

tecniche di ricerca: non si fa parola, o quasi, dei MeSH, delle strutture ad albero, di esplosioni e *subheadings*, non si parla delle complessità della ricerca semantica, ma si illustra il funzionamento di ogni singolo pulsante delle schermate del Pubmed e del NLM Gateway, e di ogni singola voce di menu di Ovid Medfive e dei Current Contents, con l'obiettivo di esplorarne le funzionalità e far capire al sanitario frettoloso che cosa queste interfacce possono offrirgli.

La seconda sezione è dedicata agli strumenti della "gestione bibliografica", ovvero alla fase successiva al reperimento di una lista di citazioni. Viene infatti illustrato il funzionamento di un software di gestione di citazioni bibliografiche, Endnote versione 4.0, e si affronta poi il tema scottante dell'accesso alle fonti primarie (più banalmente: l'accesso diretto via Internet al testo integrale dei periodici di medicina o l'accesso tramite prestito interbibliotecario) al quale la stringatezza del manuale non consente però di dedicare più di nove pagine; a coronamento della "gestione bibliografica", troviamo una pagina contenente pillole di saggezza su come archiviare la versione cartacea degli articoli reperiti, con l'accorto utilizzo di classificatori di cartone e, magari, del già elogiato Endnote!

La terza sezione, intitolata un po' ampollosamente "gli strumenti della comunicazione", è di particolare interesse in quanto spiega come si prepara una bibliografia (argomento che gli specialisti spesso non conoscono affatto) e illustra i diversi standard internazionali in materia; si spiega poi come è possibile, con l'ausilio di appositi strumenti software (Endnote, ovviamente), formattare automaticamente le citazioni scaricate dalle basi bibliografiche. Seguono poi brevi accenni ai traduttori automatici (che opinabilmente l'autrice ritiene un aiuto efficace al trasferimento dei dati da una lingua all'altra), al salvataggio e al trasferimento dei dati (ad esempio con la posta elettronica). Infine, sempre con Endnote, si impara come creare e gestire un database di indirizzi Web.

Le ultime pagine del testo prima della bibliografia contengono un elenco di siti Web di interesse biomedico, con annotazioni che sembrano desunte dai siti stessi in quanto sono redatte parte in francese e parte in inglese.

Giovanni Arganese

Biblioteca medica statale, Roma

Norman Desmarais. *The ABC of XML: the librarian's guide to the eXtensible markup language*. Texas: Houston Press: New Technology Press, 2000, 206 p. ISBN 0-96755942-0-0. \$ 25,99.

Il libro fornisce una panoramica dell'XML e di alcune sue applicazioni, anche in ambito bibliotecario, aprendo la disamina con uno sguardo alle relazioni con SGML e HTML e alle sue varie componenti. Viene poi descritta la struttura di un documento XML ed esplorati i diversi tipi di *style sheets* per le opzioni di formattazione e di presentazione dei documenti passando a esaminare le componenti *linking* e *pointing*, che già con HTML hanno reso il Web la più ricca risorsa informativa esistente. Uno dei grandi vantaggi dell'XML è la sua caratteristica di essere leggibile sia dall'uomo sia dalle macchine, con condivisione dei dati per la programmazione: un documento XML contiene infatti un misto di testo e *tags* di *markup* che ne organizzano e identificano le componenti. Il quinto capitolo discute questo aspetto, evidenziando la nuova energia che XML offre al Web, in particolare per il commercio elettronico (*e-commerce*) considerato dagli esperti la *killer application* per accelerare l'adozione dell'XML e catapultarlo nell'uso più ampio. Parte della trattazione è dedicata alla gestione dei documenti XML; un glossario e una bibliografia completano la pubblicazione. Nato prevalentemente per l'uso nel Web, XML può in realtà essere usato per qualsiasi tipo di pubblicazione elettronica: a differenza di SGML che, essendo uno standard per il *text processing*, descrive il layout e la struttura di un docu-

