Tecniche di ricerca di documenti basate sulla Classificazione Decimale Universale (CDU) in un OPAC multilingue svizzero

intervento al 3' Incontro ISKO Italia-UniMIB Milano, 22 giugno 2007

Jiři Pika

Dr. Jiří Pika, ETH-Bibliothek, ETH-Zurich, Raemistr. 101, CH-8092 Zurich, Switzerland, email: pika@library.ethz.ch, tel. 0041 44 632 6717

Introduzione

La biblioteca dell'ETH (Politecnico di Zurigo) utilizza la CDU da venticinque anni* per classificare i suoi documenti senza che gli utenti ne siano consapevoli.

Per facilitare il dialogo con OPAC, la *query* nell'attuale NEBIS** (ex ETHICS) si basa sulla ricerca con descrittori in tre lingue*** e su altri termini di ricerca corrispondenti, tra cui i sinonimi e le espressioni del gergo delle riviste scientifiche. Un singolo numero CDU, che corrisponde a questi descrittori, li collega con i relativi documenti/titoli, indipendentemente dalla lingua. Così l'utente lavora realmente con la CDU, senza accorgersene.

L'indice per soggetti della biblioteca dell'ETH contiene (al 1.1.2007) 64'000 voci; è realizzato in inglese, francese e tedesco ed è utilizzato da tutte le biblioteche svizzere che fanno parte della rete NEBIS. Tra queste, le principali sono l'ETH-Bibliothek di Zurigo, di lingua tedesca, e la controparte di lingua francese, la Bibliothèque Centrale della EPFL di Losanna, entrambe responsabili della maggior parte delle voci dell'indice per soggetti nella loro rispettiva lingua. L'inglese è la lingua di riferimento.

Ogni singola voce dell'indice per soggetti di NEBIS contiene sia i principali descrittori (in tre lingue), sia parecchi altri termini affini, come sinonimi, quasi-sinonimi, acronimi, nonché le espressioni recentemente coniate dalle pubblicazioni scientifiche – gergo scientifico – che servono ad allargare la portata della voce dell'indice per soggetti. Il descrittore principale è quello che compare nella classificazione di un documento nel catalogo, e rappresenta tutti i termini affini ad esso collegati. Il numero di questi termini può essere da quattro a otto volte superiore alla singola voce dell'indice. Questa struttura è ripetuta manualmente nelle tre lingue: ogni singolo dato è digitato, aggiornato e mantenuto come "registro NEBIS trilingue" (indice per soggetti multilingue di NEBIS) da un team di 12-15 classificatori di entrambe le biblioteche principali, Zurigo e Losanna. Inoltre, più di 80 biblioteche cooperano nella rete NEBIS nelle tre regioni linguistiche della Svizzera. In Germania, un gateway dell'informazione

denominato GERHARD utilizza da 12 anni l'indice sistematico CDU basato sul sistema dell'ETH.

Il team di classificatori della biblioteca dell'ETH amplia la CDU nelle classi e in quei punti in cui questo è logico, critico o di fondamentale importanza.

Il numero di 64'000 voci dell'indice per soggetti corrisponde per caso a quello dell'attuale edizione inglese della CDU. Tuttavia, il relativo contenuto non è necessariamente simile a quello della CDU, ma riflette esigenze locali, quali le acquisizioni locali ed internazionali, la cultura e la geografia del posto e i campi di ricerca di interesse locale. In ogni caso, la classificazione dei documenti scientifici locali ed internazionali attuali degli ultimi 20 anni può essere ritenuta responsabile sia dell'arricchimento che dello sviluppo della configurazione specifica di voci dell'indice per soggetti di NEBIS. Alcuni esempi indicheranno in che modo la CDU viene applicata nel NEBIS di Zurigo.

1. Caratteristiche CDU-CDD, ovvero diversità nell'uso della CDU e della CDD

1.1: CDD

La classificazione decimale Dewey (CDD) si è molto diffusa nelle biblioteche pubbliche e scolastiche in tutto il mondo, prima negli Stati Uniti e poi in Italia, per vari motivi:

- era la prima moderna classificazione, creata nel 1876 per utilizzo nelle biblioteche
- ogni scheda importata dalle banche dati è già fornita della notazione CDD.

Tuttavia, l'uso della CDD per la ricerca non è l'ideale, perché il numero CDD nelle banche dati non contiene la sua espressione verbale. E' quindi illeggibile per non-esperti di CDD.

