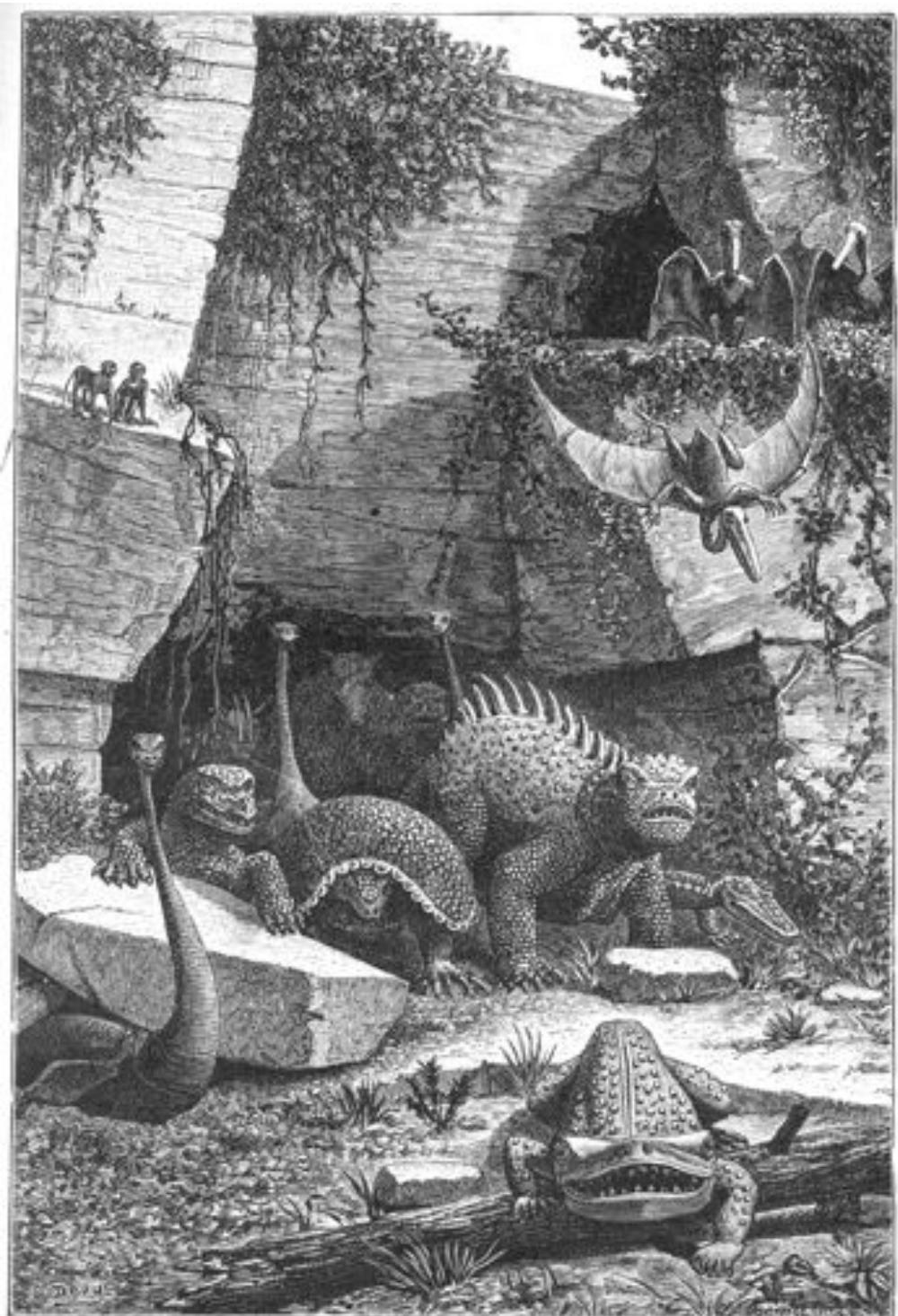


Figura 1: Il significato dei termini risiede nelle relazioni e nei documenti sottesi, in un frapporsi fra intensioni (termini) ed estensioni (documenti) [18]



Il Contesto

Capire in quale contesto si inserisce la tesi è importante. Per farlo analizzeremo prima il rapporto della biblioteconomia con l'informatica e successivamente tratteremo come l'esperienza lavorativa ci abbia portato a scegliere questo argomento di tesi.

Panorama attuale

I bibliotecari sono stati tra i primi a capire l'importanza degli strumenti informatici nel loro ambito. Se si guarda al passato fra le prime applicazioni troviamo il MARC (MAchine-Readable Cataloging), uno standard informatico per lo scambio dei dati tra le biblioteche [29]. Grazie a questo protocollo i catalogatori potevano condividere i record bibliografici con i colleghi e gli utenti erano in grado di fare ricerche su tutte le biblioteche collegate alla rete. MARC è stato il predecessore dello Z39.50, un protocollo informatico client-server atto al reperimento di documenti via rete [30].

La stessa importanza che è stata data da subito al mezzo informatico, è sempre stata riposta anche nei vocabolari controllati. Un'edizione cartacea di un sistema di indicizzazione, un thesaurus o un soggettario, è sempre stata ben curata. La prima sensazione che si ha sfogliando uno di questi libri è che sia stato progettato per essere usato. Lo si capisce dalle introduzioni, dall'impaginazione dei termini e dalle regole che spiegano come utilizzarli, dalle appendici che si integrano per completare il lavoro. La *Dewey Decimal Classification* ne è uno splendido esempio [13].

Nell'era dell'informatica il thesaurus è stato trasformato in digitale, passaggio che ha migliorato lo strumento in termini di accessibilità, condivisione e manutenzione. Tuttavia gli informatici che hanno sviluppato le interfacce utente per la traduzione di questi sistemi in digitale non vi hanno dedicato abbastanza interesse, per cui il media cartaceo è sempre stato superiore al digitale in termini di qualità nella visualizzazione dei dati.

Nel 2008, all'interno del progetto SIRS [18], abbiamo svolto un'indagine per la scelta di un software adatto alla gestione di thesauri. Sono stati valutati tutti i software disponibili sul mercato [51], attingendo specialmente alla lista stilata da due consulenti nel campo dell'*information management*, Leonard e Shena Will [58]. L'indagine ha portato al seguente risultato: nessuno dei software esistenti copriva le funzionalità richieste, in particolar modo quelle legate alla gestione di più thesauri contemporaneamente (sistema multi-thesaurus). Inoltre le interfacce utente di questi software risultavano spesso banali, prive di gusto tipografico e poco intuitive.

Nel 2009, in occasione della stesura di un articolo sull'interfaccia di SIRS [50], ci siamo confrontati con il mondo dell'information retrieval. Abbiamo preso spunto da un libro di Baeza-Yates e Ribeiro-Neto che comprende un capitolo totalmente dedicato alle interfacce utente nell'ambito dell'information retrieval [5]. Il confronto scaturito tra il campo della biblioteconomia e quello dell'information retrieval ci ha fatto riflettere sulla frattura creata nell'ambito delle interfacce tra queste due discipline. L'information retrieval pur essendo un disciplina giovane sta evolvendo in fatto di sistemi di recupero dell'informazione, tuttavia la sensazione è che ai sistemi attuali di gestione dei thesauri manchi una connessione, un'adeguata evoluzione dei sistemi precedenti, in particolare per quanto riguarda il mondo delle interfacce legate alla gestione dei thesauri. Per esempio, il gesto di sfogliare una versione cartacea della Dewey Decimal Classification dà una sensazione diversa rispetto la consultazione della versione elettronica [42] sebbene quest'ultima sia molto più comoda in termini di accessibilità. A parere nostro questo divario dipende da un'eccessiva generalizzazione delle interfacce-uomo macchina. L'impressione è che le interfacce grafiche dei sistemi informatici siano troppo uguali l'una all'altra, penalizzando le caratteristiche distinctive dei dati che visualizzano. Quello che cercheremo di fare in questa tesi è ridare personalità ai sistemi di gestione dei thesauri attraverso la creazione di un'interfaccia uomo macchina ispirata alla vecchia forma dei thesauri, quella delle edizioni cartacee.

Esempio nell'uso dei thesauri

Se questa tesi è stata scritta, uno dei motivi principali arriva dall'esperienza lavorativa. Nel 2005, per conto della Commissione Europea, è iniziato un progetto di ricerca nell'ambito dell'information retrieval. Il lavoro si è svolto in Italia nella sede di Ispra del Joint Research Centre.

L'obiettivo era quello di dotare la Central Library del JRC di un nuovo servizio, capace di agevolare il lavoro delle persone del centro. Il nuovo servizio doveva essere in grado di raggruppare diverse fonti di informazione in un'unica interfaccia, così da avere una vista completa sul lavoro di ricerca condotto al JRC.

La risposta a questa necessità è stata SIRS, Scientific Information Retrieval Service. Un sistema informativo capace di sfruttare la tecnica del mapping fra thesauri per integrare diverse sorgenti d'informazione ed offrirle all'utente in tempo reale. L'interfaccia di SIRS è coadiuvata da un'applicazione per la gestione di più thesauri contemporaneamente, *Ponto*, sviluppata in collaborazione con la FAO, il British Geological Survey e la Biblioteca Nazionale Centrale di Firenze.

SIRS

SIRS, Scientific Information Retrieval Service, è un sistema informativo capace di integrare diverse fonti di dati attraverso la tecnica del *mapping* tra indici. Il progetto nasce dall'esigenza del JRC di integrare tre diverse sorgenti di informazione che caratterizzano l'ambiente lavorativo della Commissione Europea a Ispra. Queste tre basi di dati sono:

- Azioni (organizzazione interna in temi di lavoro)
- Legislazione Europea
- Pubblicazioni del JRC

Rispettivamente indicizzate attraverso:

- Tassonomia delle Azioni
- EUROVOC
- Dewey Decimal Classification

L'idea di base era di creare un *mapping* tra i tre diversi indici, in modo da creare un canale di comunicazione che mettesse in collegamento le tre sorgenti di dati così da avere un'interfaccia popolata da informazioni contestualizzate. La tecnica del mapping consiste nell'unire due thesauri attraverso relazioni extra-thesaurus tra termini (vedi p. 85). Per mettere in atto questa idea è stata presa la decisione di visualizzare l'informazione in un'unico spazio in modo da rendere evidente l'attività all'interno del JRC. L'idea era infatti di rappresentare indici e documenti in un'unica interfaccia di information retrieval [50].

Lo spazio dell'interfaccia è stato diviso in due parti, una parte per visualizzare gli indici e una per i documenti. Nella figura 2 è possibile vedere la successione di suddivisioni che ha portato al progetto dell'interfaccia finale dove indici e documenti sono organizzati in base al numero delle sorgenti d'informazione (nel nostro caso sono tre).

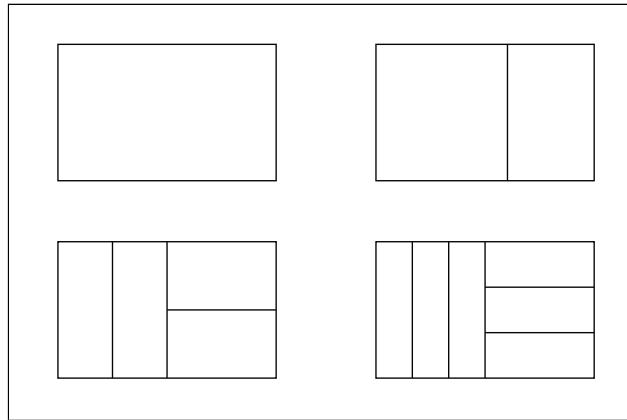


Figura 2: La successione di divisione di spazi che ha portato alla creazione dell'interfaccia utente.

Successivamente è stata introdotta la possibilità di personalizzare gli spazi dell'interfaccia in base alle esigenze dell'utente, ovvero mettere a disposizione dell'utente una serie di funzionalità per modificare gli spazi a proprio piacimento. Da un'indagine di ricerca è infatti emersa la necessità di personalizzare gli spazi, dal momento che al JRC esistono diversi profili lavorativi che richiedono viste *ad hoc*.

In particolare si è pensato di poter *allargare* e *massimizzare* lo spazio a disposizione delle aree che racchiudono l'informazione. L'allargamento prevedeva un'ampliamento di un'area senza nascondere il contenuto delle aree affiancate. La massimizzazione invece consisteva nel nascondere le aree vicine a favore di quella selezionata. Allargamento e massimizzazione sono attivabili contemporaneamente per personalizzare gli spazi dell'interfaccia (vedi fig. 4)

L'interfaccia, diventa operativa con le informazioni. L'utente ha la possibilità di navigare i dati selezionando un elemento dalle sorgenti di informazioni o dagli indici. A seguito di ogni selezione avviene il popolamento contestuale dell'interfaccia con informazioni correlate all'elemento selezionato. Questo significa che se seleziono un documento, termini e documenti visualizzati saranno inerenti al documento di partenza e allo stesso modo avverrà quando un termine è selezionato.

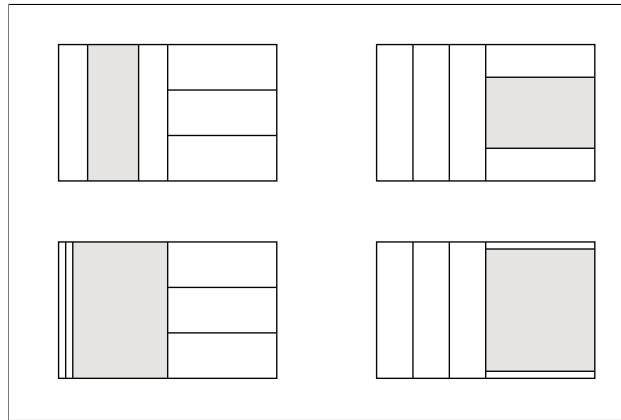


Figura 3: L'utente può personalizzare l'interfaccia. In alto abbiamo due esempi di allargamento di un'area, mentre in basso due esempi di massimizzazione.

Il popolamento si basa sulle relazioni thesaurali. Selezionato un termine, il metodo consiste nel percorrere le relazioni inter-thesaurus e successivamente extra-thesaurus per popolare prima gli indici, successivamente dagli indici verranno popolate le aree dei documenti.

Tutte le relazioni che vengono percorse con l'obiettivo di popolare l'interfaccia sono rese accessibili attraverso *web services*. Un sistema di gestione di thesauri è infatti posto alla base di SIRS per amministrare tutte le informazioni relative ai vocabolari controllati, come termini e relazioni, mentre i documenti sono estratti da sorgenti dati esterne sia al SIRS che al sistema di gestione di thesauri.

Ponto

In questo quadro si colloca un'applicazione in grado di gestire più thesauri contemporaneamente e le rispettive relazioni, siano esse intra-thesaurus o extra-thesauri. All'interno del progetto, dopo un'accurata ricerca dei software esistenti, abbiamo ritenuto opportuno sviluppare un software *ex novo* così da soddisfare tutte le esigenze per realizzare l'interfaccia di SIRS.

Il requisito fondamentale del software era che potesse gestire più thesauri allo stesso tempo e di conseguenza le relazioni all'interno e all'esterno di essi. Infatti la caratteristica del *mapping* era un requisito fondamentale per permettere l'integrazione delle sorgenti di informazione.

Un'altra caratteristica fondamentale era che sul sistema si potessero costruire dei *web service* così da poter estrarre i dati dal database percorrendo le relazioni secondo gradi di profondità. Per esempio ottenere tutti i termini relazionati a un termine ed eseguire successivamente la

richiesta per tutti i termini ottenuti. A causa di queste richieste, ottimizzare le query era fondamentale per ottenere risposte immediate nella fase di information retrieval.

Quest'ultimo motivo è stato fondamentale per orientare la scelta su un sistema *open source*. In particolar modo abbiamo focalizzato l'attenzione sul sistema AGROVOC sviluppato dalla FAO [17]. Da una collaborazione nella quale sono stati coinvolti la FAO di Roma, il British Geological Survey di Nottingham e successivamente, è nato un progetto capace di accumunare le necessità di tutti che prende il nome di *Ponto* [24].

Ponto è il nome di una delle divinità della Grecia. Nell'antichità fu il dio del mare, successivamente sostituito da Poseidone. La parola *pontus* significa mare, all'epoca il collegamento principale tra le terre affacciate sul Mediterraneo. Ai giorni nostri la parola mantiene il concetto ma cambia attori: non più mezzi su mare, ma mezzi su terra che attraversano un ponte. A noi il nome trasmette il concetto di collegamento che esiste tra i thesauri e prende corpo nel *mapping*.

Successivamente alla fase di design, al JRC si è passato all'implementazione del sistema *intra muros*: in un primo momento con la realizzazione della base di dati, poi al popolamento della stessa con i tre thesauri usati nel SIRS (Actions, EUROVOC e Classificazione Decimale Dewey), infine con la creazione di *web service* usati dall'interfaccia per popolare l'interfaccia.

Purtroppo non c'è stato tempo per studiare un'interfaccia adatta al sistema, si è solo realizzato uno sbrigativo pannello per cercare i termini e creare le relazioni. La tesi vuole essere il prolungamento di questo lavoro verso un compimento che forse potrà accadere nei prossimi anni.

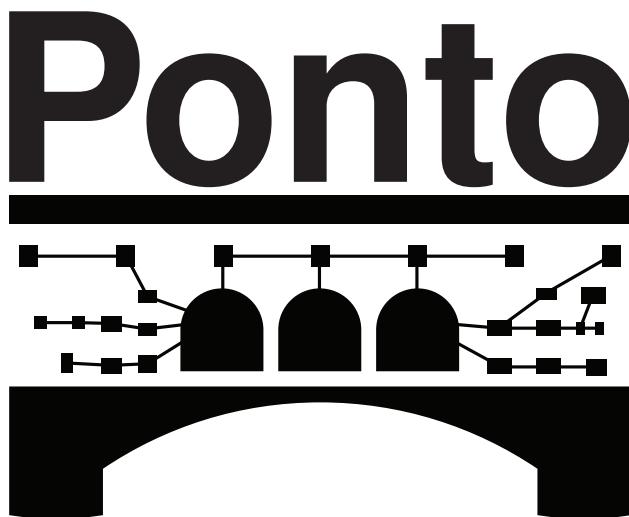


Figura 4: Il logo di Ponto.

Il Metodo

Dopo aver fatto le dovute premesse, sarà possibile addentrarci nella fase di progettazione dell’interfaccia per la gestione di thesauri. Qui verrà introdotto il metodo di lavoro, nei capitoli successivi troveranno spazio tutti gli elementi che andranno a comporre l’interfaccia. Uno alla volta andremo a inserire ogni elemento, giustificandone la presenza e posizionandolo sullo schermo. Con il progredire della lettura aumenteranno la complessità e il dettaglio dell’interfaccia cosicché alla fine si arrivi a costruire un’applicazione mirata alla gestione di thesauri.

Il capitolo corrente si posiziona all’inizio del processo di *design*. Si descriveranno l’analisi del problema, la metafora alla quale ci atterremo nello svolgimento della tesi e il thesaurus scelto per gli esempi che si incontreranno nel manoscritto.

Analisi del problema

Durante la progettazione esistono diversi metodi per analizzare ed approfondire il problema da risolvere, quali l’indagine, l’esplorazione, il *brainstorming*, etc. [32]. Nella sezione andremo a descrivere diversi approcci al lavoro, sebbene ci pare tuttora riduttivo releggere i metodi ad un semplice elenco. Infatti riteniamo che il lavoro di un *designer* sia un’attività che richiede capacità tecnologiche ma anche umanistiche, motivo per cui crediamo che ogni progetto abbia bisogno di un approccio *ad hoc*, ovvero un percorso unico di lavoro sviluppato similmente a un *canovaccio teatrale* nel quale il designer ha il ruolo di direttore. Chi svilupperà l’interfaccia dovrà operare in modo libero e creativo ispirandosi alla metodologia esistente, ma modificandola o stravolgendola a secondo delle necessità. Non esiste una progettazione migliore, piuttosto esistono diverse soluzioni nelle quali il designer è una variabile che va a influire sul risultato.

Il progetto di un’interfaccia per la gestione di più thesauri si basa su uno studio teorico. Approfondire un argomento, rendendo le nozioni parte del nostro bagaglio culturale, aiuta a muoversi con più destrezza

in tutte le fasi della progettazione. Nella lista che segue elencheremo i metodi che sono stati usati per arrivare al disegno dell'interfaccia:

- La necessità
- L'esperienza
- L'esplorazione
- La metafora

La necessità

Durante il progetto SIRS, di cui una descrizione è data a pagina 23, è stata stilata una lista di requisiti per l'acquisto di un software di gestione dei thesauri. Nessun software è stato in grado di soddisfare i punti fondamentali della lista, motivo per cui si è deciso di realizzare un software per la gestione di thesauri *intra muros*. Oltre a questo, la lista di requisiti ha avuto il pregio di mettere in maggiore evidenza l'esigenza di avere un software in grado di gestire più thesauri contemporaneamente, così da permettere le abituali operazioni di mantenimento di un thesaurus sia all'interno di un singolo thesaurus che trasversalmente su più thesauri.

La lista dei requisiti è stato il primo metodo applicato inconsciamente nella progettazione del software, grazie alla quale si è potuto subito impostare una direzione precisa allo sviluppo del progetto.

L'esperienza

In aree di studio così circoscritte l'esperienza sul campo è importan-
tissima. Il modo di ragionare quando si costruisce un thesaurus e le
azioni che portano alla creazione della sua struttura sono capacità pro-
prie dei professionisti del settore. Nel nostro caso è stato utile potersi
confrontare con i curatori del Nuovo soggettario presso la Biblioteca
Nazionale Centrale di Firenze.

Quando si inizia la progettazione di interfacce così specifiche è ne-
cessario lavorare con il supporto di esperti del settore che nel nostro
caso coincidono con gli utenti finali. Il metodo di lavoro non è impor-
tante: si possono raccogliere informazioni durante *brainstroming* oppu-
re attraverso questionari o semplici colloqui. Fondamentale invece sono
i professionisti con le loro idee, coglierle e capire i loro punti di vista
può dare enormi vantaggi al progettista.

L'esplorazione

L'esplorazione è un metodo che lascia al designer molta libertà. Consiste in una personale ricerca dell'argomento compiendo un percorso tra diverse discipline, nel nostro caso interaction design, biblioteconomia, information retrieval, graphic design, impaginazione e tipografia. Tutte le vie di lavoro sono buone per trovare uno spunto, una direzione ma soprattutto per rendere più consistente la propria consapevolezza circa l'argomento.

La metafora

Spesso i designer usano la metafora come misura di paragone. L'idea è di sovrapporre un modello concettuale già esistente e posizionarlo, adattarlo e studiarlo rispetto al progetto da sviluppare.

La metafora che useremo nello svolgimento della tesi si ispirerà al thesaurus in forma cartacea. Attraverso il paragone tra la vecchia forma e la nuova forma dei thesauri cercheremo di riallacciare i confini non più prossimi delle interfacce cartacee e digitali.

Thesaurus digitale e cartaceo

Durante la fase di esplorazione (vedi p. 29) abbiamo constatato che una delle migliori analisi della visualizzazione dei thesauri è stata fatta da Jean Aitchison nel suo libro *Thesaurus Construction and Use* [1]. L'autrice è una famosa esperta di thesauri e il suo libro è un riferimento per i professionisti che lavorano nel mondo della biblioteconomia. All'interno del testo ci ha colpito in particolar modo un frammento che paragona thesauri cartacei e digitali, tanto da influenzare la progettazione dell'interfaccia.

