

MACS (Multilingual Access to Subject) Un *authority file* virtuale multilingue

Genevieve Clavel-Merrin
Biblioteca nazionale svizzera

Introduzione: l'*authority control* diventa maggiorenne

Gli *authority file* partecipati e la cooperazione nello sviluppo di liste nazionali, sia per autore che per soggetto, hanno permesso alle biblioteche di condividere risorse e migliorare l'accesso alle proprie collezioni. Sono già state sottolineate le difficoltà di ricerca se ci si sposta da un approccio nazionale a uno internazionale nel campo della condivisione di risorse, ad esempio con il progetto TEL (The European Library, <<http://www.europeanlibrary.org>>). Variazioni nelle intestazioni per autore personale ed ente e nelle intestazioni per soggetto attraverso lingue diverse hanno comportato che di frequente l'opzione di ricerca scelta sia una parola chiave. Questo modo di procedere porta a dei risultati, ma i suoi limiti sono ben noti.

Sono in corso tentativi di ampliare l'importante lavoro, fatto in molti paesi, di costruzione e aggiornamento dei linguaggi di indicizzazione per soggetto e degli *authority file*. Questo è possibile per una serie di ragioni, sia di carattere organizzativo che tecnico: l'uso crescente di formati standard di comunicazione, la preparazione e la manutenzione di manuali ben strutturati e lo sviluppo di una gestione distribuita sofisticata e di strutture di aggiornamento basate sull'accessibilità diffusa di un ambiente di rete, che rende possibile alle istituzioni l'accesso sempre maggiore a sistemi di *authority control* centralizzati sviluppati precedentemente.

Risulta facilitata l'applicazione dell'*authority control* alle collezioni con l'uso degli stessi linguaggi di soggettazione, dal momento che le biblioteche possono scambiare e copiare registrazioni e, di conseguenza, migliorare la coerenza (monolingue) di indicizzazione nelle collezioni.

Inoltre, riconoscendo che molti linguaggi di indicizzazione per soggetto seguono costruzioni e principi applicativi simili (vedi *Principles underlying subject heading languages* dell'IFLA [1]), il lavoro preparatorio consiste nel permettere alle biblioteche di costruire sulla solida infrastruttura per l'*authority control* a un livello monolingue per poi estendere il controllo e l'accesso a un livello multilingue.

Mentre per tradizione l'accesso bilingue/multilingue è stato assicurato tramite intestazioni per soggetto multiple codificate negli *authority record*, oggi molti progetti di ricerca cercano di fornire accesso mediante processi di collegamento basati su Internet. Concetti come *language cross-linking*, o interoperabilità nei linguaggi di indicizzazione per soggetto, che sono comuni nell'odierna ricerca nell'accesso multilingue, fino a dieci anni fa erano patrimonio esclusivo dei linguisti.

MACS utilizza l'ambiente stabile dell'*authority control* e questo processo di collegamento per estendere l'uso dei linguaggi di indicizzazione per soggetto a un livello multilingue connesso su scala internazionale.

L'origine del MACS

MACS (accesso multilingue ai soggetti) è stato creato nel 1997 in risposta alla richiesta della Conference of European National Librarians (CENL) di trovare una soluzione al problema dell'accesso multilingue per soggetto nei database europei. È ovvio che la Svizzera, essendo un paese multilingue, avesse particolare interesse a trovare soluzioni, ma è significativo che altri paesi fossero interessati a questo approccio multilingue anche se considerati tradizionalmente monolingue. Questo per sottolineare i bisogni internazionali degli utenti e la necessità di cooperare nella costruzione e nell'accesso alle collezioni non solo a livello regionale o nazionale, ma sempre più in una sfera internazionale. Per discutere la questione è stato istituito un gruppo di lavoro e, tra i molti approcci discussi, l'idea di creare collegamenti tra diversi linguaggi di indicizzazione per soggetto è sembrata una soluzione promettente per quattro fra le biblioteche partecipanti. Le quattro biblioteche, la Swiss National Library (SNL), la Bibliothèque nationale de France (BnF), la British Library (BL) e la Deutsche Bibliothek (DDB) hanno accettato il compito di definire un'idea base e di condurre uno studio per stabilirne la fattibilità, i cui dettagli sono precisati più avanti. Il progetto

MACS è stato allora formalizzato con il suo obiettivo principale: «fornire gli strumenti tramite i quali si possa accedere ai database delle biblioteche su una base multilingue grazie all'uso di intestazioni equivalenti provenienti da *authority file* per soggetto». L'approccio scelto consisteva nel creare un accesso multilingue esteso che potesse essere usato il più ampiamente possibile tramite linguaggi di indicizzazione che sono comunemente usati nella maggior parte delle biblioteche in ogni determinato paese.