Recenti sforzi fatti in Germania (U. Reiner 2007a,b) hanno portato i primi frutti, con un programma per computer che consente di retro-tradurre quasi istantaneamente i contenuti della CDD nelle parole-chiave: è sufficiente metterli nel computer e farli tradurre. Ci si può chiedere perché non si può classificare in maniera più trasparente, usando la notazione e la sua espressione verbale insieme sulla scheda già dall' inizio. Sarebbe conveniente non attaccare i suoi componenti assieme: perché usare 5549426 o 554 ' 942 ' 6 se si tratta di 55 in 494 . 26, "geologia dei Grigioni"?!

Inoltre la classificazione CDD ha degli squilibri di ordine geografico e

religioso. Nonostante le frequenti revisioni, il sistema Dewey viene giudicato teoricamente inferiore rispetto ad altri sistemi più moderni che fanno uso dei caratteri alfanumerici (Wikipedia/Dewey).

1.2: CDU

I numeri utilizzati per il sistema di classificazione decimale Dewey hanno costituito la base per il sistema di <u>Classificazione Decimale Universale</u>, sviluppato da <u>Paul Otlet</u> e <u>Henri La Fontaine</u> nel 1895. Questo prevede tra l'altro la combinazione del sistema di base Dewey con una punteggiatura predefinita (virgole, punti, parentesi etc.) per distinguere le singole classi/notazioni (Wikipedia/Dewey).

Come dice Laura Severin (2004): "Otlet e La Fontaine cominciarono dunque ad apportarvi quei cambiamenti che ritenevano funzionali ai loro obiettivi. In particolare si focalizzarono sul principio del "dividi come"; tale principio permetteva di adottare, all'interno di una data classe, un certo numero di suddivisioni standard, in grado di specificare tutti quegli aspetti che potevano essere presenti nei documenti, ma che sarebbe stato impossibile prevedere ed elencare esaustivamente nelle tavole dello schema: si trattava della prima applicazione di un principio di sintesi all'interno di una classificazione gerarchica, principio riconosciuto da Dewey, anche se limitato ad ambiti piuttosto ristretti. Rendendo esplicita l'indicazione di Dewey, i due belgi enuclearono dunque una serie di "suddivisioni comuni" (cioè categorie d'ordine generale applicabili a qualsiasi contesto) nettamente distinte dalle vere e proprie suddivisioni elencate nelle tavole; rese riconoscibili da segni di interpunzione o simboli matematici, le suddivisioni comuni consentivano di esprimere sinteticamente i molteplici aspetti di un documento, dando vita a notazioni più precise ed espressive".

1.3: Conclusione

Mentre Melwil Dewey (CDD) voleva solamente un modo uniforme per sistemare i libri sugli scaffali, la CDU, come figlia legittima della CDD, è stata concepita dall'inizio per un uso universale. E' stata privata del WASP-bias e soprattutto è stata sviluppata per catalogare articoli delle riviste, singoli oggetti d'arte, naturalia etc., cioè per gli scopi documentalistici. Il numero delle classi era sempre superiore a quello della CDD ed essa era quindi considerevolmente più dettagliata.

Tuttora, la CDD serve unicamente per assegnare un posto ad un libro su uno scaffale, salvo in Germania, dove la Deutsche Bibliothek (2006) di Francoforte ha recentemente classificato 50'000 documenti utilizzando la CDD. Si vede quindi che i due concetti iniziali erano per il loro uso

completamente diversi e sorprende il continuo tentativo di confrontare due sistemi, che non sono per nulla confrontabili tra loro, sia per l'intrinseca diversità, che per gli scopi per cui sono stati creati.

2. Diversità strutturale tra CDU e CDD

La CDD, essendo una classificazione enumerativa, ha un numero di classi definito. Le istruzioni per classificare sono uguali per tutte le classi nella CDU (2 volumi, 1424 pp.), mentre cambiano da classe a classe nella CDD (4 volumi, 4193 pp.).

La struttura della CDU è stata sviluppata non soltanto per un uso universale, ma le sue suddivisioni consentono di ottenere in modo molto economico una post-coordinazione molto precisa durante il processo di classificazione, nonché durante la ricerca del documento. Questa struttura consente di combinare più descrittori, esprimendo meglio il contenuto del documento: l'uso delle suddivisioni "luogo", "lingua", "tempo", "razza", "materiale", "forma della pubblicazione" etc. permette di classificare un documento in maniera precisa e facile.