Nel confronto il thesaurus digitale e cartaceo, la Aitchison fa un'analisi dei *supporti* usati per visualizzare le informazioni. Comparando i due *media*, una delle prime cose che si nota è la grande superficie usata dal supporto cartaceo: infatti un libro è composto da molte pagine che possono essere viste come *singole unità di visualizzazione*, ma se dovessimo affiancarle una all'altra le pagine occuperebbero ampio spazio. Le pagine sono stampate su entrambi i lati e la rilegatura del libro permette la visualizzazione a coppie delle pagine, espediente che permette di raddoppiare la superficie di lettura. Inoltre la tecnica della rilegatura permette di ordinare le pagine in una sequenza. La sequenza viene numerata per dare un riferimento del punto di lettura e si rafforza se paragonata allo spessore del libro: infatti numerazione e spessore sono

due misure relazionate fra loro la cui somma è un potente strumento per comunicare la posizione al lettore. Può sembrare una riflessione banale perché l'oggetto libro ci è familiare, ma dal punto di vista della progettazione sono dei solidi riferimenti che permettono all'utente di orientarsi all'interno del thesaurus.

L'ordine, la numerazione, lo spessore, l'accoppiamento delle pagine e la posizione rispetto alla rilegatura sono singoli riferimenti che considerati assieme creano il *contesto*. Sono componenti che formano un ambiente complesso che offre all'utente una percezione dei contenuti del libro.

Il supporto video offre vantaggi diversi. A differenza del cartaceo, il cui spazio è dato dalla somma di molte pagine, il supporto digitale ha una superficie unica. Tuttavia questo divario è compensato dalla *dinamicità* del mezzo digitale, che permette allo schermo di cambiare il contenuto nel tempo. Grazie a questa caratteristica lo schermo è capace di visualizzare ingenti quantità di dati che su cartaceo occuperebbero diversi scaffali. Si pensi per esempio a come un'intera enciclopedia possa essere digitalizzata in un solo DVD.

La visualizzazione digitale introduce nuovi comportamenti nella consultazione, questo perché la fruizione di contenuti digitali è diversa dalla consultazione di un libro. In un thesaurus la dinamicità riveste un meccanismo fondamentale perché permette un passaggio tra diversi tipi di visualizzazioni, per esempio tra visualizzazione alfabetica e sistematica (vedi p. 5). Grazie al media digitale, l'utente sarà in grado di vedere un termine secondo un ordinamento e in un secondo momento vedere lo stesso termine in un nuovo ordine, passando da una visualizzazione all'altra con estrema facilità (vedi p. 78). Talvolta anche il supporto cartaceo offre volumi con diversi ordinamenti, tuttavia spostarsi da uno all'altro è tutt'altro che veloce.

Dopo questo preambolo va detto però che il thesaurus digitale non è perfetto. Il più grave problema riguarda la diffusa mancanza di un contesto che permetta all'utente di orientarsi. Abbiamo già visto come esistano molti sistemi digitali per la gestione dei thesauri (vedi pagina 21), tuttavia nessuno di essi ha sviluppato adeguatamente il contesto di lavoro. Eppure orientarsi nell'informazione è fondamentale. Gibson nel suo libro – *The ecological approach to visual perception* – spiega come l'essere umano sia in grado di dislocare gran parte della sua conoscenza all'esterno di sé: l'ambiente nella sua complessità offre una serie di *affordance*, che si manifestano in persone, luoghi o oggetti e che innescano meccanismi in grado di aiutare l'essere umano a capire quali azioni possono essere compiute in un determinato contesto [22]. Estendendo questo approccio al mondo digitale, cercheremo di creare un'interfac-

cia per thesauri in grado di fornire sufficienti *affordance* in modo che la memoria dell’utente sia dislocata negli elementi che caratterizzano l’interfaccia. Come realizzarlo? La nostra risposta sta nel fornire punti di riferimento, ricreando un contesto d’informazione e utilizzo che si ispiri all’ambiente cartaceo, offrendo all’utente una *mappa* visiva del thesaurus che espliciti l’*affordance* dell’interfaccia.

Nuovo soggettario

Per gli esempi che seguiranno, abbiamo pensato che fosse utile usare una terminologia reale, non fittizia. A tal proposito è stato scelto il *Nuovo soggettario*, l’unico sistema di indicizzazione multidisciplinare in lingua italiana in uso tutt’oggi [7].

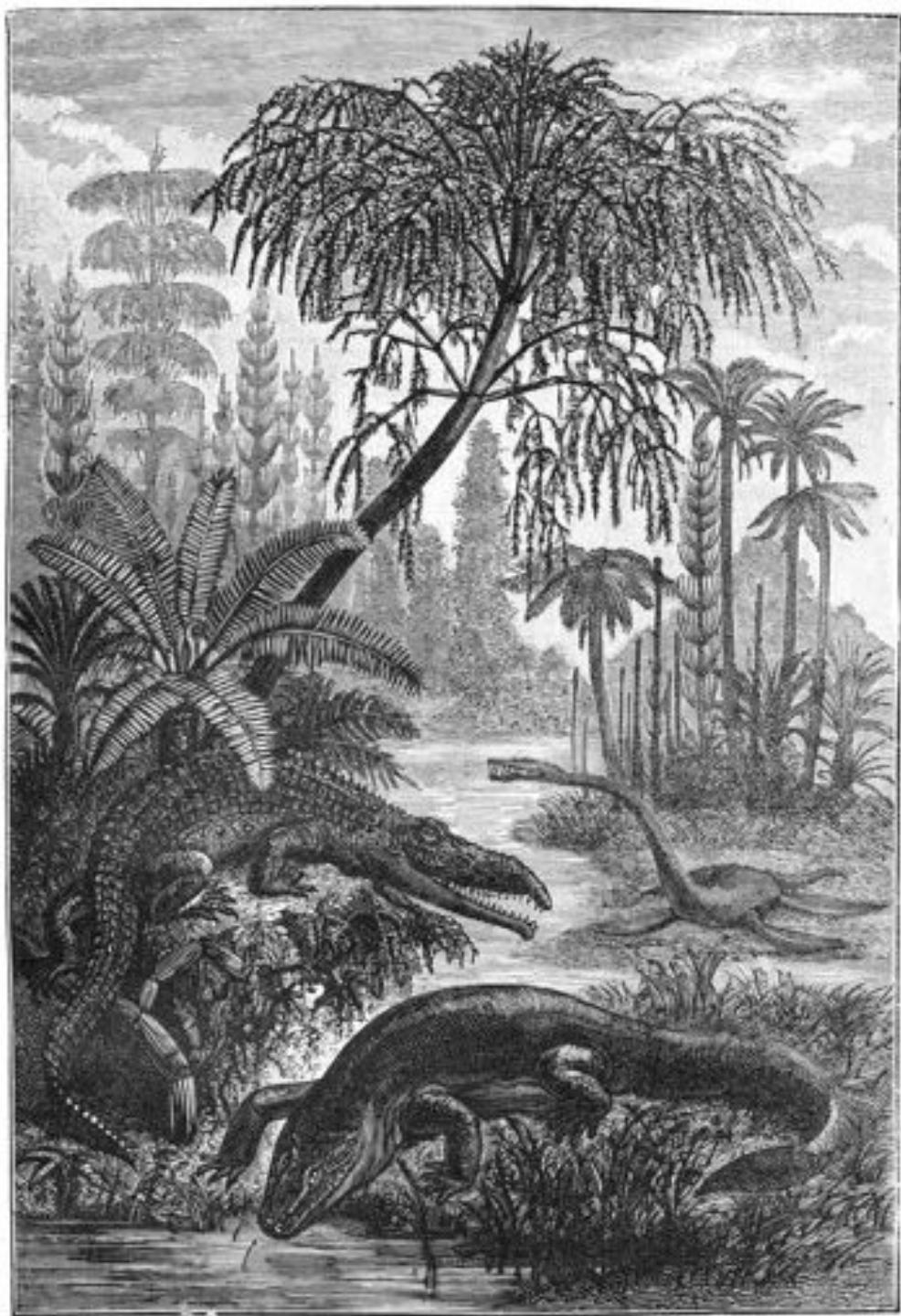
A discapito del nome, omaggio al precedente sistema di indicizzazione, il Nuovo soggettario non è un soggettario, ma un sistema di indicizzazione costituito da diversi elementi tra cui un thesaurus. Sviluppato interamente dalla Biblioteca Nazionale Centrale di Firenze, l’architettura di questo sistema comprende quattro elementi:

- Il manuale applicativo
- Le regole per semantica e sintassi
- Il thesaurus
- L’archivio di stringhe di soggetto

In questa struttura, thesaurus e archivio di stringhe di soggetto sono i contenitori dell’informazione. Entrambi sono molto importanti: il thesaurus mette a disposizione i termini con i quali si creano le stringhe di soggetto e l’archivio di stringhe è fondamentale in fase di reperimento dell’informazione [40].

Il thesaurus del Nuovo soggettario sarà utilizzato per estrarre termini e relazioni utili alla visualizzazione degli esempi che incontreremo nello svolgimento della tesi.

A sottolineare la volontà alla diffusione del Nuovo soggettario da parte della Biblioteca Nazionale Centrale di Firenze, va detto che il thesaurus del sistema è a completa disposizione di bibliotecari e studiosi, essendo essi liberamente disponibili online [6]. La possibilità di consultare il Nuovo Soggettario online ha agevolato la realizzazione degli esempi, che altrimenti sarebbe stata lunga e laboriosa.



I Costituenti

Termine

Il testo è composto da vocaboli che sono gli elementi basilari con i quali si compongono frasi, discorsi, descrizioni, dizionari e naturalmente anche i thesauri. Essendo noi interessati alla visualizzazione cercheremo di capire come meglio rappresentare i *termini* perché essi siano leggibili, accessibili e utilizzabili in un contesto di interfaccia uomo-macchina. Per raggiungere l'obiettivo approfondiremo l'argomento attraverso un occhio tipografico, andando a collocare i costituenti in uno spazio visivo che nel nostro caso corrisponde allo schermo.

Data la particolare interdisciplinarità dell'argomento, capiterà nel corso del capitolo – com'è d'altronde già capitato nel capitolo precedente – di cambiare linguaggio a seconda della disciplina più vicina all'argomento trattato. Anche per i vocaboli impiegati sarà così. In precedenza ci siamo riferiti alle parole che compongono un vocabolario controllato come a *termini*, impiegati in biblioteconomia per riferirsi alle unità logiche di un thesaurus. Se invece stiamo utilizzando un soggettario – come nel caso del soggettario della Library of Congress – si è soliti usare la parola *soggetto*. Se dobbiamo rifarcirci alla letteratura in generale i soggettari e le classificazioni sono composte da elementi che si avvicinano maggiormente a una *descrizione* – caratterizzata da una maggiore estensione – mentre nel thesaurus si usa la parola *termine* per comunicare un maggiore senso di unità linguistica [11].

Quando invece oltrepassiamo il confine del *medium* informatico, facciamo uso di *keyword*. Per esempio nei motori di ricerca inseriamo parole chiave, utilizziamo chiavi di accesso per accedere ai sistemi, assegnamo delle keyword per identificare oggetti e persone nei *social network*. In information retrieval invece si fa uso di *termini* come in biblioteconomia. Diversamente il mondo delle ontologie è caratterizzato da *concetti* (vedi p. 19).

Tipografia

La *tipografia* è definita dall’Oxford Dictionary come “*the art or process of setting and arranging types and printing from them*” [37]. Può essere vista come un processo di composizione e armonizzazione che ha come unico scopo quello di dare dignità al testo [8]. Non ha importanza quale percorso si intraprenda per raggiungere l’obiettivo – sia esso condotto dalla creatività, dalla logica o dall’esperienza – importante è raggiungere un prodotto armonico, leggibile e funzionale allo scopo che sia adeguato alla presentazione del testo.

Con le nuove tecnologie il contenuto non è riconducibile al solo testo, ma ad accompagnarlo sono anche suoni, video e la navigazione attraverso *hyperlink* che ha caratterizzato l’epoca dei *browser* [41]. In questo contesto la tipografia si adatta, forte di un’esperienza plurisecolare.

L’adattamento si attua in quattro sviluppi paralleli originati dal cambio di supporto. In primo luogo l’adozione degli schermi ha portato a un adattamento a formati diversi dalla carta stampata. Tutto ciò che negli anni la tipografia ha imparato dall’impaginazione di un libro è stato reinventato dando nuova linfa a una disciplina che sembrava circoscritta alla carta stampata. Secondariamente la densità d’immagine degli schermi, più bassa rispetto alla carta, ha portato a ridisegnare i caratteri sulla base della limitazione dei pixel portando a un generale abbassamento della qualità tipografica. Il terzo adattamento avviene con l’avvento del *desktop publishing*. Con l’avvento dei computer inizialmente l’impaginazione viene realizzata con linguaggi digitali come Groff, MacWrite, PostScript e Latex, successivamente passa alla modalità grafica WYSIWYG (What You See Is What You Get) e quindi i Word processor, Photoshop e Illustrator, Quark X Press per giungere quindi all’ambito web con linguaggi come l’HTML (HyperText Markup Language) e i CSS (Cascading Style Sheets). Questi strumenti hanno reso accessibile l’ambito della tipografia e dell’impaginazione a un pubblico molto più ampio. Il quarto punto è legato al fatto che la tipografia è cresciuta sulla staticità, ciò che è stampato conserva il contenuto immutato nel tempo. L’avvento dei supporti digitali porta la possibilità di creare una tipografia *dinamica*, mutevole soprattutto nel dialogo con l’utente, ovvero nell’*interazione*.

I punti elencati illustrano dei cambiamenti nell’uso dello strumento tipografico e saranno fondamentali nella progettazione della nostra interfaccia per la gestione dei thesauri.

Supporti ad alta definizione

La tipografia è stata capace di adattarsi nel tempo a un supporto a bassa densità di immagine come lo *schermo*. Le misure che definiscono la densità d'immagine della carta stampata e dello schermo derivano dalla stessa radice. Si parla infatti di DPI (Dots Per Inch) come il numero di punti d'inchiostro stampati orizzontalmente o verticalmente su un supporto cartaceo in uno spazio di 2,54 cm (1 in). Si parla invece di PPI (Pixels Per Inch) quando il supporto è lo schermo.

Finché si rimane nell'ambito cartaceo non vi sono ambiguità nell'uso del dpi, ma non si può dire lo stesso per quanto riguarda il ppi. Quest'ultima è una misura introdotta negli anni '80, con l'avvento dei primi sistemi grafici. Così come oggi, a contendersi il mercato erano le multinazionali statunitensi Apple e Microsoft. Quando la tipografia iniziò a spostarsi dal cartaceo al digitale, le due aziende non si affidarono a standard, ma attraverso un calcolo presupposto della distanza dell'utente dal monitor e dalla risoluzione dello schermo arrivarono a due conclusioni differenti. Il sistema Windows si sarebbe basato su un'interfaccia grafica a 96 ppi, mentre il Mac OS avrebbe avuto una risoluzione di 72 ppi. Questi valori validi per i primi computer, si sono rivelati totalmente inadatti sia nelle generazioni successive e ora più che mai nei giorni nostri.

Perché i ppi fissati da Windows ed Apple non sono affidabili? Prima di tutto perché la stima è stata fatta in una situazione dove una persona alla scrivania usa un computer di tipo *desktop*. Ora i contesti sono estremamente vari: una persona può ritrovarsi a lavorare su un computer portatile, usare uno *smartphone*, vedere delle immagini proiettate su un muro. Gli schermi sono differenti per dimensione: da pochi pollici di un telefono cellulare ai 30 pollici di un moderno schermo da scrivania. Infine gli schermi variano in risoluzione: una persona che preferisce vedere immagini più grandi, abbassa la risoluzione del monitor perché gli attuali sistemi operativi non contemplano una gestione dei ppi. Il problema è che la misurazione attuale non definisce standard per questi utilizzi.

Per esempio lo schermo di un computer da scrivania come il iMac da 21.5" ha una risoluzione nativa di 1920x1080 pixel e quindi un ppi di 102,46. Un portatile come il Macbook Pro con uno schermo da 15,44" e una risoluzione nativa di 1440x900 pixel ha un ppi di 110,27. La versione 17" dello stesso computer ha uno schermo con una densità d'immagine ancora più alta: a fronte di una risoluzione di 1920x1200 pixel ha un ppi di 133,19. Inoltre il divario cresce se ci si sposta ai nuovi supporti portatili.

Il nuovo iPad ha un'alta risoluzione se paragonato a un normale computer, infatti il suo ppi è di 131.96. Invece la nuova generazione di *smartphone* ha portato a un notevole salto di qualità in termini di qualità visiva per l'utente. Infatti la quarta generazione di iPhone ha uno schermo di 3.5" e una risoluzione di 960x640 pixel. I dati portano a un risultato di 329.65 ppi, pari a più del doppio dell'appena citato iPad.

In fotografia una stampa di alta qualità ha una risoluzione di circa 300 ppi, limite dato dal supporto cartaceo. Con la nuova generazione di schermi ad alta risoluzione si ha per la prima volta un balzo nella qualità di visualizzazione del supporto. Questo potrebbe rappresentare un fattore scatenante che porterà all'abbandono del supporto cartaceo.

Scelta del carattere

Se già con l'arrivo dei monitor di grandi dimensioni come i 30" attuali si è potuto progettare applicazioni e interfacce grafiche specializzate per cabine di regia televisive o centri di controllo, in egual modo l'alta risoluzione porterà a nuove soluzioni nel campo delle interfacce grazie a un drastico miglioramento nella qualità della visualizzazione. L'effetto di *antialiasing* in combinazione con l'alta risoluzione farà sparire completamente percezione di lavorare davanti a una matrice di pixel e alla tipografia sarà data una sua dignità, riproducendo dettagliatamente i profili di ogni carattere anche a dimensioni più piccole.

Nella fase di progettazione ci avvarremo della possibilità di scegliere il carattere più appropriato partendo dal presupposto che l'interfaccia verrà sviluppata per uno schermo ad alta definizione, dimenticandoci così dei caratteri disegnati per aumentare la leggibilità su schermo. A tal fine abbiamo praticato un'ulteriore selezione a una lista di caratteri già stilata da Robert Bringhurst [8], ottenendo così un insieme di 24 caratteri, rispettivamente ordinati in una prima serie di 16 caratteri con grazie seguita da una seconda composta da altri 8 caratteri senza grazie (vedi pp. 39-42).

Le grazie sono una forma grafica usata in tipografia per abbellire il carattere e guidare la lettura aumentano la leggibilità del testo. Rappresentano un fattore di categorizzazione principale che ordina i caratteri in due gruppi: con grazie e senza grazie, rispettivamente conosciuti come *serif* e *sans serif* (vedi fig. 5).

I caratteri senza grazie sono più antichi, le prime scritture greche risalenti a 4000 anni fa lo dimostrano. Tuttavia gli ultimi due millenni – epoca nella quale la scrittura si è diffusa con voracità – i caratteri con grazie sono stati i più usati e studiati in Europa. Nel corso del ventesimo secolo si è poi riscoperta la semplicità del carattere senza grazie per

arrivare ai giorni nostri e a Internet dove il carattere *bastoni* è sempre stato ampiamente utilizzato. Probabilmente a causa di queste ragioni storiche, il libro ha sempre avuto un'affinità particolare con il carattere graziato. Dato il nostro presupposto di ispirarci alle edizioni cartacee del thesaurus e trovando nell'inversione di tendenza tra bastoni e grazie un distacco dal vecchio mondo delle interfacce nate nei browser, la scelta del carattere ha in primo luogo escluso i caratteri *sans-serif* andando a preferire le grazie, alle quali viene dato anche il pregio di aumentare la leggibilità del testo.



Figura 5: Legacy è un carattere disponibile in versione senza e con grazie, rispettivamente a sinistra con Legacy Sans e a destra con Legacy Serif.

Il gruppo dei caratteri *serif* verrà ulteriormente scremato escludendo tutti quei caratteri che mostrano maggiore differenza di corpo tra le lettere maiuscole e minuscole. Secondo il nostro parere, avere maiuscole e minuscole più omogenee renderà più bilanciata la visualizzazione dei termini, specialmente nella forma di lista. Questo ci porta a escludere i caratteri Baskerville, Bembo, Caslo, Electra riducendo la lista a 12 elementi.

Un altro fattore importante sarà quello dello spazio. Evitare già a questo punto di utilizzare i caratteri più larghi, aiuterà a risparmiare spazio prezioso che potrà essere utilizzato per valorizzare il contenuto dell'interfaccia e visualizzare un maggior numero di termini. In questo passaggio andremo a eliminare i caratteri Berling, Galliard, Kis, Minion, Sabon e Simoncini Garamond. La lista si riduce così a 6 elementi: Albertina, Chaparral, Dante, Photinia, Spectrum e Van Dijck.

Dopo un'analisi più approfondita la scelta è ricaduta sul carattere *Photinia* (vedi p. 40) disegnato dal francese José Mendoza y Almeida [33]. Le ragioni che hanno portato a questa scelta sono diverse, tuttavia va sottolineato che i caratteri dell'ultimo gruppo avevano tutte caratteristiche conformi al nostro fine. Tuttavia *Photinia* si è rivelato par-

ticolarmente adatto per essere un carattere equilibrato nello spessore delle forme, qualità che renderà gradevole il riempimento degli spazi. Inoltre, oltre a essere riconosciuto come uno dei più riusciti caratteri post-moderni, ha anche una versione *sans-serif* che prende il nome di *Univers*. Quest'ultima sarà una qualità che sicuramente si rivelerà utile nella costruzione semantica dell'interfaccia, utilizzando il Photinia per rappresentare i termine del thesaurus e l'*Univers* per i pannelli di supporto.

Caratteri monospace

Nella scelta del carattere sono stati esclusi i caratteri *monospace*, caratterizzati dal fatto che tutte le lettere che li compongono sono eguali in larghezza. Lo scopo di questo carattere è quello di creare un grande ordine nella disposizione delle parole, creando una sorta di reticolo visivo. Infatti avendo le lettere lo stesso ingombro orizzontale, ogni lettera sarà allineata a quella posizionata sopra e quella posizionata sotto.

Alcuni thesauri utilizzano questo tipo di carattere per dare più ordine alla lista. In particolare il carattere dà ottimi risultati con le liste permutate, dove l'allineamento contribuisce a migliorare la lettura. Tuttavia i caratteri monospace non sono molto diffusi perché allo stesso tempo creano uno scompenso nelle spaziature tra le lettere.

Noi non abbiamo ritenuto abbastanza forte la necessità di un allineamento verticale a sfavore della spaziatura orizzontale, motivo per cui i caratteri monospace non sono stati presi in considerazione nella progettazione dell'interfaccia.