Questo accesso multilingue è reso possibile grazie a collegamenti di equivalenza stabiliti fra i tre linguaggi di indicizzazione per soggetto usati nelle biblioteche partner: SWD/RSWK per il tedesco, RAMEAU per il francese e LCSH per l'inglese.

Ciascuno di questi linguaggi di indicizzazione per soggetto è ampiamente usato a livello nazionale e internazionale, ad esempio l'uso di SWD è molto diffuso in Germania ma anche in Austria e in Svizzera, RAMEAU in Francia nelle biblioteche universitarie, comunali e pubbliche e sta cominciando a essere utilizzato in altri paesi francofoni, mentre la diffusione di LSCH nei paesi di lingua inglese e in istituzioni internazionali è ben nota.

La decisione di usare l'approccio dei collegamenti ha anche tenuto conto del fatto che le biblioteche (francese, tedesca e inglese) hanno investito tempo e sforzi considerevoli nella creazione e nel mantenimento di ciascun linguaggio di indicizzazione (LSCH, RAMEAU e SWD) ed è stata presa considerando anche quali linguaggi siano stati usati in maniera estesa per fornire accesso per soggetto a milioni di documenti, sia nelle biblioteche nazionali che, come notato prima, in biblioteche di tutti i paesi interessati e altrove. Creare un nuovo vocabolario multilingue basato sulla traduzione sarebbe stato poco realistico e antieconomico. L'uso delle risorse disponibili conserva l'accesso ai documenti già indicizzati ed estende l'accesso oltre i confini linguistici.

Lo studio ha portato a un sottoinsieme di intestazioni tratto dai tre linguaggi nel campo del Teatro e dello Sport, più le intestazioni usate con maggiore frequenza per indicizzare documenti nella Bibliothèque nationale de France. Il gruppo ha studiato differenti metodi di mappatura, includendo il confronto tra indici dello stesso documento in linguaggi di indicizzazione differenti, prima di stabilire manualmente collegamenti tra le intestazioni. Sono stati analizzati i metodi per risolvere la mappatura uno-a-molti oppure multi-a-molti e sono stati stabiliti i principi della creazione e della gestione cooperativa dei collegamenti. I risultati dello studio sono stati incoraggianti e hanno mostrato che il sistema del collegamento delle intestazioni è un approccio fattibile per l'accesso *cross-language* a documenti, in molti sebbene non in tutti i casi. Una combinazione o un collegamento sono stati considerati riusciti quando un concetto, rappresentato da intestazioni simili nei diversi linguaggi e collegate manualmente (intellettualmente) davano risultati quasi equivalenti nel recupero per soggetto.

Dettagli sullo studio di fattibilità possono essere trovati all'indirizzo <<http://infolab.kub.nl/prj/macs>>. Il gruppo ha raccomandato lo sviluppo di un prototipo per testare la struttura di gestione, la creazione di collegamenti e la ricerca multilingue.

Il prototipo

Dando seguito a una gara d'appalto riservata, il contratto per sviluppare il sistema è stato aggiudicato alla Index Data aps e alla Infolab Tilburg. Il sistema è disponibile per un accesso di prova all'indirizzo <<http://infolab.kub.nl/prj/macs>>.

Il prototipo contiene i seguenti dati:

- tutte le intestazioni dalle aree di soggetto Sport e Teatro i cui termini erano stati connessi nello studio di fattibilità, più le 500 intestazioni maggiormente usate identificate prima nel database RAMEAU e le intestazioni loro equivalenti negli altri linguaggi, per un totale di circa 3000 intestazioni;
- 1200 collegamenti creati tra concetti;
- registrazioni bibliografiche contenenti queste intestazioni provenienti da ciascuno dei database dei quattro partner per consentire al sistema di simulare ricerche nei cataloghi di biblioteca usando il protocollo Z39.50.