Esempio di suddivisioni comuni:

di luogo, segnalate dalle parentesi: 55 (494) **è la geologia dell'Italia**

di lingua, precedute dal segno =: 55=50 **è un testo di geologia in italiano**

di tempo, indicate dalle virgolette: 55"19" è la geologia del '900

di razza, tra parentesi con il segno =: 55 (=01.378) **è la geologia degli etruschi**

di forma, segnalate da uno 0 seguito dal numero e fra parentesi: 55 (075) **è un libro di testo di geologia**

Inoltre, la suddivisione comune "luogo" nella CDU riporta le zone geografiche dell'Europa in maniera e quantità adeguata per ogni paese, il che non si può dire nel caso della CDD. Ancora più importante è il fatto che la CDU prevede la distinzione delle singole classi: le singole notazioni sono segnalate con simboli matematici (.:/"*+'=), il che indica immediatamente dove inizia e dove finisce un elemento della classificazione del documento.

3. Uso della CDU in NEBIS

Il sistema di NEBIS, utilizzato nell'ambito della Biblioteca dell'ETH di Zurigo, consente l'uso della notazione CDU e dei descrittori, legati da un hyperlink. Nel classificare, appositi descrittori (da 2 a 7 per ogni documento) che esprimono il contenuto del documento vengono assegnati alla scheda OPAC del documento stesso. Il loro ordine rispetta l'importanza degli argomenti trattati. Le suddivisioni comuni, come tempo, luogo etc. vengono collocate verso la fine.

Ogni voce dell'indice per soggetti di NEBIS, cioè ogni classe, contiene un descrittore, vari sinonimi, termini associati, come lo slang o il gergo scientifico, e acronimi. Con questo arricchimento l'utente viene aiutato nella sua ricerca, che può essere svolta in una delle tre lingue.

4. Classificare in CDU e in CDD e reperibilità di un documento: un esempio

Esaminiamo come si classifica e come si ritrova un documento nella CDU e nella CDD.

4.1. Esempio: geologia dei Grigioni

(notazione: 55 = geologia, 494.26 = Grisons (Grigioni))

1. **CDU** 55 (494.26)

2. CDD 5549426 oppure 554'942'6

3. CDU in NEBIS 55 Geologie(Ger) Geology(Eng) Geologie (Fre)

AND (combinato con)

(494.26) Graubuenden (Ger) Grisons (Eng) Grisons (Fre)

4.2. Reperibilità dell'argomento

Se si esegue una ricerca:

I. in un sistema decimale con la notazione 494.26 = Grisons (Grigioni) e 55 = Geology si troveranno i documenti secondo gli esempi precedenti 1 e 3, II. in modo verbale, si troveranno i documenti secondo l'esempio 3,

III. in modo gerarchico decimale, si troveranno i documenti secondo l'esempio 3.

4.3. Sommario della reperibilità di un documento secondo il metodo di ricerca

Successo della ricerca:

Metodo di ricerca	Valutazione
con CDD (1) non si trova in alcun caso	0 punti
in CDU (2) si trova in due casi,	2 punti
in CDU – modo ETH-BIB (3) si trova in tutti tre i casi	3 punti

Si comprende quindi l'utilità della notazione e di conseguenza quale classificazione sia più adequata.

5. Esempio di ricerca di una carta "nei Grigioni" (zona montuosa nel sudest della Svizzera)

La ricerca nell'indice per soggetti di NEBIS sfrutta il vantaggio della sua struttura basata sulla CDU. Il termine "Grigioni" porta il numero CDU (494.26). Tutte le sue suddivisioni riportano la radice più le estensioni numeriche (494*26+numeri), il che estende il numero CDU in modo decimale. Queste possono essere rappresentate e ricercate con (494*26*) (* = query troncata).

Questo fatto può essere sfruttato per la ricerca avanzata, per esempio per tutte le carte geologiche dei Grigioni (Fig. 5.1). La carta geologica ha il numero CDU 912::55, se questo é combinato con (494*26), che corrisponde ai Grigioni, la ricerca trova rispettivamente 8619 e 1131 documenti. Il loro "cross" risulta in 40 titoli, il che significa che ci sono 40 carte geologiche dei Grigioni.

Le località piccole e "caratteristiche", come Soglio, Sils o Silvaplana, non sono state trovate, sono però contenute nella sottoclasse Grigioni - come relative suddivisioni. Queste possono quindi essere cercate con la *query* troncata (494*26 *) (= Grigioni con tutte le suddivisioni).

Iniziamo così da capo la ricerca con 912::55 combinato con (494*26*). Il risultato sarà costituito da 127 titoli (Fig. 5.2). Questo numero include tutte le suddivisioni, anche quelle che appartengono direttamente alla classe Grigioni, già trovata in precedenza.