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNPQRSTUVWXYZ
1234567890
```

Figura 6: Albertina

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNPQRSTUVWXYZ
1234567890
```

Figura 7: Baskerville

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNPQRSTUVWXYZ
1234567890
```

Figura 8: Bembo

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNPQRSTUVWXYZ
1234567890
```

Figura 9: Berling

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNPQRSTUVWXYZ
1234567890
```

Figura 10: Caslon

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNPQRSTUVWXYZ
1234567890
```

Figura 11: Chaparral

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890

Figura 12: Dante

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890

Figura 13: Electra

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890

Figura 14: Galliard

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890

Figura 15: Kis

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890

Figura 16: Minion

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890

Figura 17: Photinia

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 18: Sabon

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 19: Simoncini Garamond

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 20: Spectrum

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 21: Van Dijck

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 22: Frutiger

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 23: Futura

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 24: Gill Sans

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 25: Helvetica

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 26: Flora

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 27: Legacy Sans

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 28: Optima

```
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
1234567890
```

Figura 29: Syntax

Lista

Una volta scelto il carattere, possiamo utilizzare il termine come elemento per costituire forme più complesse. I thesauri cartacei sono soliti rappresentare i termini in *liste*, i cui termini sono ordinati in una successione verticale.

La lista di tipo alfabetico è la più diffusa, tuttavia esistono altri ordinamenti che forniscono nuove rappresentazioni dei termini di un thesaurus. L'ordine alfabetico sarà usato nelle prime prove di impaginazione, successivamente verranno introdotti anche gli altri ordinamenti (vedi p. 76).

Nella fase di progettazione useremo un thesaurus italiano, il Nuovo soggettario [6]. Ci fornirà i termini per formare una lista utile alle prime prove di impaginazione (fig. 30).



Jainismo
Jazz
Jazz nel cinematografo
Jazz nella letteratura italiana
Jazz per pianoforte
Jazzisti
Job burnout
Join-up
Joint ventures
Joint-ventures
Ju Jitsu
Ju Jutsu
Judo
Jugendstil
Jujitsu
Jujutsu
Juke-box
Juke-boxes
Jukebox
Jukeboxes
Jutland, battaglia dello. 1916

Figura 30: L'ordinamento alfabetico della lettera J del Nuovo soggettario. La dimensione del testo è di 10 pt, l'interlinea è di 12 pt.

Progettare per schermi ad alta densità di pixel significa lavorare su un supporto qualitativo pari alla carta di alta qualità che si aggira attorno ai 300 ppi. Non essendo ancora in commercio schermi ad alta risoluzione di grandi dimensioni, durante la progettazione è stato utile fare prove di stampa su carta ad alta qualità in modo da simulare un futuro impiego su computer. L'interfaccia sarà comunque adatta agli schermi tutt'oggi in commercio senza pregiudicarne il funzionamento, sarebbe sbagliato non prevederlo. Tuttavia ci piace immaginare un futuro prossimo nel quale l'interfaccia guadagnerà in termini di leggibilità fino a raggiungere o addirittura superare la qualità delle nostre bozze cartacee.

Adattare il testo alla visualizzazione

La tipografia è fatta di misure modificabili per migliorare l'aspetto del testo, ne faremo uso per rendere la lista leggibile e compatibile rispetto ai vincoli dati dal progetto dell'interfaccia. Nelle prossime pagine vedremo come siamo andati ad operare sul testo in base a dimensioni, spaziatura, interlinea e creantura.

Dimensioni

In tipografia l'unità di misura principale è il *punto tipografico* che viene impiegato sia su carta che su schermo. Questa misura viene indicata con la sigla pt e corrisponde a $\frac{1}{72}$ in (circa 0,3527 mm). In tipografia, una volta fissata la dimensione del carattere, essa sarà l'unità di riferimento per misurare la distanza tra le parole o l'altezza delle righe. Questo metodo rende le distanze generali funzionali a un cambio di grandezza del carattere. Useremo questa caratteristica per cambiare la grandezza del carattere per due ragioni. La prima è legata alla vista di ciascuna persona, dobbiamo permettere all'utente di personalizzare la grandezza del carattere in base alle proprie capacità visive. Il secondo motivo è legato invece al contesto, si dovrà permettere all'utente di ingrandire il carattere in modo da visualizzare poche parole o al contrario di rimpicciolirlo fino a poter vedere più termini possibili [53]. Se da una parte il limite di grandezza sarà la visualizzazione di un singolo termine, dall'altra sarà l'esigenza di vedere un termine selezionato nel suo contesto, ovvero cercare di visualizzare più termini possibili spingendosi fino a perdere la leggibilità dei termini per dare una visione globale di quello che è l'oggetto thesaurus. Nel design dell'interfaccia abbiamo deciso di utilizzare la scala classica impiegata in tipografia, che comprende caratteri tra i 6 pt e i 72 pt disposti in scala diatonica (vedi p. 45).

Dato che riteniamo questa interfaccia un'applicazione specifica di un ambito preciso, vogliamo permettere al professionista che la usa quotidianamente di essere rapido nelle azioni. Per raggiungere questo obiettivo uno dei metodi impiegati è basato sull'uso di *combinazioni di tasti* che permettono di accedere a una funzionalità più velocemente del mouse. Per modificare la dimensione del carattere adotteremo le combinazioni ctrl con + per aumentare il carattere alla grandezza successivo della scala, ctrl e - per diminuire.



Figura 31: Le grandezze di carattere utilizzate nell'interfaccia corrispondono alla serie tradizionale utilizzata in tipografia sin dal 16° secolo. Nella figura si vede il carattere Photinia nella successione di dimensioni 9 10 11 12 14 16 18 21 24 36 48 pt.

La spaziatura e l'interlinea

Facendo ampio uso delle liste nell'interfaccia, vogliamo che essa abbia un'alta leggibilità. I parametri che possono migliorare la visualizzazione sono la *spaziatura* e l'*interlinea*. La prima corrisponde allo spazio che

intercorre tra le lettere che compongono una parola. L'unità di misura è relativa alla grandezza del carattere in uso e viene indicata con il simbolo *em* (l'unità di misura non corrisponde a un acronimo, si legge come si scrive). Per esempio utilizzando un carattere di 16 pt, 1 em corrisponderà a 16 pt e 1,5 em corrisponderanno a 24 pt.

L'interlinea invece è la distanza che intercorre tra una linea e l'altra. Definito il *corpo* come l'ingombro massimo in altezza della riga di testo, l'interlinea misura la distanza dalla base del corpo di una linea, alla base del corpo della linea successiva. Anche per l'interlinea si preferisce usare una misura relativa, cosicché cambiando la dimensione del carattere anche la distanza tra le righe si accorderà in proporzione. Per esempio in un testo dove l'interlinea misura 1,2 em, un carattere di dimensione 10 pt avrà un'interlinea di 12 pt. Spaziatura e interlinea sono misure che contribuiscono a creare l'impaginazione di un testo. Esse dipendono una dall'altra, infatti il bilanciamento di una pagina di testo è dato anche dalla corretta combinazione delle due misure.

Ora cercheremo di applicare spaziatura e interlinea in modo da migliorare il più possibile la lettura dei termini. Per capire quale poteva essere la combinazione migliore abbiamo creato un poster che possa rappresentare le variazioni di spaziatura e interlinea (fig. 33). Il poster ci servirà per valutare quale impaginato gode della maggiore leggibilità. In particolare cercheremo di trovare il giusto bilanciamento tra lista e termine. Infatti, mentre in un libro il contenuto deve essere leggibile ed omogeneo, nella nostra interfaccia c'è la necessità di isolare maggiormente i termini, renderli parte di un tutto, motivo per cui preferiremo aumentare l'interlinea. Questa operazione ci permetterà di mantenere coesi i singoli termini e distanziarli tra loro, senza dimenticare però che aumentando l'interlinea una colonna visualizzerà meno termini.

Confrontando le varie soluzioni, si è ritenuto appropriato scegliere un'interlinea di 1,3 em e una spaziatura poco maggiore di quella standard. Nel poster la combinazione scelta corrisponde alla colonna 5, riga 5 che mostriamo in dettaglio nella figura 32.

Crenatura

I sistemi di *desktop publishing* possono applicare la *crenatura* – in inglese *kerning* – che permette di ottimizzare la spaziatura tra le lettere. Per esempio può aggiungere spazio nel caso due lettere soffochino lo spazio tra di loro, come nel caso HI, oppure togliere spazio nel caso ve ne sia in eccesso come per le lettere Tu (notate in questo caso come la u finisce sotto la maiuscola T).

La crenatura migliora sensibilmente la leggibilità del testo. Mentre la sua presenza non si nota, la sua assenza pesa molto nella presentazione e nella lettura di un testo, soprattutto su schermi ad alta risoluzione. Per questo motivo, nella fase d'implementazione dell'interfaccia, si consiglia di usare una tecnologia che supporti la crenatura come *Adobe Flash* o *Processing* [21] [47]. Al momento i CSS (Cascading Style Sheet) non impiegano ancora la crenatura e quindi al momento l'HTML puro ne è sprovvisto.

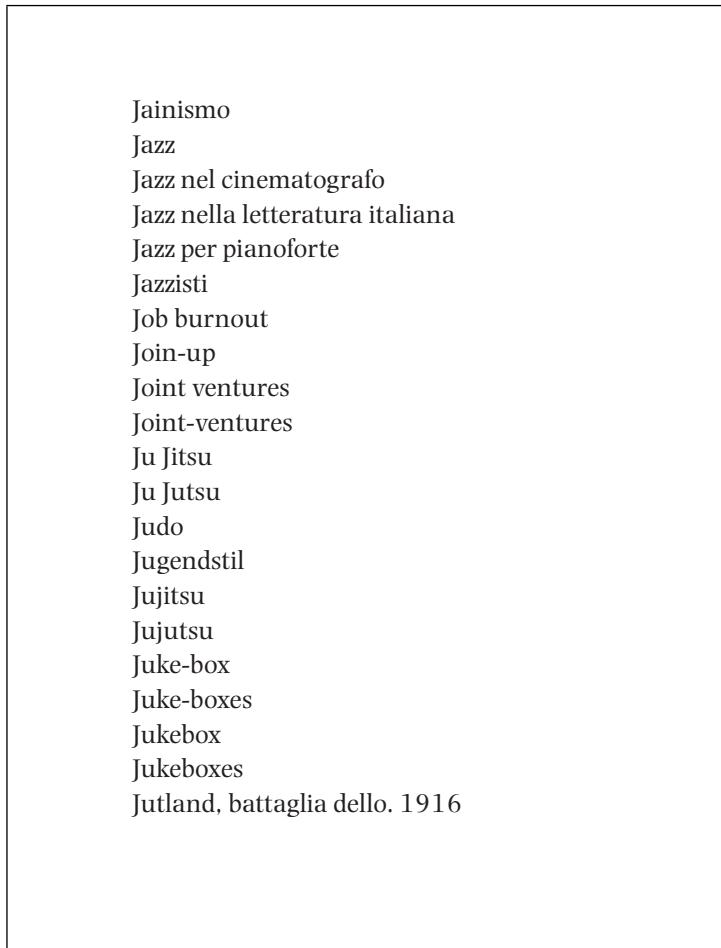


Figura 32: La figura rappresenta la proporzione tra interlinea e spaziatura scelta per la lista. Nell'esempio abbiamo un carattere di 10 pt con un'interlinea di 1,3 em e una spaziatura corrispondente al valore 5 applicato da Adobe Illustrator.



Figura 33: In scala il foglio formato A0 utilizzato per valutare il miglior rapporto tra interlinea e spaziatura per la lista. Il carattere utilizzato ha dimensione 10 pt. L'interlinea varia tra 9 pt e 19 pt (da sinistra a destra) mentre la spaziatura tra +25 e -25 secondo i canoni di applicazione di Adobe Illustrator.

Lo Spazio

Questo capitolo riguarda lo spazio che ospiterà l'interfaccia e gli elementi che lo andranno a popolare. Studieremo lo schermo e le sue dimensioni, inseriremo la lista al suo interno e vedremo come i termini verranno visualizzati al variare del carattere. Infine introdurremo il paradigma di navigazione ispirato ai thesauri cartacei.

Schermo

Per iniziare a progettare la rappresentazione del thesaurus, inizieremo dal definire i *vincoli* dello schermo. Il primo riguarda la visualizzata in uno schermo ad alta risoluzione (vedi p. 35). Nel mondo degli *smartphone* sono appena arrivati i primi schermi a risoluzioni superiori ai 300 ppi, immaginiamo che nel giro di qualche anno l'alta risoluzione raggiunga anche il panorama degli schermi per computer come sta già succedendo per i *netbook*. Con l'introduzione di questa tecnologia la leggibilità dei caratteri sarà migliore e lo sviluppo di interfacce specifiche come la nostra sarà maggiore. Il secondo vincolo è legato alle dimensioni dello schermo. Infatti riteniamo che un'applicazione, per essere usata su tutti i tipi di schermi, debba essere *scalabile* a seconda delle dimensioni. Questo significa che l'applicazione potrà essere usata indipendentemente dello schermo a disposizione dell'utente. A tal fine abbiamo deciso di studiare la compatibilità dell'interfaccia su tre schermi con diverse diagonali (vedi figg. 34, 35, 36).

Dato che riteniamo che le applicazioni professionali – come nel nostro caso di un software per la gestione dei thesauri – per raggiungere la massima resa debbano essere operative su grandi schermi, adotteremo come riferimento la misura diagonale di 27" (vedi fig. 36). Tuttavia l'intenzione è quella di disegnare un software capace di adattarsi a qualunque misura di schermo, motivo per cui terremo in considerazione anche le altre due misure minori così da assicurare una discreta usabilità dell'applicazione anche nel caso di *laptop*.



Figura 34: Schermo 15,4" in $\frac{16}{10}$ pari a 13,06 in x 8,16 in



Figura 35: Schermo 21.5" in $\frac{16}{9}$ pari a 18,74 in x 10,54 in



Figura 36: Schermo 27" in $\frac{16}{9}$ pari a 23,53 in x 13,28 in

Impaginare la lista

Il passo successivo è inserire la lista nello spazio. La superficie di lavoro è delimitata da un margine trasparente di 0,5 in per le parti superiore e inferiore, mentre il margine laterale è dato dall'intersezione dei margini superiore e inferiore con le diagonali dello schermo. Applicando il procedimento si calcolano i margini per ogni tipo di schermo.

Nelle figure 37, 38 e 39 ci sono i primi esempi di impaginazione con il carattere impostato a 12 pt. Si è scelto di disporre i termini in colonne in modo da ricostruire una sequenza familiare all'utente abituato all'uso del thesaurus cartaceo. Inoltre la scelta delle colonne permette di ottimizzare gli spazi tra una colonna e l'altra, in modo da visualizzare il numero maggiore di termini senza rinunciare alla leggibilità.

È interessante contare quanti termini è possibile vedere contemporaneamente sullo schermo. Nel 15,4" sono visibili 132 termini, nel 21,5" 264 e nel 27" 399. Visualizzare molti termini insieme non ha scopi estetici, il motivo principale è funzionale: aumenta la percezione del

contesto in cui i termini sono inseriti. Il contesto va ad arricchire la conoscenza dell’utente, permettendogli di fare analisi quantitative sui termini presenti nel thesaurus, di cercare un termine senza cambiare pagina e di acquisire una maggiore consapevolezza del thesaurus nella sua interezza. Considerando inoltre che la differenza numerica tra i 132 termini dello schermo 15,4” e i 399 termini del 27” non è minima, non c’è dubbio che l’uso del 27” agevoli gli utenti nel lavoro.

Per tal motivo nel seguito della progettazione ci concentreremo sulla visualizzazione del thesaurus su schermi a 27”. Tuttavia vogliamo precisare che il sistema sarà necessariamente compatibile con gli altri schermi illustrati, nonché con gli schermi con una diagonale compresa tra i 15,4” e i 27”.



Figura 37: Lo schermo 15,4” visualizza 132 termini di 12 pt, disposti su 33 linee e 4 colonne. La larghezza della colonna è pari a 13 em.

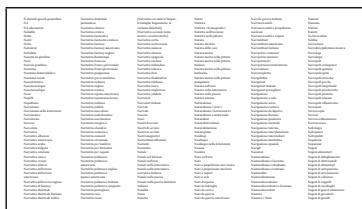


Figura 38: Lo schermo 21,5” visualizza 264 termini di 12 pt, disposti su 44 linee e 6 colonne. La larghezza della colonna è pari a 13 em.



Figura 39: Lo schermo 27” visualizza 399 termini di 12 pt, disposti su 57 linee e 7 colonne. La larghezza della colonna è pari a 13 em.

Dimensione del carattere

Una delle prime funzionalità che vogliamo applicare alla nostra visualizzazione riguarda la modifica delle dimensioni del carattere. Come già illustrato a pagina 44, è fondamentale lasciare all'utente la possibilità di adattare lo schermo alle proprie esigenze. Per esempio un utente giovane con una vista perfetta navigherà i termini anche a una dimensione di 9/10 pt, mentre una persona più avanti negli anni sceglierà una dimensione maggiore. Va considerato inoltre che questa funzionalità può essere d'aiuto alle persone con deficit visivi, perché il carattere raggiunge grandezze nei limiti delle dimensioni dello schermo. Caratteri di 6 pt potranno fornire invece all'utente una vista generale sui termini del thesaurus, rinunciando consapevolmente alla leggibilità.

Nelle figure 40, 41, 42 e 43 abbiamo applicato allo schermo 27" quattro tipi diversi di dimensione carattere. Con l'impaginazione adottata a 6 pt possiamo arrivare a mostrare 1582 termini, opzione che può essere usata ad esempio per capire quanti termini iniziano con la lettera N. Al contrario a 18 pt possiamo visualizzare un massimo di 190 termini, concentrando la nostra attenzione e apprezzando appieno le curve del carattere Photinia.

Daltonismo e accessibilità

Come si può evincere dalle figure presenti nella tesi, si è scelto di sviluppare l'interfaccia interamente in bianco e nero. Con questo non si vuole dimostrare un'avversione verso i colori, piuttosto continuare a seguire l'approccio iniziale legato all'uso tipografico. Nel caratterizzare gli elementi che appaiono nell'interfaccia abbiamo sempre scelto la via più semplice, ovvero ci siamo resi conto che spesso basta una sola caratteristica diversa per rendere un elemento distinguibile dagli altri. Abbiamo usato lo stile, il carattere, la dimensione, ma non abbiamo usato il colore.

In un discorso che va a toccare anche l'accessibilità di un'interfaccia, l'assenza del colore non è stata una scelta casuale. In verità abbiamo preferito sviluppare un'interfaccia in bianco e nero per renderla maggiormente accessibile alle persone affette da daltonismo. Spesso quando si progetta un prodotto, non ci si rende conto dei disagi che potrebbero affliggere le persone sul posto di lavoro. Progettare un'interfaccia in bianco e nero permette alle persone con impedimenti nella lettura del colore l'uso dell'applicazione.

Figura 40: A dimensione 6pt lo schermo 27" visualizza un massimo di 1582 termini, disposti su 113 linee e 14 colonne.

Figura 41: A dimensione 9 pt lo schermo 27" visualizza un massimo di 684 termini, disposti su 76 linee e 9 colonne.

Figura 42: A dimensione 12 pt lo schermo 27" visualizza un massimo di 399 termini, disposti su 57 linee e 7 colonne.

N-dimetil-guanil-guanidina	Narrativa per ragazzi	Nastri di acciaio	Naturalismo	Naturalismo <Arte>
NA	Narrativa dell'orrore inglese	Nastri in acciaio	Naturalismo <Letteratura>	Naturalismo <Letteratura>
NA alternativo	Narrativa di fantasy	Nastri magnetici	Naturalismo e misticismo	Naturalismo e misticismo
Nababbi	Narrativa dialettale	Nasturium officinale	Naturalisti	Naturalisti italiani
Nabis	Narrativa dialettale friulana	Nasturzi	Naturalizzazione	Naturalizzazione
Nadir	Narrativa dialettale ladina	Nasturzio	Naturopatia	Naturopatia
Nafta	Narrativa dialettale piemontese	Natale	Naufragi	Naufragi
Natfalene	Narrativa dialettale piemontese	Natale nel folklore	Naufragio	Naufragio nella letteratura
Natalina	Narrativa ebraica	Natale nell'arte	Nausea	Nausea
Nanetti da giardino	Narrativa erotica	Natale nella letteratura	Nautica	Nautica
Nani	Narrativa fantastica	Natale nella letteratura	Nave nell'arte	Nave nell'arte
Nani da giardino	Narrativa fantastica tedesca	narrativa italiana	Navi	Navi
Nanismo	Narrativa fantasy	Natale nella poesia	Navi a propulsione meccanica	Navi a propulsione meccanica
Nanismo dismetabolico	Narrativa fantasy americana	Natale nella poesia dialettale	Navi a propulsione nucleare	Navi a propulsione nucleare
Nanismo renale	Narrativa fantasy inglese	italiana	Navi a vapore	Navi a vapore
Nanostrutture	Narrativa flamingo	Natalità	Navi a vela	Navi a vela
Nanotecnologia	Narrativa francese	Nates	Navi d'argento	Navi da battaglia
Naos	Narrativa franco-provenzale	Natiche	Navi da carico	Navi da carico
Napelli	Narrativa giapponese	Nativi	Navi da guerra	Navi da guerra
Napolitane	Narrativa greca moderna	Natività	Navi da guerra americane	Navi da guerra americane
Narcisismo	Narrativa indiana	Natività <Iconografia>	Navi da guerra italiane	Navi da guerra italiane
Narcisismo nella letteratura	Narrativa inglese	Natività <Incisione>	Navi mercantili	Navi mercantili
Narcissismo	Narrativa iranica	Natura	Navi militari	Navi militari americane
Narcosissima	Narrativa ispano-americana	Natura dei fatti	Navi militari americane	Navi militari americane
Narcosi	Narrativa ispanoamericana	Natura delle cose	Navi militari a propulsione nucleare	Navi militari a propulsione nucleare
Narcotina	Narrativa italiana	Natura morta	Navi militari italiane	Navi militari italiane
Narrativa	Narrativa iiddish	Natura morta nella pittura	Navi porta container	Navi porta container
Narrativa albanese	Narrativa macedone	Natura morta nella pittura	Navi portaerai	Navi portaerai
Narrativa americana	Narrativa nederlandese	italiana	Natura nella letteratura	Natura nella letteratura
Narrativa araba	Narrativa neerlandese	Natura morta nella pittura	Navi romane	Navi romane
Narrativa bulgara	Narrativa neogreca	lombarda	Navi-traghetti	Navi-traghetti
Narrativa catalana	Narrativa olandese	Natura morta nella pittura		
Narrativa cinese	Narrativa osmanica	pompeiana		
Narrativa croata	Narrativa osmanli	Natura nell'arte		
Narrativa danese	Narrativa ottomana	Natura nella poesia		
Narrativa dell'orrore	Narrativa per bambini	Natura umana		
Narrativa dell'orrore	Narrativa per l'infanzia	Nastri d'acciaio		

Figura 43: A dimensione 18pt lo schermo 27" visualizza un massimo di 190 termini, disposti su 38 linee e 5 colonne.

Pagine

Non potendo visualizzare tutti i termini all'interno di una singola schermata per motivi di spazio, si dovrà introdurre una funzionalità che permetta di navigare nel thesaurus. A tal proposito vorremmo ispirarci alla fruizione di un thesaurus cartaceo, motivo per cui adotteremo la metafora delle *pagine* (vedi p. 29).