Il prototipo è un'applicazione Web basata su standard, indipendente dai sistemi dei partner e dai linguaggi di indicizzazione usati. Ha due moduli:

- Il *Link management interface* permette la creazione e la manutenzione dei collegamenti, con meccanismi per la gestione di collegamenti parziali o complessi. Il sistema è basato sul principio della *federative management*, e cioè ogni linguaggio è autonomo con ciascun partner responsabile per la gestione dei collegamenti nel proprio linguaggio (responsabilità condivisa tra DDB e SNL per il SWD), senza l'esistenza di un curatore centrale. I partner possono fare proposte sulle altre parti di un collegamento, ma queste devono essere ratificate dal partner appropriato. Una funzione di annotazione consente ai partner di commentare un collegamento o una proposta, oppure di indicare dove siano sopravvenuti dei problemi. Nella gestione dei collegamenti sono osservati i seguenti principi:
 - tutti i linguaggi hanno pari status; non c'è un linguaggio cardine. Al contrario, il sistema contiene applicazioni che sono considerate come grappoli concettuali, individuati solo da un identificatore astratto (numerico);
 - strutture gerarchiche, relazioni thesaurali o termini non accettati non sono mappati o riprodotti come parte del processo di collegamento delle singole intestazioni: sono oggetto di collegamento solo le intestazioni accettate;
 - quando non è possibile trovare un'equivalenza, un'intestazione proposta può restare da sola nel sistema per rappresentare il concetto (per possibili future applicazioni).
- La *User search interface* permette agli utilizzatori finali di effettuare, nella lingua da loro scelta, ricerche per soggetto simultaneamente in determinati cataloghi (attualmente sottoinsiemi dei cataloghi dei partner), usando il protocollo Z39.50. L'utente inserisce un termine di ricerca da una delle liste di soggetto disponibili (al momento LCSH, RAMEAU o SWD), sceglie uno o più cataloghi su cui compiere la ricerca, e in seguito può lanciare direttamente la ricerca nei cataloghi prescelti o scegliere di scorrere una lista di intestazioni corrispondenti e tra loro equivalenti, dalla quale è possibile scegliere uno o più grappoli concettuali prima di lanciare la ricerca. Controlli hanno dimostrato che quest'ultimo approccio conduce a risultati migliori e produce meno distorsione tra le risorse scelte. Dalla serie di risultati della ricerca l'utente può accedere a singole registrazioni bibliografiche in una qualsiasi delle biblioteche semplicemente cliccando sul titolo. Nel prototipo sono fornite opzioni per una visualizzazione parziale o completa nel formato MARC usato dalla biblioteca selezionata.

La sperimentazione

MACS ha reso disponibile dal 2001 la ricerca e l'accesso di sola consultazione al *Link management interface*. Durante quel periodo, oltre 500 persone di tutto il mondo si sono iscritte per testare il prototipo MACS. Dai commenti ricevuti e dalle sperimentazioni condotte dai partner MACS è derivato il perfezionamento del modello organizzativo e la progettazione di una nuova versione del sistema MACS. A dispetto dei miglioramenti nei meccanismi di collegamento, il compito fondamentale di stabilire collegamenti continuerà a richiedere molto tempo. Fortunatamente i partner possono basarsi sui collegamenti bilingue esistenti negli *authority file* di RAMEAU gestiti dalla Bibliothèque nationale de France e caricati nel sistema MACS. Approssimativamente 70.000 voci di autorità di RAMEAU contengono proposte di intestazioni di LCSH ed è stato loro assegnato un numero di dominio per soggetto. Il lavoro di aggiunta delle intestazioni SWD a queste intestazioni comincerà nel 2003 e sarà diviso tra i partner in Germania e in Svizzera a seconda dei domini. Nel frattempo, saranno studiati altri metodi per facilitare la creazione di collegamenti.

L'integrazione del MACS nell'ambiente reale

I risultati della ricerca hanno mostrato che possono essere reperiti documenti pertinenti usando collegamenti fra intestazioni per soggetto, ed è in progetto un trasferimento dall'uso dei sottoinsiemi ai database dei partner. Nell'aprile 2002 sono cominciati test con il catalogo della Deutsche Bibliothek, nell'estate 2002 la Bibliothèque nationale de France ha portato a termine con successo sperimentazioni e la Biblioteca nazionale svizzera ha condotto simili test nell'inverno 2002. Saranno necessarie modifiche nella *Search interface* per

garantire che gli utenti possano avere maschere e risultati di ricerca facilmente gestibili, e naturalmente è essenziale un numero crescente di collegamenti. L'interesse della comunità bibliotecaria verso il MACS è sottolineato dalla constatazione che le biblioteche nazionali partner del progetto TEL desiderano integrare il prodotto nel portale TEL, che nella sua fase di prototipo ha lo scopo di offrire l'accesso alle risorse di otto biblioteche nazionali europee. Ai fornitori del portale sarà chiesto di proporre modalità grazie alle quali MACS possa essere integrato, al livello più semplice, con un collegamento al sistema o forse inglobando gli stessi dati del MACS.