Le due quantità trovate possono essere facilmente identificate nella lista del processo di ricerca (Fig. 5.3) che aiuta a distinguerle. La funzione "combining"

(operatore booleano) separa 40 dei 127 titoli (Fig. 5.4). Il risultato previsto dovrebbe quindi essere 87. Gli 87 titoli trovati (Fig. 5.5) confermano la stima. Un attento controllo di questa selezione rivela che si tratta effettivamente di località minori nei Grigioni e di carte geologiche ad esse associate (Fig. 5.6 - 5.7).

Queste carte costituiscono una buona base per un'interessante escursione, ad esempio nella "finestra geologica di Engadina" che ci riporta indietro alla storia geologica della Terra e che costituisce oggi un paesaggio di singolare serenità e bellezza, apprezzata ogni anno da migliaia bibliotecari/non-bibliotecari e geologi/non-geologi.

Fig. 5.1 – 5.9 La ricerca nell'indice per soggetti di NEBIS sfrutta la sua struttura basata sulla CDU (pagine seguenti).

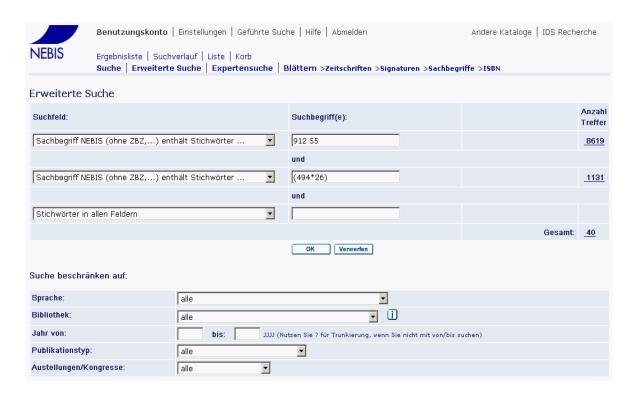


Fig. 5.1 La ricerca combinata della notazione 912 55 che corrisponde ad una carta geologica e della notazione (494*26) che corrisponde ai Grigioni risulta in 40 titoli. N.B.: la notazione CDU deve essere formulata nel modo sopra indicato.

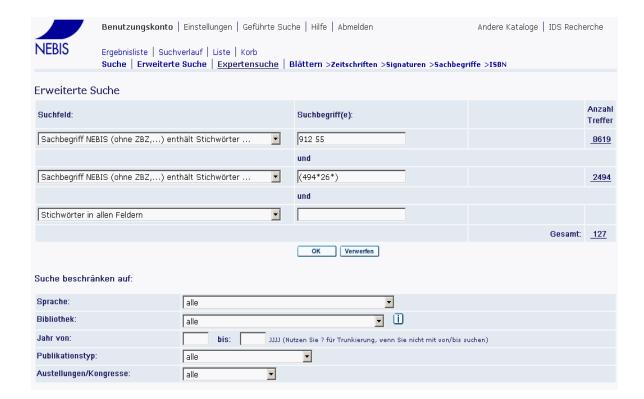


Fig. 5.2 La ricerca combinata della notazione 912 55 che corrisponde ad una carta geologica e della notazione (494*26*) che corrisponde ai Grigioni con tutte le suddivisioni risulta in 127 titoli. N.B.: la notazione CDU deve essere formulata nel modo sopra indicato, utilizzando l'asterisco (query troncata).



Fig. 5.3 Lista di progresso della ricerca con 127 e 40 titoli. (Suchverlauf = lista di progresso della ricerca)



Fig. 5.4 Il nostro scopo è ottenere il secondo set (127), meno il primo (40) (operazione booleana), cioè tutti i titoli riguardanti le suddivisioni dei Grigioni (127, tutte le unità geografiche minori), tranne la classe stessa Grigioni (40).



Fig. 5.5 Risultato: 87 titoli - tutte le suddivisioni dei Grigioni, cioè le loro unità geografiche minori esclusa la classe stessa dei Grigioni.

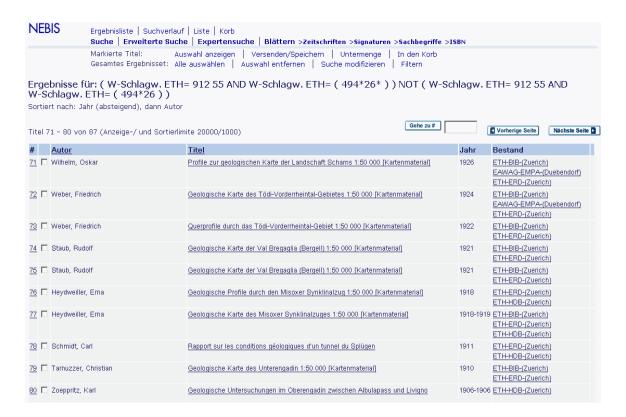


Fig. 5.6 Schermata di titoli delle carte geologiche delle aree geografiche minori dei Grigioni.