Il concetto di pagina deve essere così adattato allo schermo: gli impaginati di termini introdotti saranno le nostre *pagine*, le unità elementari sulle quali la navigazione sarà basata. Le pagine saranno inoltre ordinate in successione in modo da formare una sequenza navigabile.

Mentre i sistemi operativi che conosciamo sono basati sul paradigma della *pergamena* – meglio conosciuto con il termine inglese *scroll* – secondo cui un ordinamento verticale dà ordine al testo, nel nostro caso l'ordinamento sarà dunque orizzontale, similmente alle pagine del libro disposte prima a sinistra e poi a destra per ripetere lo stesso *pattern* ogni volta che la pagina viene girata. Così come talvolta capita nei musei, quando vengono esposte le pagine di libri rimosse dalla propria rilegatura, le nostre schermate saranno posizionate in una successione di tavole da sinistra verso destra (vedi fig. 44).



Figura 44: I 1000 termini del Nuovo Soggettario che iniziano con la lettera **B** rappresentati su schermo 27" in tre schermate ordinate orizzontalmente.

La navigazione consiste in uno scorrimento delle pagine in senso orizzontale. La transazione da una pagina all'altra è gestita da un'effetto di scorrimento che trasmetta la sensazione spaziale dello scorrere delle pagine. L'effetto viene gestito attraverso tre differenti modalità a seconda della *device* che si sta utilizzando. Con il *mouse* appaiono delle frecce sui bordi laterali dello schermo, premendole si attiverà la transazione che porterà a una delle pagine limitrofe. Con la *tastiera* sarà sufficiente usare le combinazioni *ctrl* e \rightarrow oppure *ctrl* e \leftarrow per attivare la transazione. Con strumenti *touchscreen* come l'iPad sarà sufficiente trascinare un dito orizzontalmente sulla superficie sensibile, portando fuori la pagina corrente a favore della successiva che spunterà dalla parte opposta alla direzione di trascinamento.

A supportare questo tipo di navigazione verrà inserita nell'interfaccia una *barra di navigazione* (fig. 45). Attraverso una maniglia posto alla base dello schermo, l'utente potrà attivare un menù contestuale di na-

vigazione delle pagine (vedi figg. 46 e 47). Lo spazio della barra di navigazione sarà di altezza fissa e visualizzerà il thesaurus in maniera astratta, cercando di arricchire il contesto del vocabolario controllato.

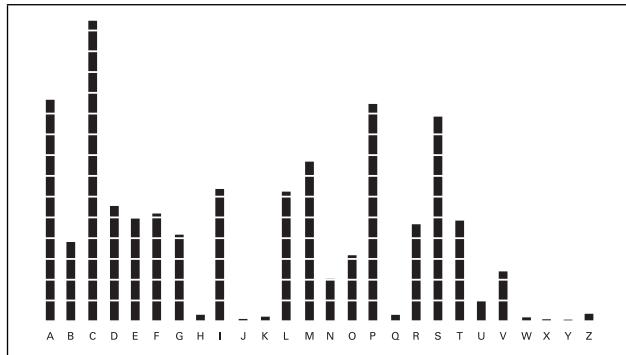


Figura 45: La barra di navigazione suddivide i termini prima in classi, poi in pagine.

Molti software moderni come *iPhoto* per la gestione delle immagini o programmi di montaggio video come *Premiere* dispongono di ormai consolidate interfacce di navigazione per fotografie o *frame video* che si basano su una serie di *thumbnail* selezionabili e navigabili. Abbiamo pensato di applicare una funzione simile alla nostra interfaccia, astraendo completamente il thesaurus e trasformandolo in un grafico informativo capace anche di funzionare come navigazione.

Con l'aggiunta di questa funzionalità andremo ad ampliare il contesto della pagina dei termini, così da permettere un'orientamento nella completa estensione del thesaurus. Infatti la barra di navigazione consente di muoversi tra le pagine attraverso una rappresentazione basata sulle seguenti informazioni:

- La divisione in classi
- Le pagine in una classe
- Il numero di termini presenti nella pagina

La barra di navigazione viene creata applicando una prima divisione per classi, per esempio nella figura 45 sono le lettere dell'alfabeto ad essere usate come classi. Le barre verticali, che indicano la quantità di termini presenti nella singola classe, sono frammentate a seconda del numero di pagine necessarie alla visualizzazione dei termini. Il rettangolo più prossimo alla lettera indica la prima pagina della sequenza. Inoltre i rettangoli si comportano come bottoni: fermando il mouse sopra di essi si aprirà una finestra di dettaglio che mostra la prima e

l'ultima parola della pagina, cliccando sul bottone la pagina corrente verrà sostituita da quella selezionata.

A questo punto va comunque fatta notare una questione di visualizzazione. Avevamo preso in considerazione l'opportunità di inserire una *numerazione delle pagine*. Tuttavia abbiamo ritenuto la cosa fuorviante dal fatto che abbiamo diversi parametri che vanno a modificare il numero di termini in una pagina e quindi il numero di pagine totale. Utilizzare una numerazione come nei libri sarebbe stato come dare dei riferimenti instabili, che equivale a fornire indicazioni fuorvianti all'utente. Abbiamo così deciso di escludere il numero pagina e di andarlo a sostituire con elementi come la barra di navigazione delle pagine che secondo noi andranno meglio a descrivere il contesto di un thesaurus non più statico come nelle versioni cartacee, ma dinamico al pari dell'ambiente digitale che lo descrive.

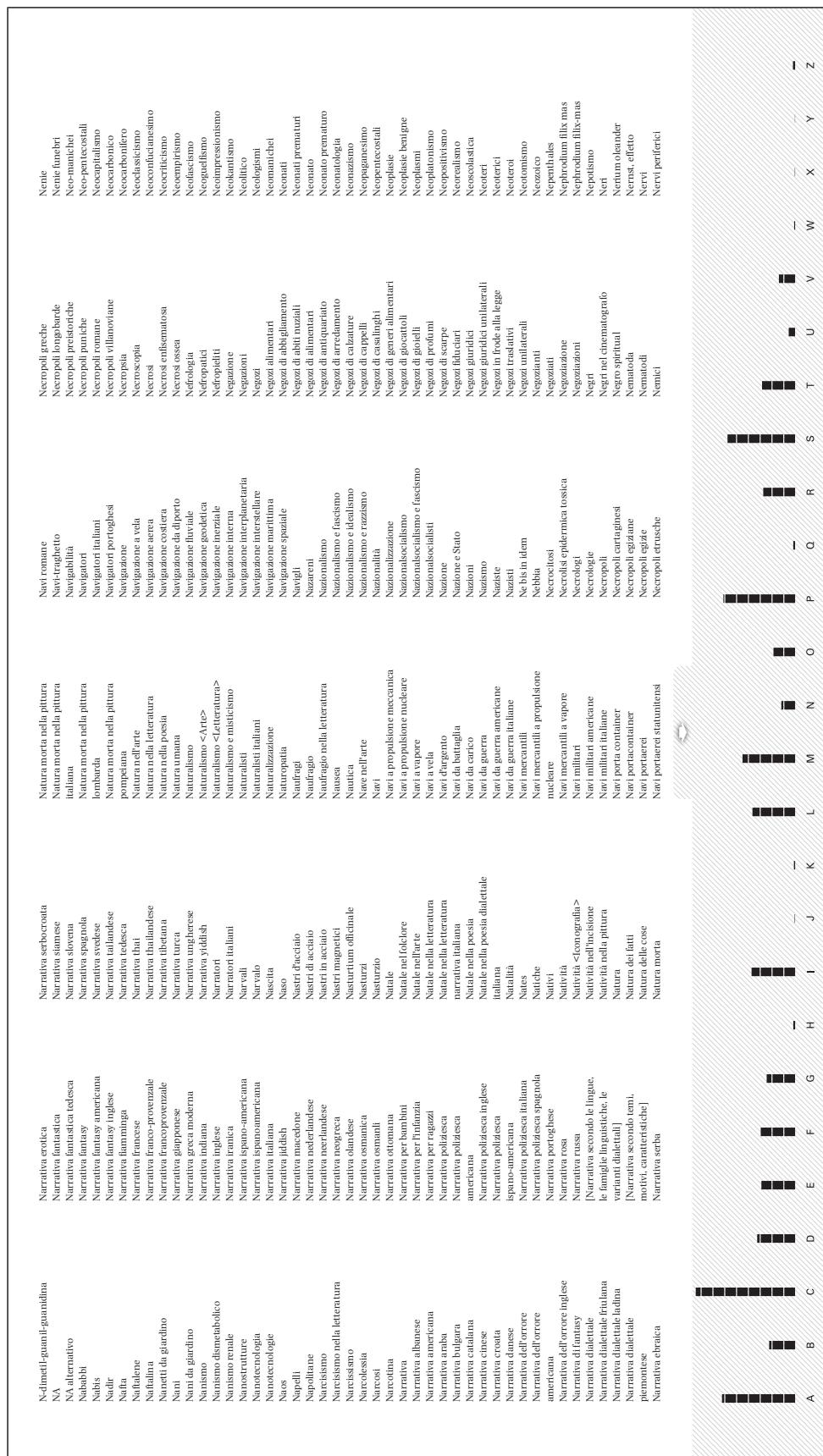


Figura 46: Lo schermo 27" con la barra di navigazione aperta riduce lo spazio dedicato alla liste da 399 a 329 termini. La linguetta posta tra i termini e la navigazione permette la chiusura della barra.

Figura 47: Lo schermo 27" con la barra di navigazione chiusa attivabile attraverso la linguetta posta alla base dello schermo.



Le Caratteristiche del Termine

I termini rappresentati nella visualizzazione sono solo l'estensione visiva della struttura che compongono all'interno del thesaurus, dietro ad ogni termine esiste una complessità che deve essere a disposizione dell'utente per comprenderne il significato. La lista è un elemento semplice, tuttavia fornisce informazioni che aiutano nella comprensione del termine.

Preferiti e non preferiti

Per esempio una lista può essere arricchita introducendo una distinzione tra *termini preferiti* e *termini non preferiti* (vedi p. 9). I thesauri su carta differenziano i termini attraverso l'uso della tipografia: il carattere *normale* denota i termini preferiti, mentre il *corsivo* è associato ai termini non preferiti (vedi fig. 48).

<i>N-dimetil-quanil-guanidina</i>	<i>Narrativa cinesi</i>	<i>Narrativa nederlandese</i>	<i>Narrativa thai</i>
<i>NA</i>	<i>Narrativa croata</i>	<i>Narrativa neerlandese</i>	<i>Narrativa thailandese</i>
<i>NA alternativo</i>	<i>Narrativa danese</i>	<i>Narrativa greca</i>	<i>Narrativa tibetana</i>
<i>Nababbi</i>	<i>Narrativa dell'orrore</i>	<i>Narrativa olandese</i>	<i>Narrativa turca</i>
<i>Nabis</i>	<i>Narrativa dell'orrore americana</i>	<i>Narrativa osmanica</i>	<i>Narrativa ungherese</i>
<i>Nadir</i>	<i>Narrativa dell'orrore inglese</i>	<i>Narrativa osmanli</i>	<i>Narrativa yiddish</i>
<i>Nafta</i>	<i>Narrativa di fantasy</i>	<i>Narrativa ottomana</i>	<i>Narratori</i>
<i>Naftalene</i>	<i>Narrativa dialettale</i>	<i>Narrativa per bambini</i>	<i>Narratori italiani</i>
<i>Naftalina</i>	<i>Narrativa dialettale friulana</i>	<i>Narrativa per l'infanzia</i>	<i>Narvali</i>
<i>Nanetti da giardino</i>	<i>Narrativa dialettale ladina</i>	<i>Narrativa per ragazzi</i>	<i>Narvalo</i>
<i>Nani</i>	<i>Narrativa dialettale</i>	<i>Narrativa poliziesca</i>	<i>Nascita</i>
<i>Nani da giardino</i>	<i>piemontese</i>	<i>Narrativa poliziesca americana</i>	<i>Naso</i>
<i>Nanismo</i>	<i>Narrativa ebraica</i>	<i>Narrativa poliziesca inglese</i>	<i>Nastri d'acciaio</i>
<i>Nanismo dismetabolico</i>	<i>Narrativa erotica</i>	<i>Narrativa poliziesca</i>	<i>Nastri di acciaio</i>
<i>Nanismo renale</i>	<i>Narrativa fantastica</i>	<i>ispano-americana</i>	<i>Nastri in acciaio</i>
<i>Nanostrutture</i>	<i>Narrativa fantastica tedesca</i>	<i>Narrativa poliziesca italiana</i>	<i>Nastri magnetici</i>
<i>Nanotecnologia</i>	<i>Narrativa fantasy</i>	<i>Narrativa poliziesca spagnola</i>	<i>Nasturtium officinale</i>
<i>Nanotecnologie</i>	<i>Narrativa fantasy americana</i>	<i>Narrativa portoghese</i>	<i>Nasturzì</i>
<i>Naos</i>	<i>Narrativa fantasy inglese</i>	<i>Narrativa rosa</i>	<i>Nasturzì</i>
<i>Napelli</i>	<i>Narrativa fiamminga</i>	<i>Narrativa russa</i>	<i>Natate</i>
<i>Napolitane</i>	<i>Narrativa francese</i>	<i>[Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]</i>	<i>Natate nel folclore</i>
<i>Narcisismo</i>	<i>Narrativa franco-provenzale</i>		<i>Natate nell'arte</i>
<i>Narcisismo nella letteratura</i>	<i>Narrativa franco-provenzale</i>		<i>Natate nella letteratura</i>
<i>Narcissismo</i>	<i>Narrativa giapponese</i>	<i>[Narrativa secondo temi, motivi, caratteristiche]</i>	<i>Natate nella letteratura narrativa</i>
<i>Narcolessia</i>	<i>Narrativa greca moderna</i>		<i>italiana</i>
<i>Narcosi</i>	<i>Narrativa indiana</i>	<i>Narrativa serba</i>	<i>Natate nella poesia</i>
<i>Narcotina</i>	<i>Narrativa inglese</i>	<i>Narrativa serbocroata</i>	<i>Natate nella poesia dialettale</i>
<i>Narrativa</i>	<i>Narrativa iranica</i>	<i>Narrativa siamese</i>	<i>italiana</i>
<i>Narrativa albanese</i>	<i>Narrativa ispano-americana</i>	<i>Narrativa slovena</i>	<i>Nataltà</i>
<i>Narrativa americana</i>	<i>Narrativa ispanoamericana</i>	<i>Narrativa spagnola</i>	<i>Nates</i>
<i>Narrativa araba</i>	<i>Narrativa italiana</i>	<i>Narrativa svedese</i>	<i>Natiche</i>
<i>Narrativa bulgara</i>	<i>Narrativa jiddish</i>	<i>Narrativa tailandese</i>	<i>Nativi</i>
<i>Narrativa catalana</i>	<i>Narrativa macedone</i>	<i>Narrativa tedesca</i>	<i>Natività</i>

Figura 48: I caratteri normale e corsivo distinguono termini preferiti e non preferiti, nell'esempio la prima pagina della lettera *N* su uno schermo di 15,4"

Selezione

Poter selezionare un termine è un comportamento che è necessario introdurre per visualizzare le caratteristiche di un termine e per manipolare la struttura del thesaurus, in particolare le relazioni. Posizionare il fuoco su un elemento è una comportamento comune ai sistemi informatici [56], ora vedremo come implementarlo in un sistema di gestione per thesauri.

Per definizione la selezione distingue un termine da tutti gli altri. Un buon principio della tipografia consiste nell'applicare una sola variazione di stile per definire un sola differenza. Questo significa che un solo carattere distintivo è sufficiente per riconoscere il termine selezionato. Nella nostra interfaccia il termine selezionato sarà in *grassetto* (vedi fig. 49). La scelta fatta evita di introdurre l'uso dei colori all'interno dell'interfaccia, rendendo così l'applicazione accessibile anche alle persone affette da daltonismo [23].

<i>N-dimetil-quanil-quanidina</i>	<i>Narrativa cinese</i>	<i>Narrativa nederlandese</i>	<i>Narrativa thai</i>
<i>NA</i>	<i>Narrativa croata</i>	<i>Narrativa neerlandese</i>	<i>Narrativa thailandese</i>
<i>NA alternativo</i>	<i>Narrativa danese</i>	<i>Narrativa neogreca</i>	<i>Narrativa tibetana</i>
<i>Nababbi</i>	<i>Narrativa dell'orrore</i>	<i>Narrativa olandese</i>	<i>Narrativa turca</i>
<i>Nabis</i>	<i>Narrativa dell'orrore americana</i>	<i>Narrativa osmanica</i>	<i>Narrativa ungherese</i>
<i>Nadir</i>	<i>Narrativa dell'orrore inglese</i>	<i>Narrativa osmanli</i>	<i>Narrativa yiddish</i>
<i>Nafta</i>	<i>Narrativa di fantasy</i>	<i>Narrativa ottomana</i>	<i>Narratori</i>
<i>Naftalene</i>	<i>Narrativa dialettale</i>	<i>Narrativa per bambini</i>	<i>Narratori italiani</i>
<i>Naftalina</i>	<i>Narrativa dialettale friulana</i>	<i>Narrativa per l'infanzia</i>	<i>Narvali</i>
<i>NaNetti da giardino</i>	<i>Narrativa dialettale ladina</i>	<i>Narrativa per ragazzi</i>	<i>Narvalo</i>
<i>Nani</i>	<i>Narrativa dialettale</i>	<i>Narrativa poliziesca</i>	<i>Nascita</i>
<i>Nani da giardino</i>	<i>piemontese</i>	<i>Narrativa poliziesca americana</i>	<i>Naso</i>
<i>Nanismo</i>	<i>Narrativa ebraica</i>	<i>Narrativa poliziesca inglese</i>	<i>Nastri d'acciaio</i>
<i>Nanismo dismetabolico</i>	<i>Narrativa erotica</i>	<i>Narrativa poliziesca</i>	<i>Nastri di acciaio</i>
<i>Nanismo renale</i>	<i>Narrativa fantastica</i>	<i>Ispano-americana</i>	<i>Nastri in acciaio</i>
<i>Nanostrutture</i>	<i>Narrativa fantastica tedesca</i>	<i>Narrativa poliziesca italiana</i>	<i>Nastri magnetici</i>
<i>Nanotecnologia</i>	<i>Narrativa fantasy</i>	<i>Narrativa poliziesca spagnola</i>	<i>Nasturium officinale</i>
<i>Nanotecnologie</i>	<i>Narrativa fantasy americana</i>	<i>Narrativa portoghese</i>	<i>Nasturzi</i>
<i>Nao</i>	<i>Narrativa fantasy inglese</i>	<i>Narrativa rosa</i>	<i>Nasturzio</i>
<i>Napelli</i>	<i>Narrativa fiamminga</i>	<i>Narrativa russa</i>	<i>Natale</i>
<i>Napolitane</i>	<i>Narrativa francese</i>	<i>[Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]</i>	<i>Natale nel folclore</i>
<i>Narcisismo</i>	<i>Narrativa franco-provenzale</i>	<i>[Narrativa secondo temi, motivi, caratteristiche]</i>	<i>Natale nell'arte</i>
<i>Narcisismo nella letteratura</i>	<i>Narrativa giapponese</i>	<i>Narrativa serba</i>	<i>Natale nella letteratura</i>
<i>Narcissismo</i>	<i>Narrativa greca moderna</i>	<i>Narrativa serbocroata</i>	<i>Natale nella letteratura narrativa italiana</i>
<i>Narcolessia</i>	<i>Narrativa indiana</i>	<i>Narrativa siamese</i>	<i>Natale nella poesia</i>
<i>Narcosi</i>	<i>Narrativa inglese</i>	<i>Narrativa slovena</i>	<i>Natale nella poesia dialettale italiana</i>
<i>Narcotina</i>	<i>Narrativa iranica</i>	<i>Narrativa spagnola</i>	<i>Natunità</i>
<i>Narrativa</i>	<i>Narrativa ispanica</i>	<i>Narrativa svedese</i>	<i>Nates</i>
<i>Narrativa albanese</i>	<i>Narrativa ispano-americana</i>	<i>Narrativa talandese</i>	<i>Natiche</i>
<i>Narrativa americana</i>	<i>Narrativa ispanoamericana</i>	<i>Narrativa tedesca</i>	<i>Nativi</i>
<i>Narrativa araba</i>	<i>Narrativa italiana</i>		
<i>Narrativa bulgara</i>	<i>Narrativa jiddish</i>		
<i>Narrativa catalana</i>	<i>Narrativa macedone</i>		

Figura 49: Il termine selezionato appare in grassetto. Una sola caratteristica è sufficiente a distinguere il termine selezionato.

La selezione si attiva con un click del mouse o il tocco di un dito su uno schermo *touchscreen*, e può essere modificata scegliendo un altro termine oppure muovendosi con i tasti cursore della tastiera: i tasti \uparrow e \downarrow serviranno a muovere la selezione da una riga all'altra, mentre \leftarrow e \rightarrow permetteranno di muoversi tra le colonne. Questo tipo di navigazione permetterà di muovere la selezione come se si trattasse di una lista unica: arrivati all'ultimo termine di una colonna e premendo \downarrow , la selezione sarà spostata al primo termine della colonna successiva.

va. Se il termine selezionato è nell'ultima colonna, premendo il cursore destro l'interfaccia provvederà automaticamente a caricare la pagina successiva, sfruttando la successione delle pagine del thesaurus.

Dettaglio

Selezionare un termine significa anche accedere alle caratteristiche che non è possibile mostrare nella lista. Una riflessione è sufficiente per capire che i dettagli di ogni termine necessitano di uno spazio dedicato. All'interno della nostra interfaccia abbiamo deciso di raggrupparle in un *pannello di dettaglio* posizionato sul lato destro dello schermo, similmente a quanto realizzato per pannello di navigazione (vedi p. 57).

Per esempio le caratteristiche del termine *Narrativa Russa*, il cui dettaglio è visibile nella tabella qui rappresentata, all'interno dell'interfaccia appariranno come illustrato nella figura 50.

Dettaglio di <i>Narrativa russa</i>	
Categoria/Faccetta	Cose:Forme
Relazioni	UF Letteratura narrativa russa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali] NT Racconti russi RT Letteratura russa UF+ Fantastico nella letteratura narrativa russa HSF Letteratura narrativa russa
Nota storica	Precedentemente: Letteratura narrativa russa; sulla base della voce esemplificativa: Letteratura narrativa italiana (americana, etc.) (BNI 1986-1998)
Fonte	BNI non registr.: Letteratura narrativa russa
Dewey, Ed. 21	891.73
Agenzia	BNI
cat./Proponente	
Status del record	Termine strutturato
Identificativo	3553

I pannelli devono poter essere visualizzati contemporaneamente, così da essere usati senza vincoli uno dall'altro (vedi fig. 51).

Figura 50: Il dettaglio del termine selezionato è disponibile in un pannello laterale dell'interfaccia. L'utente ha la possibilità di visualizzarlo o nasconderlo secondo le sue necessità.