Incorporare una effettiva gestione di autorità (*vedi e vedi anche*)

Come da progetto iniziale, solo le intestazioni accettate sono oggetto di collegamento. I partner ammettono che non è sempre facile per gli utenti trovare i termini accettati per la ricerca senza l'aiuto dei rinvii *vedi e vedi anche*. L'obiettivo è di incorporare un accesso agli *authority file* nazionali come opzione per gli utenti che desiderino controllare le intestazioni e raffinare la loro ricerca prima di inoltrarla ai database scelti. Inoltre, nessun linguaggio di indicizzazione rimane statico, le intestazioni possono essere modificate in risposta a cambiamenti linguistici o a nuove scoperte. È essenziale che tali cambiamenti siano trasferiti nel *Link management system* (e anche, naturalmente, nei cataloghi delle rispettive biblioteche). Nel caso di una semplice sostituzione potrebbe essere possibile un aggiornamento del *Link management system*: ciò richiederà che i gestori dei linguaggi di indicizzazione siano in grado di individuare ed estrarre le intestazioni modificate e il loro ID, e che il *Link management system* carichi e faccia combaciare questi cambiamenti. Dal momento che il *Link management interface* contiene gli ID per le intestazioni accettate, i collegamenti per la navigazione e gli aggiornamenti, nel caso di intestazioni accettate modificate, possono essere incorporati senza il bisogno di duplicare le informazioni.

Il futuro

Il valore immediato per gli utenti della Biblioteca nazionale svizzera (SNL) è la possibilità reale che nel prossimo futuro le ricerche per soggetto siano condotte in più di una lingua. Come istituzione federale, alla Biblioteca nazionale svizzera è affidato il compito di fornire servizi ai propri utenti nelle tre lingue ufficiali della Svizzera (tedesco, francese e italiano). La decisione della SNL nel 1998 di adottare SWD come proprio standard di indicizzazione fu presa assicurandosi che anche gli utenti francesi e italiani avrebbero tratto beneficio da quello strumento di indicizzazione, sebbene multilingue. Andando oltre i confini nazionali, altri utenti della SNL e di altre biblioteche nazionali troveranno le proprie ricerche facilitate da un *gateway* (portale) di accesso multilingue ai loro cataloghi. La Deutsche Bibliothek considera il progetto come un modo per promuovere il valore di SWD e renderà i risultati del MACS disponibili ai suoi partner in Germania e in Austria. La Bibliothèque nationale de France ha già cominciato a valutare l'esame dei modi di estendere i collegamenti del MACS a livello nazionale.

Estendere i collegamenti ad altri linguaggi di indicizzazione per soggetto è un obiettivo, anche se i partner futuri devono essere consapevoli che l'investimento in termini di risorse umane è considerevole, dal momento che, allo stato attuale, il lavoro di collegamento è analizzato ed eseguito manualmente. Finché il sistema operativo è in continuo cambiamento è ancora difficile per i partner valutare con precisione il tempo necessario per la creazione e il controllo di un collegamento, ma stime preliminari mostrano che il tempo richiesto per un collegamento può variare dai due agli otto minuti, a seconda dell'area semantica e della complessità. Perciò ogni partner potenziale dovrebbe avere una infrastruttura simile a quelle presenti in Francia o in Germania, dove la gestione del linguaggio di indicizzazione per soggetto è diffusa tra istituzioni in partecipazione.

Un *authority file* virtuale?

Attualmente MACS è limitato alle intestazioni autorizzate dei linguaggi di indicizzazione per soggetto partecipanti e, come Freyre e Naudi hanno mostrato, si può dire che svolge, sotto alcuni punti di vista, la funzione di un dizionario multilingue di linguaggi di indicizzazione per soggetto, un «ponte tra sistemi che sono essi stessi progettati per organizzare e designare concetti [2]». Incorporare un collegamento agli altri

elementi degli *authority record* all'interno degli *authority file* nazionali dei partner porterà MACS a creare un *authority file* multilingue virtuale.

Note bibliografiche

[1] IFLA Section on Classification and Indexing, Working group on principles underlying subject heading languages. *Principles underlying subject heading languages (SHLS)*, editors Maria Inês Lopes and Julianne Beall. München: Saur, 1999.

[2] Elisabeth Freyre - Max Naudi. *MACS: Subject access across languages and networks*. In: *Subject retrieval in a networked environment: papers presented at an IFLA satellite meeting sponsored by IFLA Section on classification and indexing & IFLA Section on information technology, OCLC, Dublin, OHIO, USA, 14-16 August 2001*. Dublin (OH): OCLC, 2001.