mento, XML è in grado di descriverne il contenuto informativo. Rispetto invece a HTML, che ha un set predefinito di *tags*, permette all'utente di definire proprie *tags*, con una flessibilità notevole: dal momento che il numero di *tags* è praticamente infinito, il contenuto di un documento può essere descritto con un alto livello di precisione, fino alla descrizione della singola parola o termine e i documenti creati in XML possono essere usati in applicazioni diverse, senza necessità di conversione. In riferimento alla comunità bibliotecaria, Desmarais sostiene che il formato MARC potrebbe essere facilmente utilizzato come *XML Data Schema* (ossia come *Document Type Definition*) per applicazioni bibliotecarie, dal momento che contiene tutte le informazioni necessarie a coloro che operano nel campo editoriale, al punto da definirlo «the grand-daddy of all DTDs» (la Library of Congress Network Development e MARC Standards Office sta sviluppando un framework che consente di lavorare con dati MARC (e MARC 21) in ambiente XML <<http://www.loc.gov/marc/marcxml.html>>. Per la conversione automatica e la validazione è disponibile un ricco toolkit all'URL <<http://www.loc.gov/standards/marcxml/marcxml.zip>>. Nel secondo capitolo Desmarais presenta una panoramica delle *tags* XML e del vocabolario usato dai documenti così codificati: ogni documento presenta infatti una struttura logica, che ne determina l'organizzazione concettuale, e una struttura fisica che determina come le varie parti sono memorizzate e organizzate. Uno dei punti di forza dell'HTML è stata la sua abilità di rendere i documenti interpretabili da molti Web browser per la visualizzazione *on screen* o riformattabili per la stampa, l'e-mailing oppure esportabili per uso in altre applicazioni software. Un documento HTML contiene attributi che specificano come esso deve apparire sullo schermo; le *tags* devono definire sia il contenuto sia le funzioni di visualizzazione. Ciò implica che qualsiasi cosa sul Web è strutturata per un determinato mezzo di comunicazione (*medium*) tramite il quale si può "vedere" l'intera pagina come è stata progettata: cambiando le caratteristiche della tecnologia di visualizzazione il documento è soggetto a molte distorsioni (caso di un documento contenente colonne o tabelle: se lo convertiamo in un altro formato modificando margini, font e grandezza del carattere, cambia il layout e la presentazione di colonne e tabelle fino a renderlo incomprensibile). XML è innovativo in quanto le Document Style Sheets separano il contenuto dalle sue caratteristiche di visualizzazione permettendo la selezione di una specifica di stile per un certo mezzo di output senza che il contenuto o il *markup* ne risentano. Quando contenuto e stile sono uniti è infatti molto difficile personalizzare una pagina Web per differenti viste destinate a classi di utenti diverse o uno stesso utente in visite diverse al sito. Riguardo alla funzionalità di *linking*, l'autore sottolinea che la massiccia diffusione di HTML nel mondo Internet è fortemente legata alla sua capacità di sviluppare l'*hyperlinking*: per essere usato come tecnologia Web l'XML ha dovuto sviluppare un proprio meccanismo di *linking* compatibile con i meccanismi esistenti di HTML, pur conservando la sua estensibilità e *robustness*. Le differenze tra i meccanismi di *linking* dei due linguaggi risiedono fondamentalmente nel fatto che XML supporta link unidirezionali e multidirezionali e inoltre i link HTML accedono unicamente ai documenti a livello di pagina mentre quelli XML arrivano a livello di oggetto (singolo paragrafo o sezione). Relativamente alla gestione dei documenti in XML, Desmarais sottolinea che esso è di aiuto nella risoluzione di problemi legati alle varie fasi di sviluppo dei documenti sul Web, grazie alle funzionalità che il software può utilizzare per la loro creazione e gestione a lungo termine. Un database relazionale usato per memorizzare documenti XML può segmentare i documenti nelle loro componenti gerarchiche: il sistema più semplice usa una tabella relazionale che assegna un identificatore unico a ogni record e uno pseudo-elemento a ogni elemento. Ma i limiti dei database relazionali sono noti, *in primis* la difficoltà di rappresentare gerarchicamente informazioni strutturate in un modello tabellare. In alternativa i più recenti database a oggetti