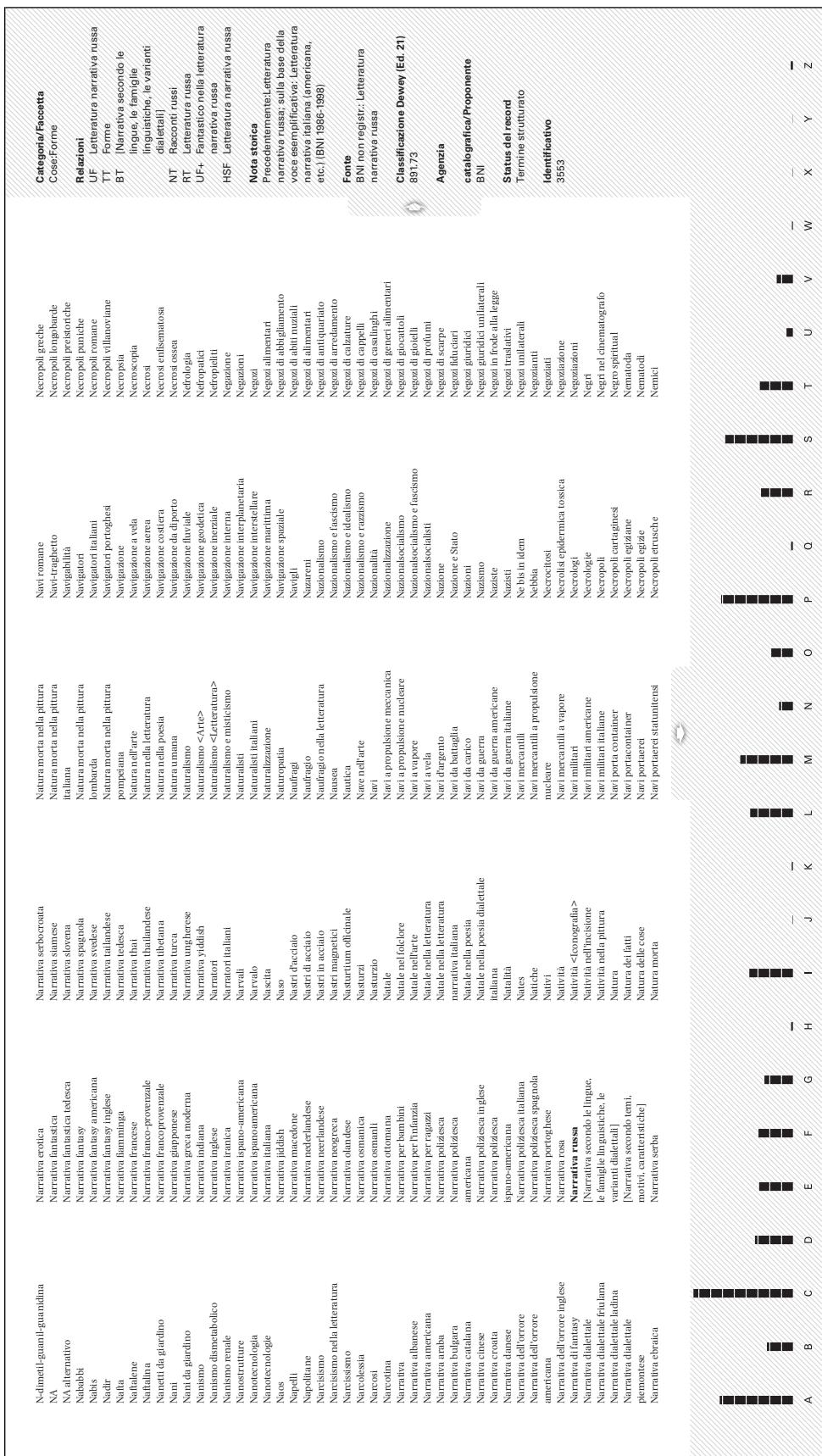
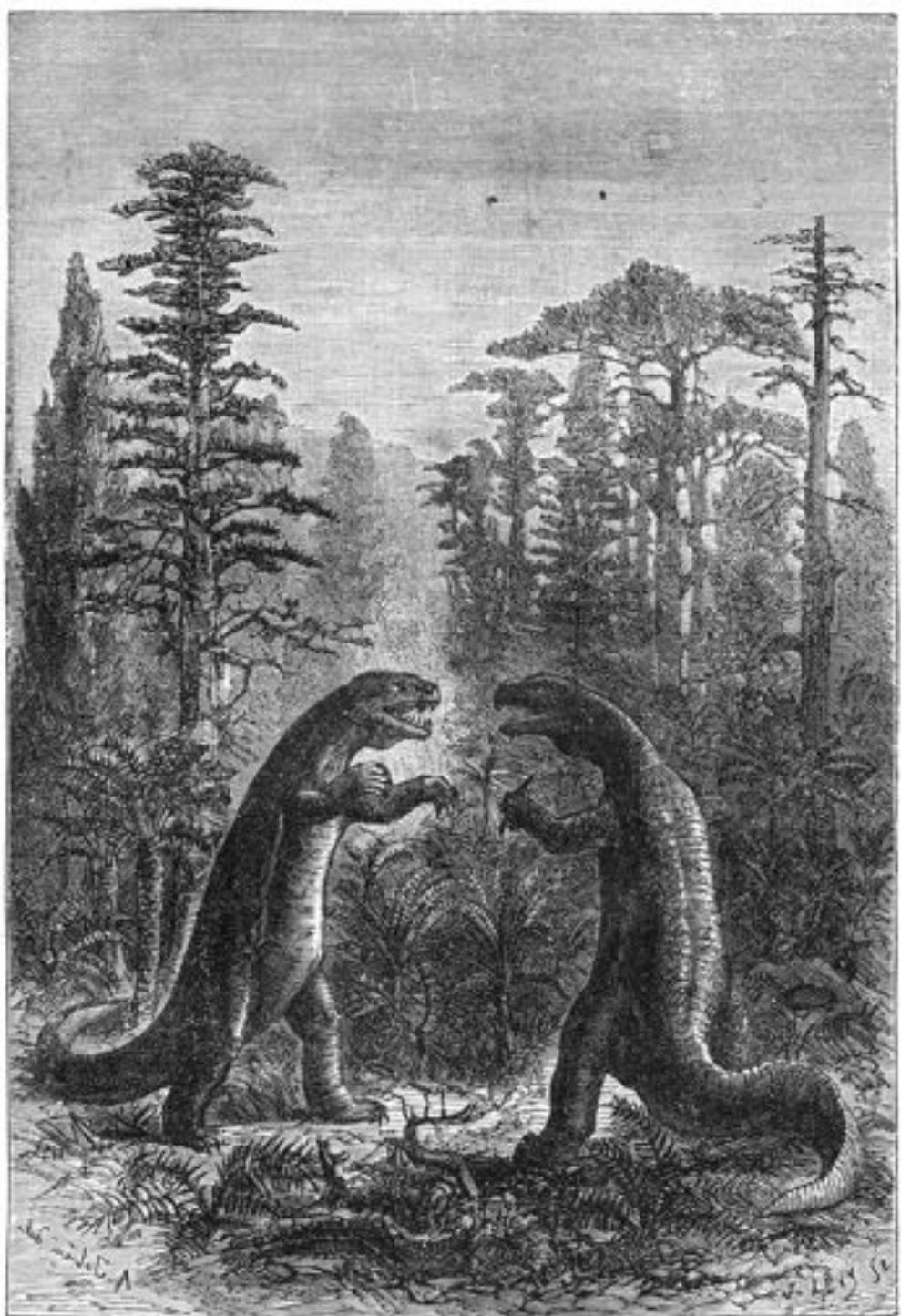


Figura 51: Le barre di supporto all’interfaccia devono essere in grado di operare in accordo all’interno dell’interfaccia.



L'Ordinamento

Uno dei grandi vantaggi del supporto digitale consiste nel rappresentare i contenuti in maniera dinamica. Quindi, se in passato il thesaurus veniva stampato scegliendo un'unico modo di organizzare i termini e le relazioni, progettare un thesaurus su supporto digitale significa consentire diversi tipi di organizzazione di termini e relazioni, permettendo all'utente il passaggio da una all'altra nella più assoluta libertà.

Essendo la visualizzazione di termini e relazioni strettamente correlata, la visualizzazione risultante dovrà trasmettere questo vincolo di dipendenza. Ricordiamo infatti che nel thesaurus il significato di un termine è espresso dalle relazioni e dai documenti che sottende. Data che il valore semantico di un termine è dato dalle relazioni, nella fase di visualizzazione sarà importante considerare termini e relazioni congiuntamente.

Introdotto questo vincolo sussidiario, esemplificheremo prima i metodi di rappresentazione delle relazioni. Solo successivamente approfondiremo l'ordinamento dei termini in tutte le declinazioni utili.

Relazioni

Le relazioni servono a definire il significato e l'uso del termine. Visualizzare le relazioni nell'interfaccia diminuisce il numero di termini visibili nello schermo ma completa il significato di ognuno di loro.

Le relazioni sono divise in tre gruppi distinti: le relazioni di equivalenza, gerarchiche e associative (vedi p. 9). Nei thesauri cartacei le relazioni vengono elencate in un'ordine predefinito. Il Nuovo soggettario aderisce allo standard ISO e lo estende inserendo le proprie relazioni specifiche [2]. Nella tabella di pagina 70 vengono visualizzate le relazioni per il termine *Narrativa russa*.

Relazioni associate a <i>Narrativa russa</i>	
UF	Letteratura narrativa russa
TT	Forme
BT	[Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
NT	Racconti russi
RT	Letteratura russa
UF+	Fantastico nella letteratura narrativa russa
HSF	Letteratura narrativa russa

Di norma, le relazioni relative ad un termine vengono visualizzate sotto ad esso. Termini e relazioni si distinguono per posizione grazie all'*indentazione*, cioè lo spazio vuoto posto all'inizio della riga di testo. Complessivamente ogni unità visiva è costituita da diverse righe indennate in maniera differente: la prima riga corrisponde al termine, esso è caratterizzato da un'indentazione di zero pixel; a seguire troveranno posto tutte le relazioni composte da acronimo della relazione e termine corrispondente.

Il termine con le sue relazioni assumerà forma come nella figura 52, mentre la schermata completa risulterà come nella figura a pagina 72.

Narrativa russa
UF Letteratura narrativa russa
TT Forme
BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
NT Racconti russi
RT Letteratura russa
UF+ Fantastico nella letteratura narrativa russa
HSF Letteratura narrativa russa

Figura 52: I termini e le relazioni vengono raggruppati in un'unico blocco visuale. L'indentazione viene utilizzata per enfatizzare la differenza tra termini e relazioni.

La rappresentazione standard delle relazioni non è la sola, esistono altri modi di visualizzazione, specialmente per le relazioni gerarchiche. Nella nostra indagine, abbiamo ritenuto utile raccogliere altri tipi di visualizzazione che potrebbero contribuire a migliorare l'interfaccia:

- Relazioni gerarchiche
- Relazioni gerarchiche a più livelli
- Relazioni gerarchiche a doppia direzione

Relazioni gerarchiche

Nel primo caso le relazioni visualizzate sono di tipo gerarchico. Si tratta di una semplificazione della visualizzazione standard con l'obiettivo di mostrare più termini e allo stesso tempo porre più attenzione alla gerarchia del thesaurus (vedi fig. 54).

Relazioni gerarchiche a più livelli

La visualizzazione dei termini nidificati può essere arricchita inserendo un livello di profondità aggiuntivo, rappresentato dalle abbreviazioni BT2 e NT2 (vedi fig. 55).

Relazioni gerarchiche a doppia direzione

Un'ultima alternativa sta nel visualizzare le relazioni gerarchiche sfruttando una doppia direzione in sostituzione di quella singola. L'idea consiste nel posizionare le relazioni padre al di sopra e le relazioni figlio al di sotto del termine corrispondente, così da ricostruire una sequenza simile ad un'alberatura. Proprio grazie a questa struttura potremo eliminare le abbreviazioni BT e NT, sostituendole con simboli grafici che opportunamente moltiplicati andranno ad indicare il livello della relazione (vedi fig. 56).

Figura 53: La prima pagina della lettera N appare così se si sceglie di visualizzare tutte le relazioni di un termine.

Nababbi	TT Personae BT [Personae secondo la condizione sociale]	Narrativa bulgara TT Forme BT [Narrativa secondo le dialetticità del sistema nervoso]	Narrativa dialettale piemontese TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa inglese TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nabis	TT Personae BT Pittori	Narrativa TT Materia BT Alcalolidi	Narrativa TT Forme BT [Cerchi laterali] secondo la forma]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nadir	TT Processi	Narrativa TT Forme BT [Cerchi laterali] secondo la forma]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nalla	TT Materia	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nathalie	TT Materia	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nari	TT Personae BT [Personae secondo le condizioni fisiche o mentali]	Narrativa albanese TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nari da giardino	TT Oggetti BT Statue	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Naufragio	TT Processi BT [Mallate di organi, apparati, sistemi, fondelli]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Naufragio dismetabolico	NT	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Naufragio dismetabolico	TT Processi	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Naufragio	BT Naufrago	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Naufragio	TT Discipline BT Tecnologie	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Nas	NT	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Natureologie	TT Strutture BT Edifici di culto	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]
Narcissimo	TT Processi BT [Processi relativi alla psiche e alla mente]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]	Narrativa TT Forme BT [Narrativa secondo le lingue, le famiglie linguistiche, le varianti dialettali]

Figura 54: La prima pagina della lettera N appare così se si sceglie di visualizzare le relazioni gerarchiche di un termine.

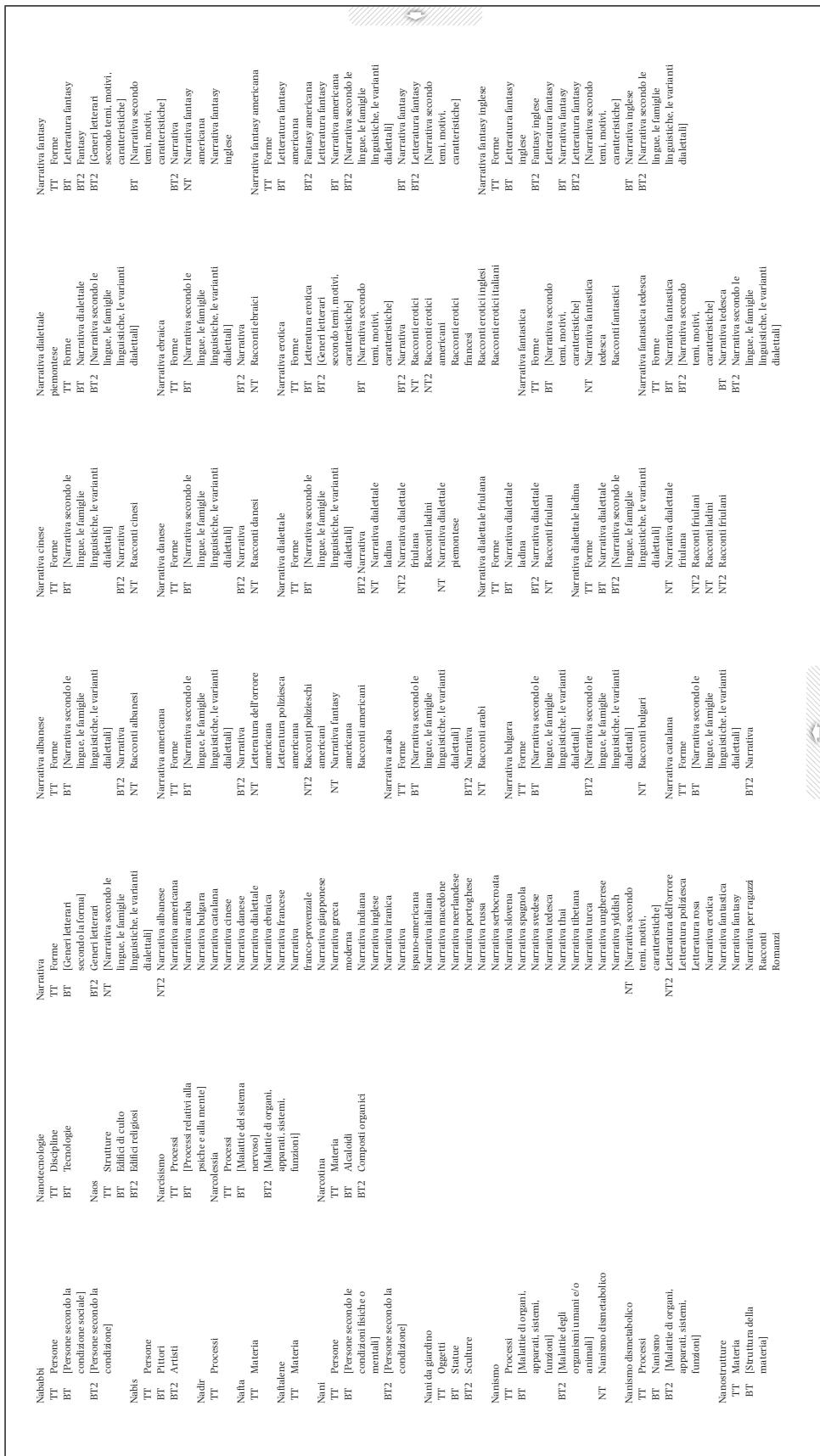


Figura 55: La prima pagina della lettera N appare così se si sceglie di visualizzare le relazioni a più livelli di un termine.

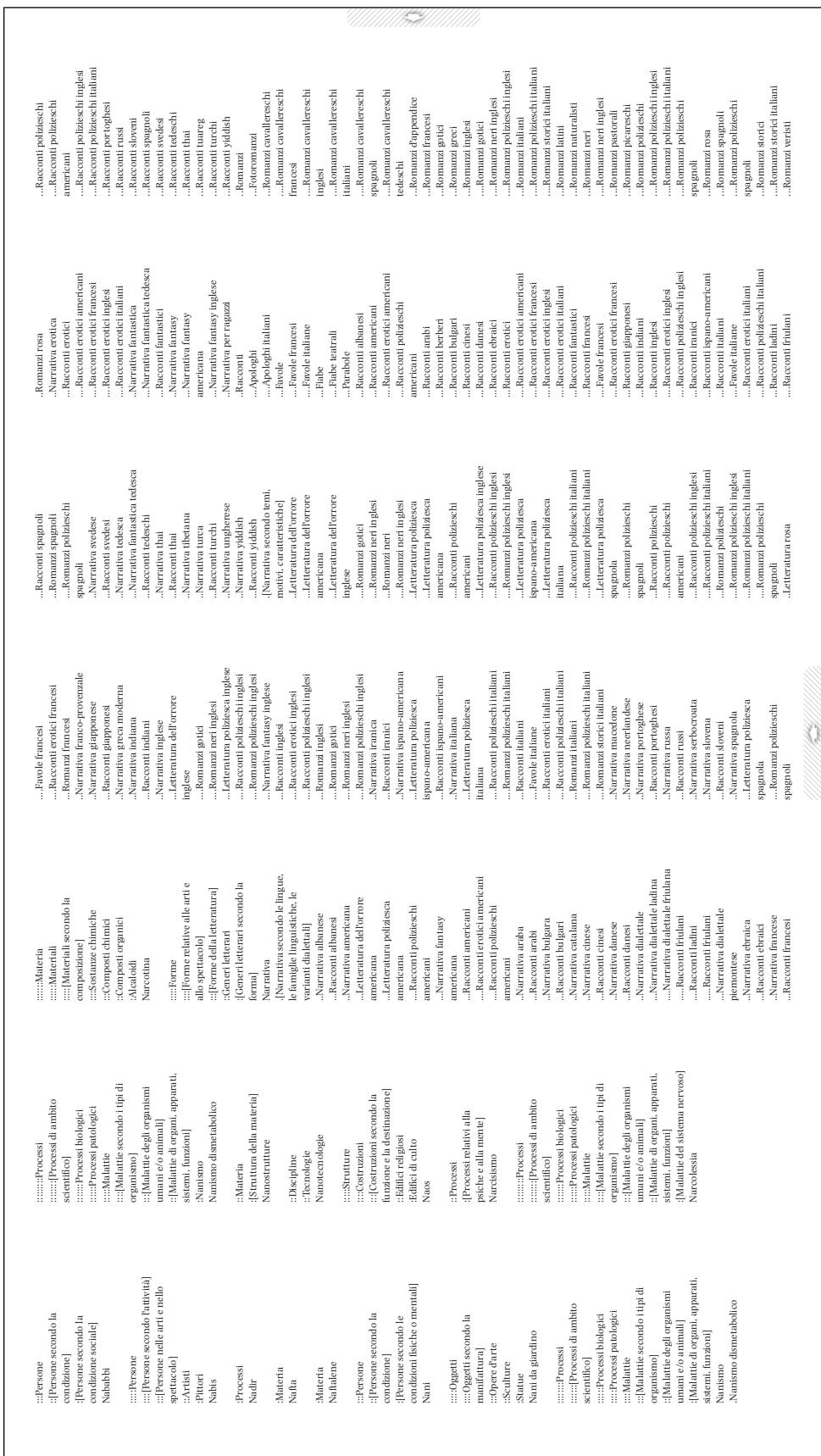


Figura 56: La prima pagina della lettera N appare così se si sceglie di visualizzare le relazioni a doppia direzione di un termine.

Termini

Dopo aver introdotto vari modi di visualizzare le relazioni, ora possiamo occuparci dell'ordinamento dei termini. Avere la possibilità di ordinare i termini diversamente permette all'utente di accedere al thesaurus da diversi punti di vista. L'importanza della varietà di ordinamenti sta nell'integrare le visualizzazioni, così da trasmettere la struttura complessa che costituisce il thesaurus.

Dopo aver letto, annotato, elencato, analizzato e sfoltito gli studi di Aitchinson e Ranganathan [1] [46] [44], sono stati individuati cinque tipi diversi di ordinamento dei termini qui di seguito elencati:

- Ordine alfabetico
- Ordine alfabetico permutato
- Ordine sistematico
- Ordine cronologico
- Ordine quantitativo

Ordine alfabetico

L'ordine alfabetico è il più diffuso ed è stato utilizzato finora nelle figure a supporto del testo, ma esistono due tipi diversi di ordinamento alfabetico:

- Parola per parola
- Lettera per lettera

Per essere precisi, finora negli esempi è stato usato un ordinamento *parola per parola*, ovvero un ordinamento dove lo spazio conta come una lettera dell'alfabeto e pesa nella sequenza in quanto carattere. Tuttavia esiste anche un ordinamento meno diffuso, che potrebbe essere preferito in ambiti di lavoro specifici o per particolari lingue: è il *lettera per lettera*, dove lo spazio non sarà valutato nella creazione della sequenza di termini. Riteniamo sia interessante permettere all'utente di scegliere tra questi due ordinamenti, permettendogli così di personalizzare la visualizzazione dei termini a seconda delle proprie esigenze.

Ordine alfabetico permutato

In alternativa all'ordinamento alfabetico si può usare l'ordine alfabetico permutato che ordina alfabeticamente i termini e le rispettive permutazioni. Per esempio il termine *Narrativa russa* non sarà rappresentato da un solo termine, ma anche dalla sua unica permutazione (*Russa, Narrativa*). In questo modo la lista dei termini verrà estesa con tutte le permutazioni, andando a costruire una nuova sequenza di termini. Dato che questo ordinamento introduce una notevole espansione della lista, riteniamo che la funzionalità debba essere attivata manualmente dall'utente così da lasciare la scelta di utilizzarla nei momenti necessari.