	Benutzungskonto Einstellungen Geführte Suche Hilfe Abmelden And
NEBIS	Ergebnisliste Suchverlauf Liste Korb
	Suche Erweiterte Suche Expertensuche Blättern >Zeitschriften >Signaturen >Sachbegriffe >ISBN
Vollanzeige d	es Titels
Anzeigeformat: St	andardformat Katalogkarte Zitierformat MARC
Titel	Geologische Karte der Val Bregaglia (Bergell) 1:50 000 : Ausgabe A: mit Gebirgszeichnung / Rudolf Staub; Schweiz, Geol. Kommission [Kartenmaterial]
Kartendaten	1:50 000 (E 09°29-E 09°45/N 46°25-N 46°15)
Impressum	Winterthur: Kartographia, 1921
Umfang	1 Karte: mehrfarbig; 40 x 44 cm
Reihe	(Geologische Karte der Schweiz, Spezialkarten ; Nr. 90, 1921)
Gehe zu	Geologische Karte der Schweiz, Spezialkarten
Notiz	Topogr. Grundlage: Topogr. Atlas der Schweiz, Blätter No. 510, 513bis, 520, 521, 523, 523bis
Gesamtbestand	Alle Exemplare
Bibliothek	ETH-BIB (Zuerich) K P 101017: 90 KS Kartensammlung 🗻
Bibliothek	ETH-BIB (Zuerich) K P 610025; 90 KS Kartensammlung
Bibliothek	ETH-ERD (Zuerich) I KG 204090 []
Bibliothek	ETH-ERD (Zuerich) I KG 204090 ex.A
Bibliothek	ETH-ERD (Zuerich) I KG 204 090 ex.B
Bibliothek	ETH-ERD (Zuerich) KG 204 090 ex.C
Bibliothek	ETH-ERD (Zuerich) I KG 204 090 ex.D []
Bibliothek	ETH-ERD (Zuerich) L KG 602 090 []
Bibliothek	ETH-ERD (Zuerich) I KG 602 090 ex.A 🗻
Bibliothek	ETH-ERD (Zuerich) I KG 602 090 ex.B
Bibliothek	ETH-ERD (Zuerich) I KG 602 090 ex.D
Bibliothek	ETH-ERD (Zuerich) KG 602 090 ex.C Raritaeten
	——————————————————————————————————————
Sacherschliessun	g GEOLOGISCHE KARTEN: 912%65
	BERGELL (KANTON GRAUBUENDEN): (494/261.72)
Autor/-in	Staub, Rudolf, Geologe, Schweiz, 1890 - 1961, ger

Fig. 5.7 Dettaglio di uno dei risultati: carta geologica della val Bregaglia (Bergell) di R. Staub, classificata come suddivisione dei Grigioni, essendo esclusa dalla classe stessa dei Grigioni.

Note

- * In OPAC; nella bibliografia da più di 35 anni.
- ** Rete NEBIS delle biblioteche e dei centri d'informazione in Svizzera; include oltre 80 biblioteche delle università, delle scuole professionali e degli istituti di ricerca. Esiste dal 1999. ETHICS, il sistema di controllo delle informazioni della Biblioteca dell'ETH è stato utilizzato dal 1983 al 1999. (ETHICS =ETH Library Information Control System)
- *** (tedesco, inglese e francese). L'ETH (Eidgenössische Technische Hochschule) di Zurigo e l'EPFL (Ecole Polytechnique Fédérale) di Losanna costituiscono il Politecnico Federale Svizzero.

Nota storica: secondo le DK-Tafeln (Tabelle CDU) (1978, 1986: p. 18, 16), il primo contatto per domande riguardanti l'applicazione dell'elaborazione dei dati elettronici della CDU è stata la Biblioteca dell'ETH, insieme all'Ufficio Centrale per la Documentazione Automatica (ZMD, Zentralstelle für die Maschinelle Dokumention) nella Società per l'Informazione e la Documentazione (GID, Gesellschaft für Information und Dokumentation) di Francoforte. Anche la guida della CDU di McIlwaine (2000, 2007) riferisce che la Biblioteca dell'ETH di Zurigo utilizza la CDU e cita l'indice per soggetti multilingue e le opzioni multilingue delle domande. Successivamente, negli anni Novanta, il sistema GERHARD (GERman Harvest Automated Retrieval and Directory) di Oldenburg (Germania) ha reso la versione CDU dell'ETH accessibile in internet.