usano una struttura più libera rispetto a tabelle costituite da righe contenenti i campi: gli oggetti costituiscono le unità informative (semplici dati o puntatori ad altri oggetti) identificabili univocamente e contengono attributi, ciascuno con un nome e un valore, esattamente come in XML. Quindi il database a oggetti appare più appropriato per registrare documenti XML poiché le relazioni sequenziali e gerarchiche possono essere descritte più facilmente. Un sito Web HTML/XML risulta facilmente gestibile in tale ambiente ma le difficoltà sorgono al crescere del sito, quando è necessario contare sul supporto software per il mantenimento, esattamente come nel Document Management System. Quanto ai motori di ricerca, la tecnologia base usata per localizzare un documento contenente parole o frasi specifiche è l'indice invertito, in cui le parole sono ordinate alfabeticamente e hanno link ai documenti che le contengono, con la posizione esatta dei termini nel documento originale: una versione XML dell'indice può registrare anche la posizione iniziale e finale di ogni elemento incluso nel file. Il vantaggio è che l'XML rende possibile lo *storage* di documenti come parti piuttosto che come ampi blocchi di informazioni spesso indecifrabili: un Document Management System in XML tiene quindi conto dell'organizzazione fisica (entità) e logica (elementi) del documento e può memorizzare i documenti come set di elementi all'interno di un database organizzato gerarchicamente, fornendo un'efficace gestione delle risorse informative ed evitando la ridondanza. Molti editori fanno uso di tali strumenti, detti sistemi editoriali, di solito arricchiti dalla funzione di *workflow*, con i quali possono gestire file di testo, immagini e file di layout di pagine e includere funzioni per quantificare il tempo che un operatore impiega su ogni documento. Nel caso di modifiche al singolo paragrafo di un libro, un sistema editoriale in XML non richiede di aprire l'intero libro né il capitolo, ma consente l'accesso direttamente al paragrafo per rieditarlo in modo autonomo. Come già anticipato, il settimo capitolo tratta estesamente delle potenzialità dell'XML nell'e-commerce, con un *excursus* storico del fenomeno, a partire dal 1948, fino al riconoscimento dell'X12 EDI (Electronic Data Interoperability) per operazioni bancarie, trasporti e funzioni di vendita, operazioni per cui il governo forniva formulari elettronici e messaggi per la spedizione e l'acquisto. Si tratta di un set di standard che facilitano i processi B2B quali ordini, adempimenti e transazioni finanziarie in cui XML apporta innovazione tramite l'uso delle sue *meta-tags* per identificare i *data elements*: le società possono così "etichettare" i loro dati una volta per tutte e usarli in diverse applicazioni, dalla preparazione dei cataloghi, all'evasione degli ordini, alla fatturazione. Ciò ottimizza i processi di vendita dei beni e apre il mercato a qualsiasi partner che disponga di una connessione a Internet e di un browser in grado di usare XML, facilitando l'ingresso di società più piccole sul mercato elettronico e permettendo ai *vendors* di offrire prodotti direttamente a un acquirente, eludendo forme di intermediazione, poiché tutte le procedure vengono effettuate dalla parte client, il computer del *buyer*. Le biblioteche che effettuano un numero enorme di transazioni ma a costo limitato per item, non hanno alcun interesse a un'estesa implementazione di EDI mentre l'uso di XML ridurrebbe comunque i costi e il carico di procedure poiché permette lo scambio dei dati indipendentemente dal sistema operativo o dalle applicazioni di *accounting* in uso.

Concludendo, a XML viene riconosciuto un grande potenziale d'uso, soprattutto nell'e-commerce, cui è dedicato il capitolo più esteso sulle potenzialità di XML. L'ambito bibliotecario è trattato in fondo in modo marginale, nonostante il titolo focalizzi l'attenzione su questo settore: probabilmente più che per i bibliotecari il libro parla dell'XML visto da un bibliotecario dal momento che l'autore ricopre questo incarico presso il Providence College (Rhode Island, USA) e si occupa di acquisizioni.

Daniela Canali
ISPRI – CNR, Roma