Ordine sistematico

Nel Nuovo soggettario l'ordine sistematico è legato all'appartenenza dei termini a categorie e faccette qui elencati:

- Agenti:Organismi
- Agenti:Organizzazioni
- Agenti:Persone e gruppi
- Azioni:Attività
- Azioni:Discipline
- Azioni:Processi
- Cose:Forme
- Cose:Materia
- Cose:Oggetti
- Cose:Spazio
- Cose:Strumenti
- Cose:Strutture
- Tempo

Se nel thesaurus cartaceo questa vista è rappresentata solitamente con la divisione dei termini in paragrafi, nella nostra interfaccia ben si lega allo spazio di navigazione basato sulle classi che abbiamo già descritto a pagina 57. La barra di navigazione potrà quindi essere usata anche per navigare le classi (vedi fig. 58).

Ordine cronologico

I termini hanno una data di inserimento nella struttura del thesaurus, questa può essere utilizzata per ordinare gli elementi a seconda siano più o meno recenti. È un modo utile per andare a reperire gli ultimi termini inseriti (vedi fig. 59).

Ordine quantitativo

Se il thesaurus è connesso a un sistema bibliotecario, allora l'interfaccia può accedere a dati riguardanti l'uso dei termini. I dati nell'ordinamento possono essere impiegati per organizzare i termini dal più al meno utilizzato e viceversa. Si andrebbe così a creare una lista per cercare i termini con una maggiore ricorrenza tra i documenti. Ampliando il discorso al pannello di dettaglio del termine già introdotto a pagina 65, potremmo addirittura immaginare che questa connessione possa anche visualizzare i libri classificati col termine selezionato, così da ampliare la comprensione circa l'uso del termine (vedi fig. 60).

Dinamicità degli ordinamenti

Trasformare uno strumento da statico a dinamico apre la porta a soluzioni che prima parevano impossibili. Nel thesaurus cartaceo non è stato mai possibile modificare l'ordinamento dei termini, al massimo si ricorre a ordinamenti stampati su volumi diversi. Il mondo digitale ha il pregio di essere dinamico anziché statico e noi sfrutteremo la caratteristica in modo da permettere all'utente di muoversi da un ordinamento all'altro con semplicità.

Dopo aver individuato i quattro tipi di ordinamento cercheremo di capire come muoversi tra essi. Ci sono innanzitutto due casi da distinguere: quello in cui un termine è selezionato e quello in cui nessun termine è selezionato. Il primo è decisamente interessante e sarà maggiormente approfondito, il secondo invece corrisponde a un comportamento più generico e meno interessante.

Proviamo a pensare una pagina del thesaurus dove l'ordinamento applicato ai termini è alfabetico. Siamo alla prima pagina della lettera Q e abbiamo selezionato la parola *Quadri*. Se desiderassimo cambiare l'ordinamento dei termini da alfabetico a sistematico, la parola *Quadri* fungerebbe da perno nel cambiamento, ricalcolando il contesto attorno ad essa. Nel paragrafo in cui si analizzano le differenze tra thesaurus cartaceo e digitale si è data grande importanza al contesto di un singolo

termine (vedi p. 29). Applicando il cambio di ordinamento riusciamo a dare più profondità alla comprensione della struttura del thesaurus.

Il cambiamento di ordinamento nel caso nessun termine sia selezionato è meno utile. Infatti, se si confronta con l'esempio precedente, non essendoci nessun termine su cui fare perno il contesto non potrà essere calcolato. Come risultato l'utente sarà mandato alla prima pagina dell'ordinamento prescelto senza sfruttare la dinamicità del cambio di contesto.

Scelta delle relazioni

La dinamicità ci permette anche di personalizzare le relazioni a seconda della necessità. Abbiamo visto come nelle figure 53, 54, 55 e 56 per le relazioni siano disponibili diversi tipi di visualizzazione e come esse siano correlate agli ordinamenti. Sarà quindi fondamentale, ai fini dell'ottimizzazione dello spazio disponibile, mettere l'utente in grado di poter scegliere le relazioni da visualizzare. Per quanto riguarda il Nuovo soggettario, potranno essere attivate queste relazioni:

- Relazioni gerarchiche di primo livello
- Relazioni gerarchiche di primo e secondo livello
- Relazioni gerarchiche a doppia direzione
- Relazioni d'equivalenza
- Relazioni associative
- Variante storica
- Termini composti non preferiti

L'aspetto della pagina cambierà a seconda delle relazioni che verranno selezionate dall'utente. Le opzioni saranno visualizzate in una lista, dove sarà possibile attivarle e disattivarle. Ovviamente non tutte le opzioni saranno attivabili contemporaneamente. Per esempio le *relazioni gerarchiche di primo livello* e le relazioni gerarchiche di primo e secondo livello potranno essere scelte in maniera esclusiva. Invece, per esempio, le *relazioni gerarchiche di primo livello* saranno visibili assieme alle *relazioni d'equivalenza* o in aggiunta anche alle *relazioni associative*. Spetterà al sistema sapere quali combinazioni sono possibili, il che vuol dire che ad ogni selezione dell'utente verranno disattivate le opzioni in conflitto.

Visualizzazione grafica

La lista di termini non è l'unica modalità con cui rappresentare un thesaurus. Un'alternativa usata nel thesaurus EURATOM consiste nella rappresentazione attraverso i grafi. Nell'anno 1957, in cui viene istituita la Comunità Economica Europea, nasce anche la Comunità Europea dell'Energia Atomica [15]. Il Joint Research Centre è l'istituzione che si occupa di portare avanti la ricerca nel campo nucleare. Nel contesto di questo lavoro si decide di avviare lo studio e la realizzazione di un thesaurus specializzato, che prenderà il nome di thesaurus EURATOM [16].

Il thesaurus EURATOM è composto da un insieme di tavole divise a seconda dell'argomento. Nella figura 57 possiamo vedere la tavola 53 *Reaction Operation*. All'interno della tavola prendono posto i termini inerenti all'argomento, ogni termine preferito viene posizionato in una cornice nella quale vengono indicate le occorrenze del termine all'interno dell'archivio. A fianco del termine preferito trovano spazio gli eventuali termini non preferiti, che si distinguono tipograficamente dal termine preferito stampato in grassetto maiuscolo. Le linee che uniscono i termini rappresentano le relazioni tra essi. Maggiormente due termini sono inerenti, maggiore è lo spessore della linea che li unisce. Inoltre i contorni della tavola indicano relazioni esistenti tra i termini vicini al bordo e altre tavole del thesaurus. Questo meccanismo non solo permette di vedere se ci sono relazioni pertinenti in altre tavole, ma anche di muoversi agevolmente tra una tavola e l'altra, senza ricorrere sempre a un indice principale.

Il thesaurus EURATOM è spesso citato come buon esempio di design nella rappresentazione dei termini.

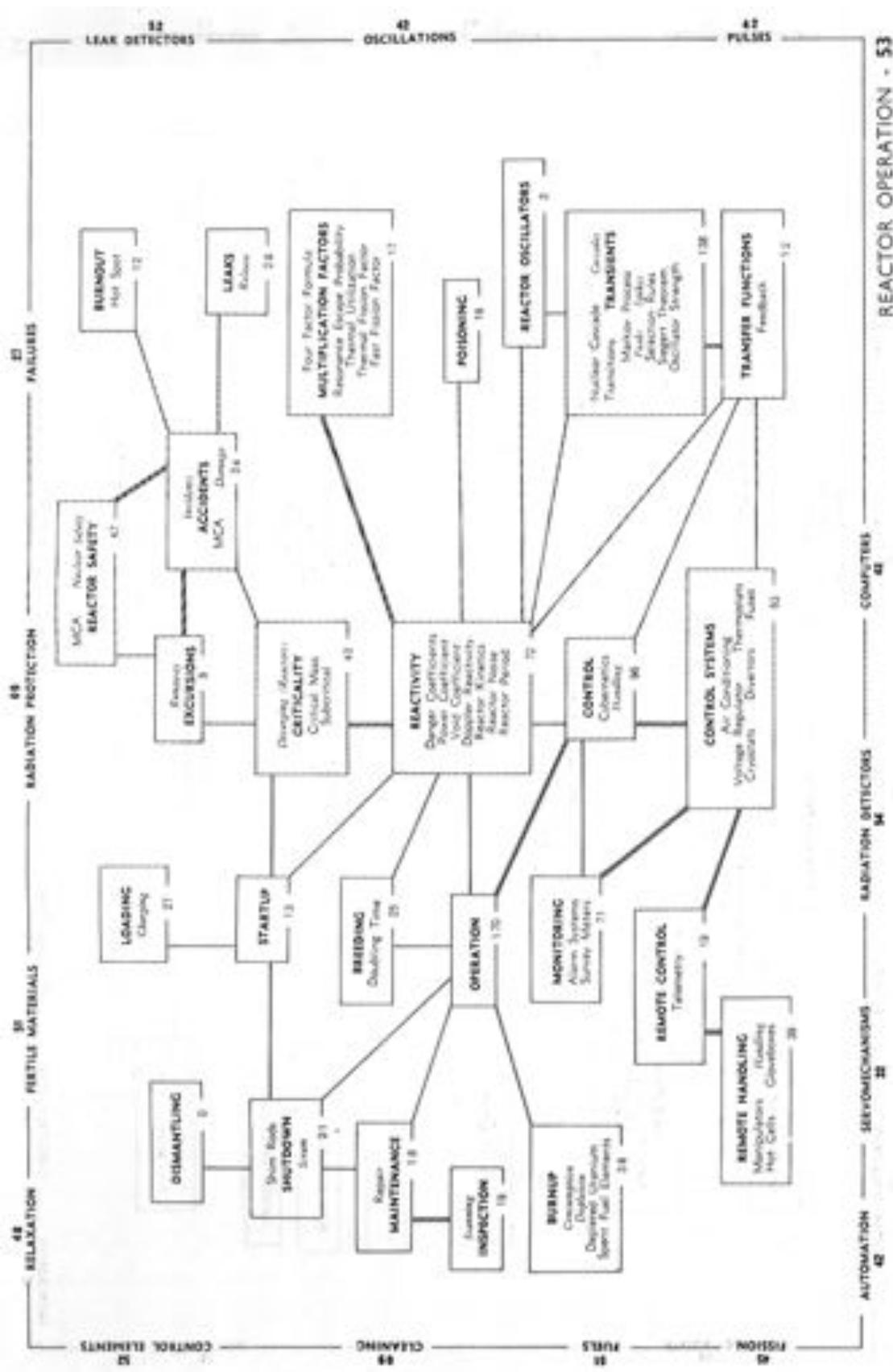


Figura 57: Una tavola del thesaurus EURATOM

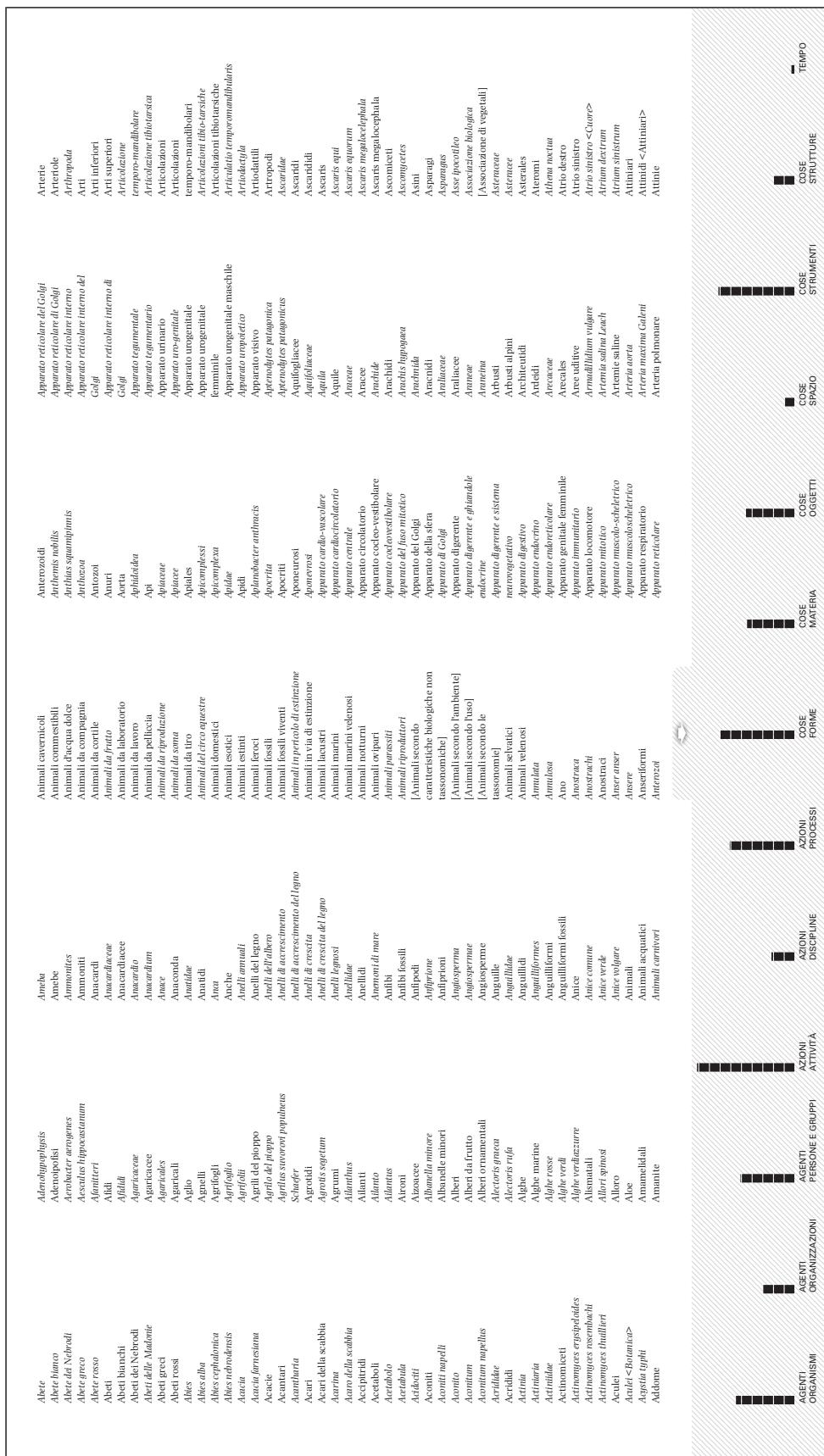


Figura 58: Un esempio di ordine sistematico basato sui termini del Nuovo soggettario.

Generatori elettrici	Tricromatisti	Immigrati pakistani	Scrittori campani	Mummificazione	Gomiti
Impianti elettrici	Sistema decastato probatorio	[Vasi secondo la formola]	Zoarcheologia	Sacrifici cruenti	Gineggiamenti
Aerogeneratori	Sistemi tauri	[Vasi secondo il materiale]	Detenuto/a: a inizio e	Macchine volanti	Prestidrammatica
Insetticidio a provvisoria	Suoli collinari argillosi	[Vasi secondo la civiltà, il	[Scrittori secondo l'attività, il	Volo unesco	Matturato
Spirogyra	Suoli collinari	periodo storico, lo stile]	comportamento]	Luoghi paratarchichi	Materia solida condensata
Antonimia	Suoli incolti	Detergenti	Archaeologia	Volo meccanico	Poli gamma
Articoli < giornalismo>	Suoli incolti	Scrittori egiziani	Letteratura religiosa	Forze armate multinazionali	Tradizioni greche
Crocco	Suoli e aceri	Scrittori indiani	Letteratura	Soluzioni di azoto	Mortalità maschile
Grafitari	Suoli compatti	Scrittori indiani	Confisca (Diritto penale)	Malattie citomegalica	Mortalità femminile
Grafitisti	Suoli bonificati	Scrittori indonesiani	[Malattie degli occhi]	Infeczione	Incarenzione
Writers	Cooperativa tra le due grandi	Scrittori dominicani	Malattie del sangue	Cocionali	Cocionali
Stecco di Roma <127>	Vaghi internazionali	Scrittori dominicani	Concezione calvinista	Infibulazione	Infibulazione
Libri di famiglia	Nord attacco ferrosi	Scrittori austriani	Opere storiche	Dipendenza statistica	Dipendenza
Denominazione d'origine	Relazione frautto	Scrittori austriani	Opere religiose	Registri di stato civile	Registri
Architettura capitolina	Ongosinteriorialità	Autori ingrebbini	Opere politiche	Discendenti	Discendenti
Pittura americana	Genitorialità	Scrittori giapponesi	Opere pedagogiche	Prinavera	Prinavera
Schimpansen, primato	Proprietà editoria	Arca	Opere mediche	Fiere	Fiere
Schimanzano, mano	Tumore, epato-cellulare	[Scrittori secondo il genere	Opere materniche	Acido neuroamminico	Acido neuroamminico
Parapassus	Piellamente in dico argentoato	letterario]	Opere giuridiche	Acido neuroamminico	Carte glaciologiche
Bonobo	Preziosi metalli a disco d'argento	Studi longitudinali	Opere di entomologia	Gradi di qualitato	Gradi di qualitato
Fornitieri africani	Attorniamenti assoluti a disco	Capitale circolante netto	Letteratura religiosa	Fattore III	Fattore III
Oryctoperus ater	d'argento	Capitale circolante	Lingua etzja	Trombofistina	Trombofistina
Lucania secca	Attorniamenti assoluti ad uso	Aeronautica militare	Linea di parentela	Trombofistina	Trombofistina
Cremonea, trasformazioni di	argentato	Personale di credita	Collo degli antenati	Protonibassi	Protonibassi
Picard, vanda di	Attorniamenti a disco d'argento	Industria	Progenitori	Trombofistina	Trombofistina
Cubisti	Attorniamenti a disco d'argento	Studi longitudinali	Avanza popolare	Trombofistina	Trombofistina
Ortitori	Attorniamenti assoluti a disco	Capitale circolante netto	Scrittori celebri genovesi	Shuorobinianus	Shuorobinianus
Pettuzzanni	d'argento	Capitale circolante	Ecologia cristiana	Giudice penale di appello	Giudice penale di appello
Momento territoriale	Attorniamenti assoluti ad uso	Aeronautica militare	Cultura francese	Corti d'appello penale	Corti d'appello penale
Accent gratici	argentato	Personale di credita	Mechandising	Ceramiche egiziane	Ceramiche egiziane
Accenti	Attorniamenti assoluti a disco	Industria	Big, modello di	Aberrato/traumatico	Aberrato/traumatico
Circostanze	Attorniamenti a disco d'argento	Stessi	Catge, teore di	Tutti luci	Tutti luci
Spettacoli di "Penissonsone"	Attorniamenti a disco d'argento	Extanterritorialità	Imprenditoria	Tutti luci	Tutti luci
Archimedie, spalle di	Arte prestaria	Serietà	Costi indiretti	Semenzadi condanna	Semenzadi condanna
[strumenti per misure dello	Malattia di chetococco	Istituzioni pubbliche di	Costi indiretti	CTG.	CTG.
spazio]	Cisti idatidea	assistenza e beneficenza	Ex-detenuti	Tococardiografia	Tococardiografia
Igiene navale	Cisti idatidea	Muscedia	Ex-detenuti	Cardiografia	Cardiografia
Medicina militare	Piatura siciliana	Moscedi	Sociale e Stato	Coprolito	Coprolito
Igiene militare	Cisti da chetococco	Cleistidi	Sociale e Stato	Irriugatori	Irriugatori
Accenti gratici	Stenocisti	Ortofagia	Prediche	Musci napolitani	Musci napolitani
Circostanze	Subellito	Fibrosi	Opere in prosa	Canti reggiani	Canti reggiani
Spettacoli di "Penissonsone"	Fibre di gastera	Ortofagia	Medagli del bocca p	Funk	Funk
Archimedie, spalle di	Spastrum lanceum	Ottalina infettiva	Radici quadre	Acidiane	Acidiane
[strumenti per misure dei	Gastrula di Spagna	Occhio rosa	Crotalus atrox	Scarpe per uomini	Scarpe per uomini
fenomeni elettrici e magnetici]	Gastrula odorosa	Ortofagia conglutinante infettiva	Immunità <diritto	Scarpe da uomo	Scarpe da uomo
Istrumenti di misurazione	Gastrula comune	botuina	Internazionale	Scarpe maschili	Scarpe maschili
della luce, della temperatura]	Cistiflora tessili	Tumori di Erdheim	Immunità di diritto internazionale	Acciai rapidi al tungsteno	Acciai rapidi al tungsteno
tempo]	venezili >	Adamanitum	Ovolo	Costumi da bagno	Costumi da bagno
[strumenti per misure dello	Bilance	Ovolo	Pendolo gravimetrico	Acciai fusi	Acciai fusi
spazio]	Istrumenti per misure della	Pendolo gravimetrico	D'Allemberg, paradosso di	Feste per l'infanzia	Feste per l'infanzia
Igiene navale	gravidità	Istrumenti per misure della	Petizione <diritto	Popolare rurale	Popolare rurale
Medicina militare	gravidità	gravidità	Immunità <diritto	Imbalsamatura	Imbalsamatura
Igiene militare	Metabolismo vegetale	Istrucciflora	Internazionale		
Storia ecclastica	Plankton fossile	Istruttatura	Conceristi musicali		
Catrami	Tessuto sottocutaneo	Echinococchi <bestioli>	Paligia		
Derivati dei catrami	Foruncoli	Echinococchi <bestioli>	Paludaria		
Comunicazioni marittime	Antracite	Echinococcosi	Parasiti		
Servizi marittimi	Istruttori	Echinococcosi	Popoli pescatori		
Distomatosi	Malattia idiotica	Echinococcosi	Popoli cacciatori		
Adenoviridie	Atresia	Echinococcosi	Cacciatori-raccolitori		
Anatomia veterinaria topografica	Certificato prenurionale	Echinococcosi	Folkloristi		
Iea	Certificato medico prenurionale	Echinococcosi	Feste pubbliche		
Anatomia topografica veteri-	Tubercolosi genitoriantraria	Echinococcosi	Ricerca della maternità		
naria	Tubercolosi genitoriantraria	Echinococcosi	Verenze canadensesche		
Trichomonas	Tubercolosi genitoriantraria	Echinococcosi	Trasferimenti di popolazione		

Figura 59: L'ordine cronologico è utile per ritrovare gli ultimi elementi inseriti nel thesaurus.

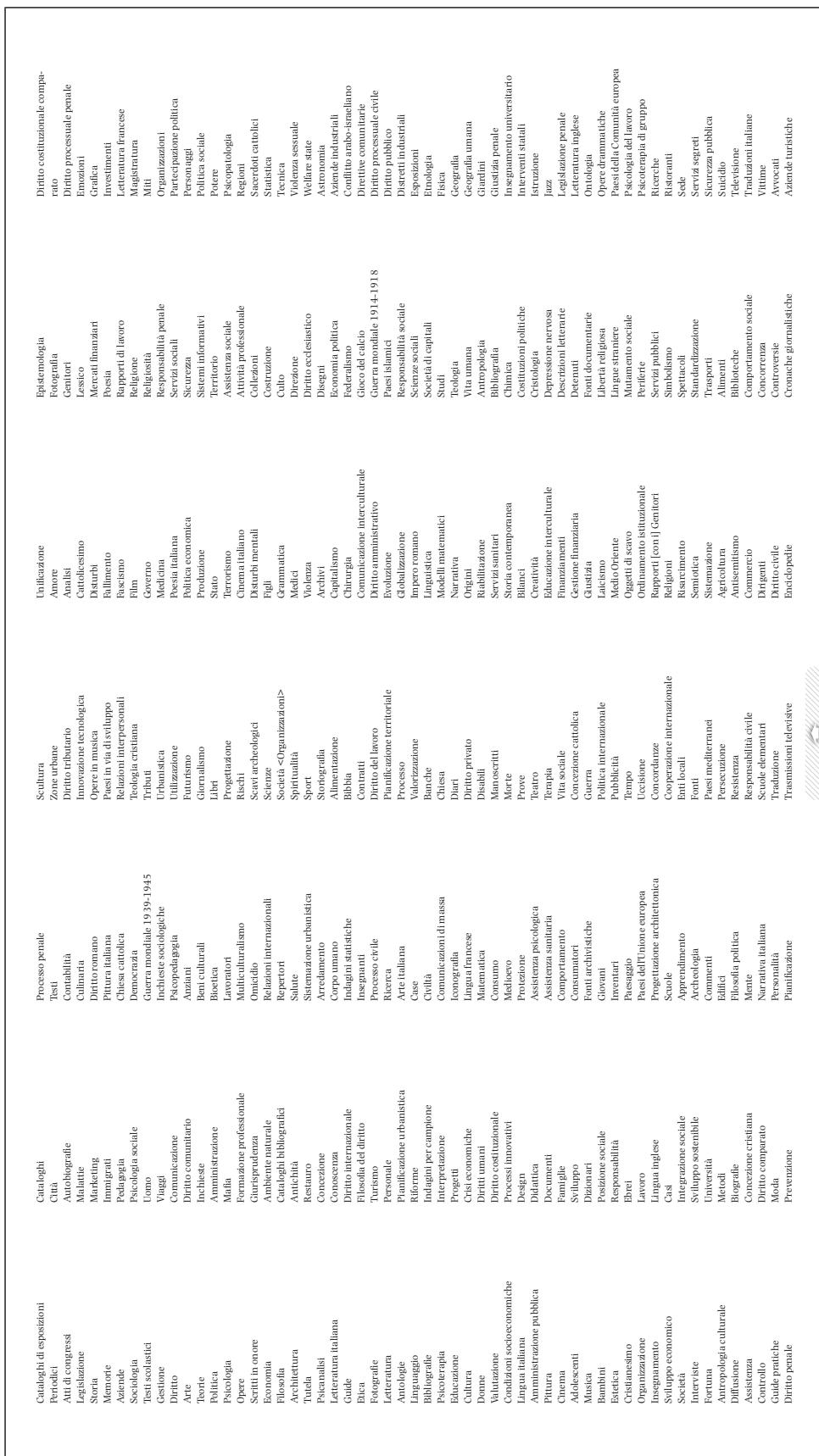


Figura 60: L'ordine quantitativo permette di scoprire i termini più utilizzati.

Il Sistema Multi-Thesaurus

Il motivo per cui le interfacce per la gestione di thesauri sono mono-thesaurus deriva dal fatto che nei contesti lavorativi si ha spesso avuto l'esigenza di usare un unico thesaurus. Negli ultimi anni qualcosa è cambiato, si è investito molto nell'integrazione di sistemi di indicizzazione e anche nell'integrazione di thesauri per poter eseguire ricerche distribuite su diversi sistemi informativi. Una delle tecniche più utilizzate per l'integrazione di vocabolari controllati è il *mapping*, attraverso cui si mettono in relazione termini appartenenti a diversi thesauri così da poter interrogare diverse sorgenti di dati [14]. Il *mapping* si avvale di relazioni extra-thesaurus divise in relazioni di equivalenza esatta, parziale, generica e specifica. Le prime due appartengono alle relazioni di equivalenza, usate per indicare casi di sinonimia completa o parziale. Le seconde due relazioni sono di tipo gerarchico e funzionano come le relazioni BT e NT.

Per gestire due thesauri introduciamo l'idea di suddividere l'interfaccia in due *viste* come illustrato nella figura 64. Nella parte sinistra apparirà un primo thesaurus, nella destra un secondo thesaurus. Ognuna di queste parti sarà indipendente, per permettere all'utente di navigare l'una o l'altra a seconda della necessità.

Pur visualizzando due thesauri contemporaneamente l'attenzione dell'utente sarà posta solo su uno di loro, sarà questo il thesaurus su cui l'utente potrà operare. Il thesaurus messo a fuoco ha due caratteristiche visive che lo distinguono da quello non selezionato. Per le aree è stato introdotto un titolo, che corrisponde al nome del thesaurus. Se il thesaurus è selezionato la scritta appare in grassetto e i termini sono racchiusi da un contorno che li lega al titolo, altrimenti la scritta è in corsivo e i termini senza contorno (vedi fig. 64).

Per esplicitare le relazioni tra la parte sinistra e la parte destra ci si avvale di una visualizzazione grafica. Le relazioni, che collegano i termini delle due viste, saranno rappresentate da una linea leggera in modo da non compromettere la leggibilità. In questo modo si potrà avere un colpo d'occhio complessivo delle relazioni presenti tra i termini (vedi fig. 64).

Uno, due, tre thesauri

Dividere lo schermo per lavorare sui thesauri permette di adottare configurazioni diverse a seconda delle necessità. Sarà infatti possibile combinare thesauri e aree di visualizzazione per gestire sistemi mono-thesaurus o multi-thesaurus. Le combinazioni che andremo a vedere sono le seguenti:

- Un thesaurus, una visualizzazione
- Un thesaurus, due visualizzazioni
- Due thesauri, due visualizzazioni
- Tre visualizzazioni

Un thesaurus, una visualizzazione

Questa configurazione è standard, lo schermo non viene suddiviso e esiste una sola vista che viene dedicata completamente alla visualizzazione di un thesaurus. Questo è il sistema più diffuso nelle interfacce che si trovano sul mercato.

Un thesaurus, due visualizzazioni

Lo schermo è condiviso da due visualizzazioni dello stesso thesaurus. Essendo le aree indipendenti tra loro, l'utente può navigare come se avesse due interfacce separate a disposizione. Lo scopo di questo tipo di visualizzazione è quello di gestire le relazioni all'interno di un singolo thesaurus. Per esempio, una volta trovato il termine da relazionare nella prima area, l'utente può spostarsi liberamente nella seconda area per cercare il secondo termine da collegare. Un volta trovato, sarà in grado di creare la relazione e immediatamente vederla esplicitata nella linea che collega i due elementi.

Due thesauri, due visualizzazioni

In questo caso abbiamo due visualizzazioni, un thesaurus nell'area sinistra e un thesaurus nell'area destra. Questa situazione è pensata per gestire le relazioni extra-thesaurus.

Tre visualizzazioni

Questa configurazione aggiunge una nuova area alle due già esistenti. L'utilità di questo tipo di visualizzazione sta nella possibilità di navigare tre thesaurus contemporaneamente come si vede nella figura 65. Un esempio concreto è portato dal progetto SIRS discusso a pagina 23. In questo caso abbiamo tre vocabolari controllati – Actions, EUROVOC e Dewey Decimal Classification – che integrano tre diversi archivi elettronici attraverso due *mapping* [13]. Actions e EUROVOC, EUROVOC e Dewey sono le coppie tra cui si applica la tecnica del *mapping*. Grazie a tre visualizzazioni e tre thesauri coesistenti, sarà possibile vedere le relazioni extra-thesaurus in modo da esplicitarle graficamente. In questo modo si potranno vedere le relazioni che partono dal primo e arrivano al terzo indice passando dal secondo. Nel caso di SIRS la visualizzazione grafica è particolarmente utile perché riproduce visivamente il *mapping* usato per integrare i thesauri (vedi p. 23) [50].

Navigare le relazioni

La divisione dell'interfaccia introduce un metodo di navigazione: sarà possibile consultare i termini percorrendo le relazioni. Mettiamo che l'utente stia navigando sulla parte sinistra e arrivi a studiare le relazioni appartenenti al termine *Narcolessia*. Intenzionato a vedere come il termine viene strutturato all'interno del thesaurus, cliccando su una relazione – ad esempio RT *Sonno* – vedrà comparire *Sonno* nella parte destra dell'interfaccia esplicitando così la relazione. Successivamente potrà selezionare RT *Allucinazioni*, ripopolando così l'area destra dell'interfaccia in base alla nuova relazione selezionata.

Viceversa, cliccando sulla parte destra dell'interfaccia si ripopolerà la sinistra. Questo tipo di navigazione può dare vita a un cambio di contesto che verte prima su un'area e quindi sull'altra per poi ricominciare.

Questo metodo di navigazione può essere applicato anche alla configurazione basata su tre aree. In questo caso la navigazione delle relazioni in successione rimbalzerà da sponda a sponda seguendo il *pattern*: sinistra, centro, destra, centro, sinistra per poi ricominciare.

Lavorare con le relazioni

La gestione delle relazioni in un sistema multi-thesaurus è importante. Vediamo come eseguire la cancellazione e la creazione di una relazione.

Cancellare una relazione

Soffermandosi col puntatore su una relazione, dopo qualche secondo apparirà leggermente defilato un bottone con una x (vedi fig. 61). Premendolo, la relazione sottesa verrà cancellata e eliminata dal database. L'azione si annullata premendo ctrl e z contemporaneamente.

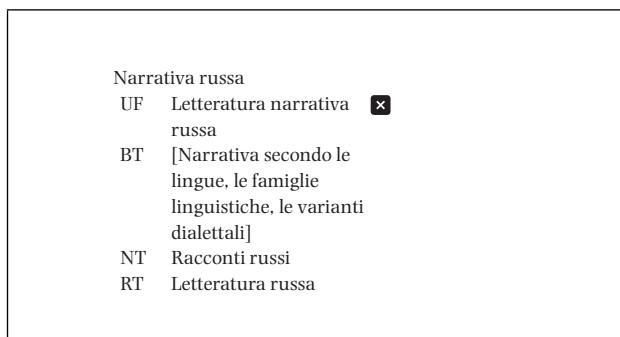


Figura 61: Le relazioni si cancellano usando il bottone che appare a fianco.

Creare una relazione

Per creare una relazione è necessario che siano selezionati due termini, visualizzati in due aree diverse. Una volta selezionati e quindi visibili in grassetto, premendo il tasto destro del mouse su uno dei due termini, sarà possibile accedere ad un menù contestuale con la lista delle relazioni. Selezionandone una dalla lista, verrà creata la relazione del tipo prescelto fra i due termini.



Figura 62: A un termine possono essere aggiunte relazioni attraverso il menù contestuale attivabile con il tasto destro del mouse.

La lista delle relazioni è variabile a seconda che i termini selezionati appartengano allo stesso thesaurus o a due thesauri diversi. Infatti relazioni inter-thesaurus e extra-thesaurus sono diverse tra loro.

Appena creata la relazione verrà tracciata la linea che unisce i due termini, questa visualizzazione funziona come *feedback* permettendo all'utente di capire che la relazione è stata effettivamente creata. La conclusione dell'operazione sarà esplicitata non solo attraverso la tracciatura della linea, ma anche attraverso l'inserimento tipografico della relazione in corrispondenza dei termini interessati.

Se cancellare e creare relazioni su un computer desktop è stato descritto finora utilizzando il mouse, attraverso gli schermi *touchscreen* la stessa operazione si esegue tenendo posato un dito per qualche istante negli stessi punti interessati dal mouse negli esempi precedenti. Il flusso di azioni che ne seguirà sarà identico a quello già illustrato.

Tasti rapidi

La navigazione può essere più veloce utilizzando combinazioni di tasti abbinate finora ad ogni funzionalità introdotta. Lo spostamento di fuoco da un'area all'altra si potrà eseguire con l'uso del tasto *tab*. Premendolo, il fuoco si sposterà all'area successiva. Una volta giunti all'area più a destra dello schermo, la selezione ricomincerà dalla prima area.

Aprire e chiudere le aree

La gestione delle aree si trova nelle etichette dei titoli delle rispettive aree (fig. 63). Cliccando su una di esse si apre un menù contestuale con tre simboli legati a funzionalità differenti:

- Cambio di thesaurus
- Chiusura dell'area
- Apertura di una nuova area

Il cambio di thesaurus è indicato da un triangolo rovesciato come si usa comunemente nei menù a tendina. Il comportamento innescato è lo stesso, si apre una lista di thesaurus dalla quale si può selezionare il nuovo thesaurus da visualizzare nell'area.

La chiusura dell'area si attiva cliccando sul segno x. La chiusura è disponibile esclusivamente se esiste più di un'area, il che significa che se c'è una sola area di visualizzazione essa non può essere chiusa.

La terza funzione apre una nuova area all'immediata destra della visualizzazione da cui la funzione viene chiamata. Per esempio se ci sono due aree e si crea una nuova area dalla prima, essa verrà creata esattamente tra le due aree esistenti. Al contrario se si crea dalla seconda, essa verrà aggiunta come terza area.

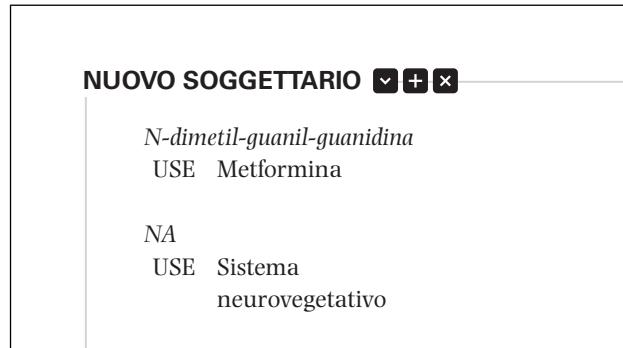


Figura 63: Le aree si gestiscono attraverso una pulsantiera che compare di fianco ai titoli di ognuna delle aree.

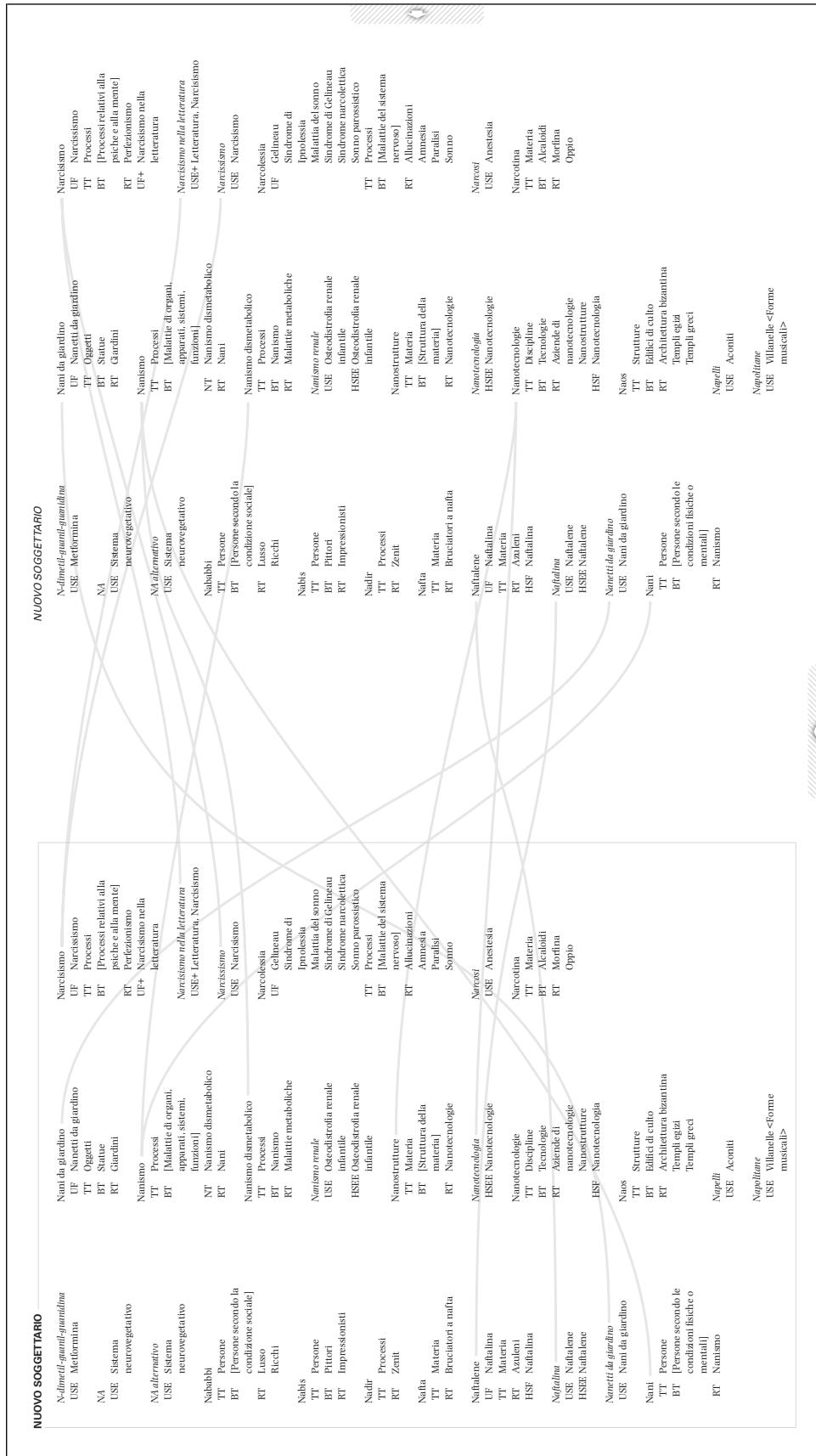


Figura 64: Lo schermo è diviso in due parti indipendenti, il contorno indica che la parte selezionata è quella sinistra. Le linee trasparenti rappresentano le relazioni da una parte all'altra.

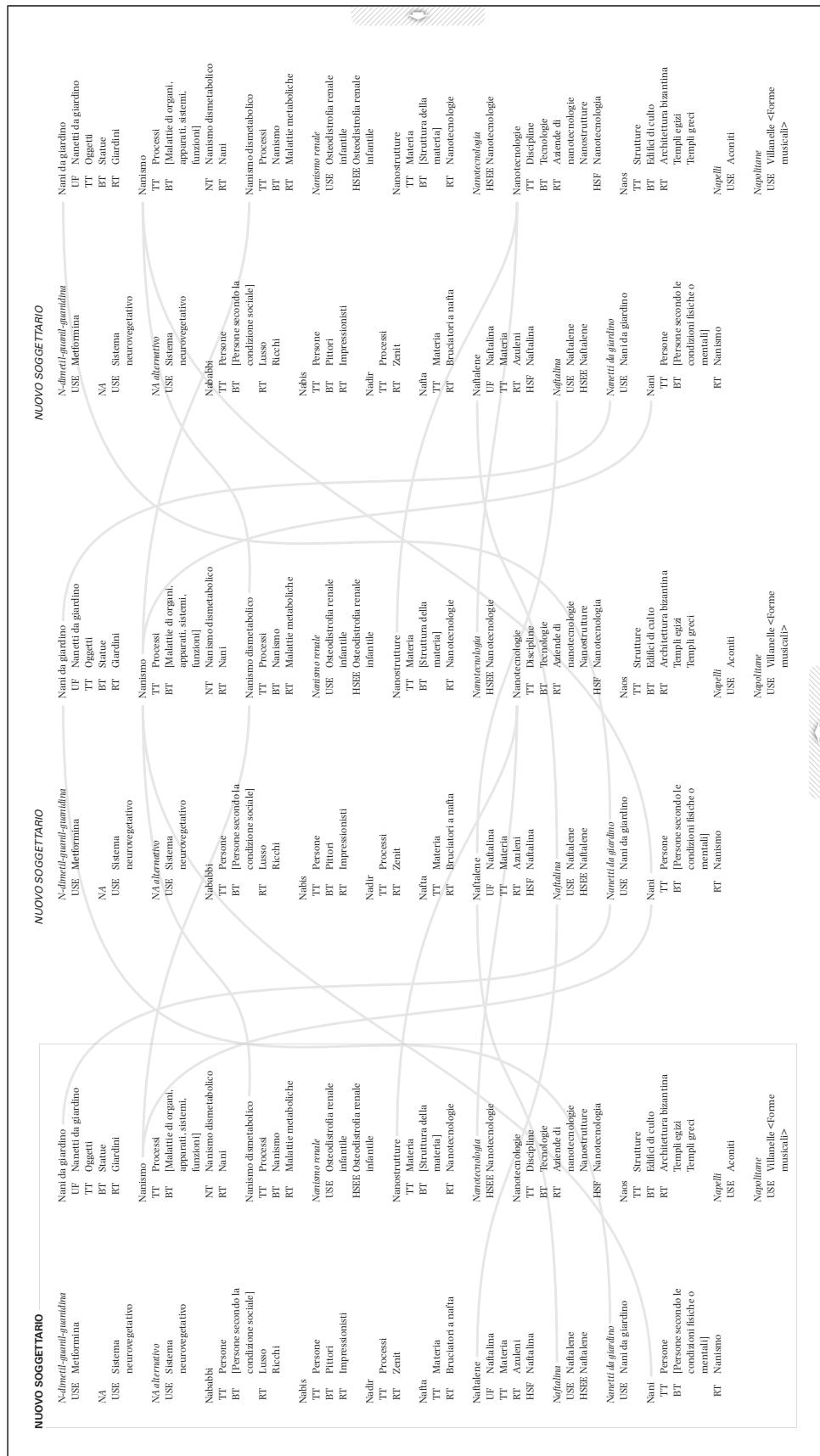


Figura 65: Lo schermo è diviso questa volta in tre parti indipendenti, l'area selezionata è quella sinistra. Le linee che rappresentano le relazioni sono state marcate maggiormente per questioni di leggibilità sulla carta stampata.

Lo Spazio Personale

Ogni utente ha un uso personale dell’interfaccia. Come ognuno di noi ha la propria personalità nel mondo reale, così essa si riflette sui nostri comportamenti nell’uso del computer. Se quando si sale in un’autostrada sedile e specchietti vengono sistemati, allo stesso modo l’interfaccia ha bisogno di personalizzazioni capaci di rendere più confortevole il lavoro.

Abbiamo già visto come nell’interfaccia l’utente sia in grado di modificare la dimensione dei caratteri a proprio piacimento (vedi p. 52, oppure come sia in grado di gestire il numero di aree (vedi p. 89), tuttavia questi sono solo alcuni accorgimenti per migliorare l’usabilità personale dell’interfaccia. Oltre a questo riteniamo fondamentali altri due elementi che riguardano lo spazio personale: il *percorso di navigazione* e i *segnalibri*.

Muoversi nel percorso di navigazione

Una sessione nella nostra interfaccia per la gestione di thesauri si caratterizza da spostamenti di selezione da un termine all’altro e da cambi di configurazione delle aree a seconda di quanti spazi si utilizzano per sfogliare i thesauri. Questi due elementi sono le tracce che vengono registrate dal sistema per risalire il percorso di navigazione. Attraverso il tasto *backspace* sarà possibile andare indietro, ripercorrendo i passi eseguiti su termini e aree. Diversamente potrà andare in avanti premendo la combinazione *ctrl* e *backspace*. Inoltre l’utente, dopo essere tornato indietro ad un punto qualsiasi del percorso, sarà in grado di modificarlo col semplice gesto di una nuova selezione o di una nuova area. Si andrà così a creare una nuova navigazione che andrà a cancellare la vecchia ramificazione del percorso così come avviene nei *browsers* moderni.

Segnalibri

Un’altro standard delle interfacce è rappresentato dal segnalibro o *bookmark*. Se nelle versioni cartacee dei thesauri si usava mettere un segnalibro di cartone per ricordarsi la pagina, l’interfaccia digitale mette a disposizione la stessa funzionalità con una maggiore precisione. Infatti i segnalibri digitali indicano con precisione un singolo termine.

Trascinando un termine nello spazio destinato ai segnalibri, sarà possibile memorizzarlo. Aggiungendo un termine alla volta nello spazio dedicato, si arriverà a formare una lista di termini preferiti. La lista sarà ordinata alfabeticamente o cronologicamente così da permettere all’utente un veloce reperimento con un semplice click.

La Vita di un Termine

Il thesaurus cambia nel tempo così come gli elementi che lo costituiscono. I termini infatti mutano nella loro forma con il mantenimento del thesaurus: nuovi termini vengono aggiunti, modificati nei metadati e nelle relazioni, cancellati per mantenere aggiornato il vocabolario controllato.

I cambiamenti che caratterizzano la vita di un termine vengono attivati dell’utente. Egli, in quanto responsabile del mantenimento del thesaurus, ha il compito di operare sulla struttura, modificandola secondo le esigenze di indicizzazione e reperimento.

L’utente ha quindi bisogno di un’interfaccia che permetta di attivare azioni in maniera veloce. Non crediamo che strumenti specializzati come un’interfaccia per la gestione dei thesauri debbano essere forzatamente intuitivi, pensiamo piuttosto che dopo una fase di apprendimento da parte dell’utente essi debbano essere specialmente veloci, così da permettere il compimento di compiti nel più breve tempo possibile.

Una caratteristica particolare legata alla velocità d’azione è la semplicità. Infatti crediamo che se la progettazione dell’interfaccia è stata realizzata a dovere, la successive introduzioni di azione basate sul *design* iniziale debbano inserirsi con semplicità nella struttura del sistema. Qui di seguito andremo ad analizzare una serie di azioni da costruire *sopra* l’interfaccia progettata fino ad ora. Ci limiteremo a dare delle indicazioni, perché a questo punto si dovrebbe accedere a una fase di progettazione più avanzata per introdurre definitivamente queste funzionalità. Ai fini della tesi è più importante capire come le azioni vengano inserite nella struttura esistente, successivamente – se si vorrà – si potranno portare le azioni qui elencate a una fase di progettazione più avanzata.

Creazione

Ogni area di visualizzazione è caratterizzata da un menù contestuale che si attiva cliccando sul titolo dell'area selezionata. Il menù contiene una serie di funzioni applicabile all'area selezionata, una di queste è la creazione di un termine attivando il tasto **T** (fig. 66).

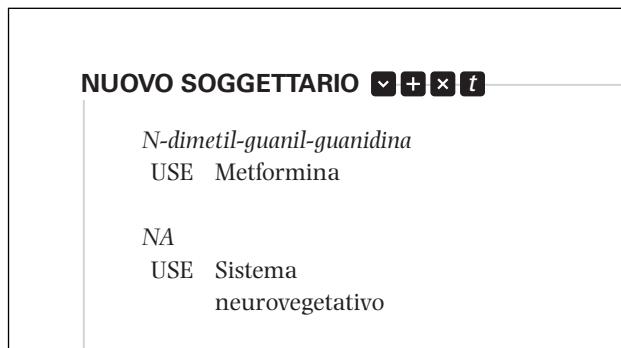


Figura 66: L'ultimo pulsante a destra del titolo è dedicato alla creazione di nuovi termini.

Selezionando la creazione di un termine si accede subito a un campo di inserimento testo posto nella lista. Man mano che si digitano le lettere che compongono il termine, la posizione del campo va via via definendosi in modo sempre più preciso collocandosi all'interno dell'ordine alfabetico corrente. Una volta finito di digitare interamente il nome del termine, la posizione che il campo di testo assume sarà già quella definitiva del termine in quello speciale ordinamento. La possibilità di vedere il contesto prima della creazione del termine permette all'utente di evitare inutili duplicazioni, o semplicemente essere aiutato nella compilazione dai termini attigui in base alla radice che li caratterizza. Solo con la conferma attraverso il tasto *invio* il campo non sarà più modificabile e il termine verrà inserito all'interno della base di dati. Viceversa premendo il tasto *esc* l'operazione viene immediatamente annullata.

Caratterizzazione

Se il processo di creazione può considerarsi semplice, più complessa sarà la cadenza degli eventi che definiranno il dettaglio di un termine. Sotto questo aspetto le relazioni rivestono una grande importanza, abbiamo già visto a pagina 88 come è possibile creare relazioni tra termine e termine. L'altro modo per caratterizzare un termine è attraverso l'uso di metadati. Nel nostro caso i metadati sono stati elencati nel

menù di dettaglio attivabile da un menù a scomparsa sul lato destro dello schermo (vedi p. 65). Per inserire un nuovo metadato basta selezionarlo dal menù contestuale posizionato nella parte alta del pannello, quindi digitare il contenuto e confermarlo. La modifica di un metadato già esistente invece è più semplice, basta fare doppio click sul testo e modificarlo.

Cancellazione

Similmente a quanto succede per le relazioni (vedi fig. 61), soffermandosi con il mouse su un termine qualsiasi, apparirà un bottone che premuto porterà alla cancellazione del termine dopo aver risposto affermativamente a una schermata di conferma.



Figura 67: La cancellazione di un termine si esegue attraverso l'ultimo bottone che appare a fianco del termine quando il mouse si sofferma sull'area.

Ricerca

I thesauri possono contenere decine di migliaia di termini. In questi casi diventa difficile trovare un termine, un'azione chiave nell'uso di sistemi di gestione per vocabolari controllati. Il nostro sistema offre all'utente due possibilità nella ricerca dei termini. La prima è legata alla fruizione dei thesauri cartacei. La seconda è più in linea con i sistemi di ricerca a disposizione delle interfacce digitali.

Ricerca nel contesto

Il primo percorso che un utente può seguire nella ricerca di un termine è ispirato alla fruizione di un thesaurus cartaceo. Da qui abbiamo tratto l'idea di ordinare i termini per pagine, introducendo la barra di naviga-

zione in sostituzione di informazioni contestuali come lo spessore del libro o la divisione in pagine (vedi p. 57).

Immaginiamo un utente che cerca un termine che inizia con la lettera **N** usando la barra di navigazione. Osservando le informazioni disponibili può vedere di quante pagine sia composta la sezione, successivamente sfogliarle per cercare il termine come se si fosse di fronte ad un libro aperto.

Consultare un thesaurus digitale significa utilizzare le funzionalità e le informazioni messe a disposizione dall'interfaccia. Significa scoprire e rendere proprie le diverse modalità con cui si può accedere a un termine perché, così come i libri, anche alle interfacce digitali bisogna dedicare tempo per apprenderne le potenzialità e utilizzarle al meglio.

Questo percorso può essere agevolato da una visualizzazione contestuale. In qualsiasi momento, all'utente è permesso di accedere a una funzionalità di ricerca che si attiva con la combinazione di tasti **ctrl** e **F**. Si potrà così digitare una parola chiave on una stringa di testo nel campo di ricerca apparso. Automaticamente l'interfaccia andrà ad evidenziare nella pagina corrente tutti i termini nei quali la parola chiave appare. Inoltre, con il proposito di ampliare la ricerca a tutto il thesaurus, sarà possibile osservare la presenza di altri termini corrispondenti ai parametri di ricerca nelle altre pagine nella barra di navigazione.

Ricerca fuori dal contesto

Al contrario del primo percorso, la ricerca extra-contestuale ha uno stile più compatto, ma estranea i termini dal contesto. Attraverso la combinazione di tasti **ctrl**, **shift**, **F** si accede a un campo di inserimento dati simile al precedente. Digitando una parola chiave o una stringa di caratteri, verrà composta in tempo reale una pagina ordinata alfabeticamente con i termini che soddisfano la richiesta (fig. 68).

Cliccando sui singoli termini sarà possibile accedere alla pagina del thesaurus dove il termine selezionato è inserito in una lista. Premendo il tasto **backspace** si tornerà alla pagina dei risultati premendo e si potrà scegliere un altro termine.

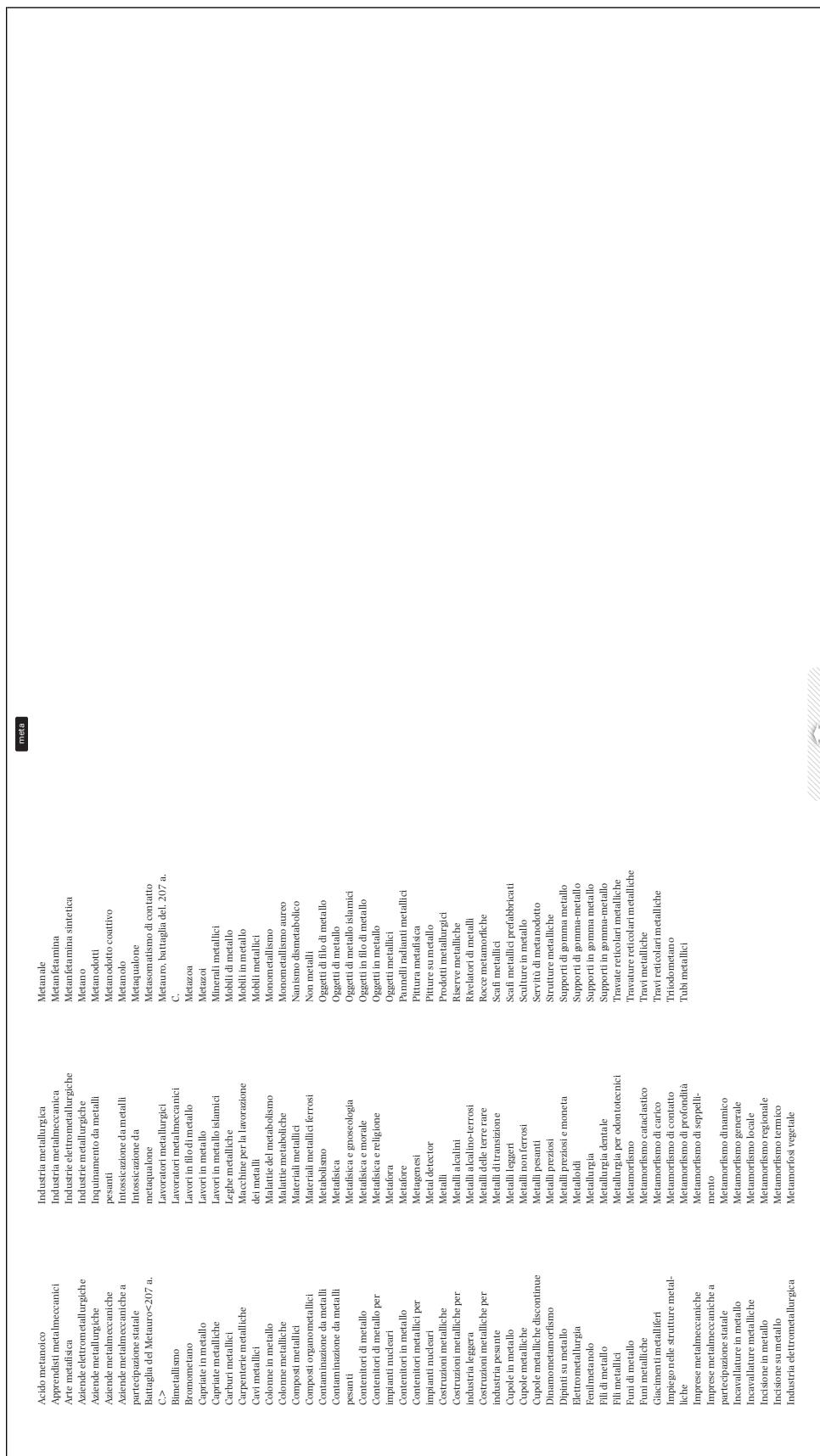


Figura 68: Nell'esempio viene eseguita una ricerca per tutti i termini che contengono la stringa *meta*.



Conclusioni

Dopo tre anni di lavoro presso il Joint Research Centre di Ispra per lo sviluppo di un sistema di information retrieval (SIRS) basato sull'uso di vocabolari controllati e di un'applicazione per il mantenimento dei thesauri (Ponto), la tesi vuole essere la conclusione al ciclo lavorativo svolto presso la Commissione Europea.

L'interfaccia progettata nella tesi è il risultato di un lavoro che ha diminuito la lontananza tra i thesauri digitali e cartacei, dimostrando che, nonostante l'incalzare di ontologie informatiche, il thesaurus non può essere definito un dinosauro ma può ancora contribuire alla realizzazione di sistemi di reperimento e indicizzazione dell'informazione sia manuali che automatici.

L'interfaccia per la gestione di thesauri descritta è l'unica che unisce studi tecnologici, tipografici e biblioteconomici allo stesso tempo. L'idea di riversare l'esperienza del passato in un'interfaccia digitale è stato l'obiettivo perseguito nella tesi.

Il processo di design non si ferma qui. Crediamo che la progettazione debba essere ciclica, debba vivere dell'esperienza dei designer futuri e degli utenti che usano l'applicazione in modo da migliorare il prodotto con il passare del tempo.

Lo studio fatto sin qui è la preparazione alla fase implementativa, in cui l'interfaccia può essere realizzata usando tecnologie attuali come *jQuery* o *Processing* [48] [21]. La tesi ha dimostrato come sia possibile realizzare un'interfaccia specializzata in biblioteconomia da un solido impianto teorico e quindi quanto sia utile la fase di ricerca prima dell'implementazione di un progetto. Se si troveranno le risorse e ci sarà la volontà, l'interfaccia potrà così essere utile nel mantenimento di thesauri come AGROVOC, EUROVOC o il Nuovo soggettario.



Ringraziamenti

Se il percorso di laurea è giunto a una conclusione, lo devo anche alle persone che qui ringrazio.

Carlo Ferigato che sta dalla parte delle biblioteche, Gabriella Pasi la cui pazienza supera la lunghezza di questa tesi e Luca Bernardinello per le racchette da tennis vintage.

La Central Library del JRC, Beppe Merlo e la cerimonia del tè, Daniela e lo yoga, Elena e la famiglia, la signora Giaretta e la pensione, Françoise e i cioccolatini, Una e la British Library, An e Adriaan. Il capo dell'unità Marc Wilikens, Marco e Gianni. Igor e la corsa. Matina e la Biennale di Venezia.

La Biblioteca Nazionale di Firenze, Marta Motta e la famiglia Kraemer, Anna Lucarelli, Maria Grazia, Isolina, Luciana. La Biblioteca Mario Rostoni della LIUC di Castellanza, Piero Cavalieri. La FAO e Margherita Sini. Il British Geological Survey e Rachel Heaven. La Library dell'EC, Carol Bream, Ron Davies.

Gli amici per l'ospitalità.

Gli zii per i dinosauri [19].

I genitori per la fiducia.

La nonna.



Bibliografia

- [1] J. Aitchison, A. Gilchrist, and D. Bawden. *Thesaurus construction and use: a practical manual*. Aslib London, 4th ed. edition, 1997.
- [2] American National Standard Institute - National Information Standards Organization. *ANSI/NISO Z39.19: Guidelines for the construction, format and management of monolingual controlled vocabularies*. NISO Press, Bethesda (USA), 2005.
- [3] W.Y. Arms. *Digital libraries*. The MIT Press, 2001.
- [4] D. Austin. *PRECIS: a manual of concept analysis and subject indexing*. London, 1974.
- [5] R. Baeza-Yates, B. Ribeiro-Neto, et al. *Modern information retrieval*. Addison-Wesley Reading, MA, 1999.
- [6] Biblioteca nazionale centrale di Firenze. Nuovo soggettario online. <http://thes.bncf.firenze.sbn.it/>.
- [7] Biblioteca nazionale centrale di Firenze. *Nuovo soggettario: guida al sistema italiano di indicizzazione per soggetto*. Bibliografica, Milano, 2007.
- [8] R. Bringhurst. *The elements of typographic style*. Hartley & Marks, Point Roberts, WA, 3rd ed. edition, 2004.
- [9] British Standards Institution (BSI). *Structured vocabularies for information retrieval (BS 8723:2005-2008)*. British Standards Institution, London, 2005-2008.
- [10] V. Broughton. *Costruire thesauri*. Editrice Bibliografica, Milano, 2008.
- [11] M.P. Carosella and M. Valenti. *Documentazione e biblioteconomia*. F. Angeli, 1982.
- [12] G. Chierchia and S. McConnell-Ginet. *Meaning and grammar: An introduction to semantics*. The MIT Press, 2000.

- [13] M. Dewey. *Dewey decimal classification and relative index*. OCLC Online Computer Library Center, Dublin, Ohio, ed. 22 edition, 2003.
- [14] M. Doerr. Semantic problems of thesaurus mapping. *Journal of Digital information*, 1(8), 2006.
- [15] European Communities. *Euratom Treaty: Treaty Establishing the European Atomic Energy*. Office for Official Publications of the European Communities, Luxembourg, 1957.
- [16] European Communities. *Euratom Thesaurus: part I, Indexing terms used within EURATOM's nuclear documentation system; part II, Terminology Charts Used in Euratom's Nuclear Documentation System*, volume EUR report 500.e. Information and Documentation Center, Brussels, 1966,1967.
- [17] FAO. Agricultural Information Management Standards: Interoperability, Reusability and Cooperation. www.fao.org/agrovoc/.
- [18] C. Ferigato, G. Merlo, D. Panfili, and D. Rodighiero. Role of thesauri in a scientific organisation. In *Networks of Design: Proceedings of the 2008 Annual International Conference of the Design History Society (UK) University College Falmouth, 3-6 September*, page 301. Universal-Publishers, 2009.
- [19] C. Flammarion. *Il mondo prima della creazione dell'uomo*. Società Editrice Sonzogno, Milano, 1934.
- [20] W.B. Frakes and R. Baeza-Yates. *Information retrieval: Data structures & algorithms*, volume 152. Prentice Hall Englewood Cliffs, NJ, 1992.
- [21] B. Fry and C. Reas. Processing. <http://processing.org/>.
- [22] J.J. Gibson. *The ecological approach to visual perception*. Lawrence Erlbaum, 1986.
- [23] P. Gray and D. Conti. *Psicologia*. Zanichelli, 2004.
- [24] R. Heaven, D. Rodighiero, and M. Sini. Ponto, multi-thesaurus tool. <http://www.assembla.com/wiki/show/ponto/>.
- [25] A. Iacona and E. Paganini. *Filosofia del linguaggio*. R. Cortina, 2003.
- [26] J. Kaiser. *Systematic indexing*. Pitman, 1911.

- [27] F.W. Lancaster. *Vocabulary control for information retrieval*. Information Resources Press, Washington, DC, 1972.
- [28] F.W. Lancaster and A.J. Warner. *Information retrieval today*. Information Resources Press Arlington, VA, USA, 1993.
- [29] Library of Congress. Marc standards. <http://www.loc.gov/marc/>.
- [30] Library of Congress. Z39.50. <http://www.loc.gov/z3950/agency/>.
- [31] Library of Congress. *Subject headings*. Library of Congress, Subject Cataloging Division, Processing Department, Washington, DC, 1975.
- [32] J. Löwgren and E. Stolterman. *Thoughtful interaction design: A design perspective on information technology*. The MIT Press, 2004.
- [33] M. Majoor and S. Morlighem. *José Mendoza y Almeida*. Ypsilon Editeur, Paris, 2010.
- [34] C.D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge University Press New York, NY, USA, 2008.
- [35] L. Manovich. *The language of new media*. The MIT press, 2001.
- [36] D. Marconi. *La Filosofia del linguaggio: da Frege ai giorni nostri*. Utet Libreria, 2005.
- [37] E. McKean, editor. *New Oxford American Dictionary*. Oxford University Press, New York, 2005.
- [38] C.W. Morris. *Foundations of the Theory of Signs*. University of Chicago Press, 1938.
- [39] M. Motta. La struttura del thesaurus: monogerarchie e poligerarchie. *Biblioteche oggi*, 25(6):103–104, 2007.
- [40] M. Motta and D. Rodighiero. Il Thesaurus del Nuovo soggettario interpreta SKOS. *Associazione Italiana per la Documentazione Avanzata (AIDA)*, 2010.
- [41] J. Nielsen. *Hypertext and hypermedia*. Academic Press Professional, Inc. San Diego, CA, USA, 1990.
- [42] OCLC. Webdewey. <http://connexion.oclc.org/>.

- [43] Y. Qiu and H.P. Frei. Concept based query expansion. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 160–169. ACM, 1993.
- [44] S.R. Ranganathan. *The five laws of library science*. Madras Library Association (Madras, India) and Edward Goldston (London, UK), 1931.
- [45] S.R. Ranganathan. *Colon classification: basic classification*. Asia Publishing House, 1963.
- [46] S.R. Ranganathan. *Prolegomena to library classification*. Asia Publishing House, 1967.
- [47] C. Reas and B. Fry. *Processing: a programming handbook for visual designers and artists*. The MIT Press, 2007.
- [48] J. Resig. jQuery. <http://jquery.com/>.
- [49] R. Ridi. *Il mondo dei documenti*. Laterza, Roma-Bari, 2010.
- [50] D. Rodighiero, M. Halkia, and M. Gusmini. Mapping for Multi-source Visualization: Scientific Information Retrieval Service (SIRS). In *Proceedings of the 13th International Conference on Human-Computer Interaction. Part IV: Interacting in Various Application Domains*, page 605. Springer, 2009.
- [51] S.K. Roe and A.R. Thomas. *The thesaurus: review, renaissance, and revision*. Cataloging & Classification Quarterly. Haworth Press, 2004.
- [52] P.M. Roget and B. Kirkpatrick. Roget's Thesaurus, 1911.
- [53] R. Spence. Information Visualization. Harlow, 2001.
- [54] E. Svenonius. *The intellectual foundation of information organization*. The MIT Press, 2000.
- [55] S. Tessaris, C. Gutierrez, R.A. Schmidt, S. Handschuh, T. Eiter, E. Franconi, and M.C. Rousset. *Reasoning Web. Semantic Technologies for Information Systems: 5th International Summer School 2009, Brixen-Bressanone, Italy, August 30-September 4, 2009, Tutorial Lectures*. Springer Berlin Heidelberg, 2009.
- [56] H. Thimbleby. *Press On: Principles of interaction programming*. The MIT Press, 2010.

- [57] C.J. Van Rijsbergen. *The geometry of information retrieval*. Cambridge University Press, Cambridge NY, 2004.
- [58] L.D. Will and S.E. Will. Willpower information. <http://www.willpowerinfo.co.uk/>.
- [59] L. Wittgenstein and G.E.M. Anscombe. *Philosophical investigations: the German text, with a revised English translation*. Wiley-Blackwell, 2001.