

# Con DoIS l'informazione professionale è in linea

*Un progetto per descrivere e rendere disponibile  
su web la documentazione di ambito LIS*

di Antonella De Robbio

DoIS (Documents in Information Science)<sup>1</sup> è un servizio internazionale posto su web che raccoglie la letteratura di ambito LIS (Library and Information Science), articoli di periodici, comunicazioni presentate a convegni e conferenze, rapporti di ricerca.

DoIS internazionale nasce dallo sforzo di un gruppo di volontari composto da bibliotecari e professionisti del settore al fine di creare una risorsa bibliografica specializzata in Library and Information Science che possa offrire contenuti scientifici selezionati, gratuitamente accessibili agli addetti ai lavori e agli utenti in generale, ricercabili attraverso un motore di ricerca non commerciale.

DoIS è presente e indicizzato in: Google's Web Directory for Library and Information Science. È da considerare che si tratta di un servizio giovane, sorto poco più di un anno fa.

Come segnalato nel "D-Lib Magazine" del settembre 2000,<sup>2</sup> alla data del 9 settembre 2000 DoIS conteneva appena mille descrizioni provenienti dalle seguenti sette riviste: "Ariadne", "Boletin de RedIris", "Bulletin des Bibliothèques de

France", "CLIR Reports", "Cuadernos de Documentacion Multimedia", "D-Lib Magazine", "Journal of Electronic Publishing".

Attualmente, dopo un anno, DoIS comprende le descrizioni di 7.582 documenti: 5.648 articoli (di riviste) e 1.622 altri paper, di cui 3.702 accessibili liberamente a testo pieno (dati relativi al primo novembre 2001).

Tra le numerose riviste e serie di paper indicizzate da DoIS ricordiamo:

- "ALA, Association for Library Collections & Technical Services"
- "Ariadne"
- "Aslib proceedings"
- "Cataloging & classification quarterly"
- "CLIR Reports"
- "Council on Library and Information Resources"
- "Cybermetrics"
- "D-Lib Magazine"
- "First Monday"
- "High Energy Physics Libraries Webzine"
- "IFLA journal"
- "Information technology and libraries"
- "Journal of documentation"
- "Journal of Education for Library and Information Science"

- "Journal of Electronic Publishing"
- "Journal of information science"
- "Journal of the American Society for Information Science"
- "Knowledge organization"
- "Library trends"
- "Reference librarian"
- "Reports of the AAU Task Forces"
- "SBN Notizie"
- "Technical services quarterly"

Il team di DoIS internazionale è composto da: Julio Alonso Arévalo, University of Salamanca, Faculty of Information Sciences; José Manuel Barrueco Cruz, University of Valencia; Antonella De Robbio, Università degli studi di Padova; Thomas Krichel, assistant professor at the Palmer School of Library and Information Science, Long Island University (New York, USA); Natividad Noverges, Generalitat Valenciana, Valencia.

DoIS non è un catalogo tradizionale o una classica bibliografia, benché contenga attualmente anche riferimenti a documenti presenti in forma cartacea; ciò è dovuto al fatto che, in una prima fase di conversione dei dati - iniziata con il riversamento dei documenti LIS di ambito spagnolo - sono confluiti anche record che non possiedono link alla versione elettronica.

Lo standard di metadati adottato da DoIS è ReDIF (Resource Description Information Format).<sup>3</sup> DoIS è un progetto di catalogazione di risorse documentarie basato su standard ReDIF e protocollo Guildford.<sup>4</sup>

ReDIF è un formato "relazionale" di metadati. Questo significa che ogni record rappresenta univocamente le caratteristiche di un determinato oggetto. Ogni record ha un identificatore, che viene chiamato *handle* (maniglia). Un record può contenere identificatori di altri record: le relazioni fra gli elementi del mondo descritto vengono così rappresentate in forma compatta, senza duplicare le stesse informazioni fra record interessati dalle relazioni.

L'infrastruttura del servizio DoIS si basa sul modello di tipo distribuito utilizzato per la rete RePEc,<sup>5</sup> progetto internazionale per i preprint di economia.

Va anche ricordato che in ReDIF sono stati convertiti i formati di altri sei server OAI all'interno del prototipo UPS (Universal Preprint Server).<sup>6</sup> Questo significa che esiste una *mapping* tra ReDIF e gli altri formati OAI compatibili, tra cui Dublin Core.

Si è scelto questo modello per il DoIS appunto per la sua alta compatibilità e trasportabilità.

La descrizione delle risorse elettroniche avviene attraverso metadati internazionali ReDIF all'interno di un servizio che consente il recupero della versione elettronica, se esistente in rete, da parte dell'utente finale, in questo caso il bibliotecario, o chiunque sia interessato alla letteratura di settore LIS. DoIS ha due lati, una vista "utente" per il recupero dei documenti presenti in rete e una vista "per gli addetti ai lavori", zona grigia dove si creano e si trasportano i metadati.

La parte riservata agli addetti ai lavori non è altro che un insieme di aggregazioni informative organizzate sulla base di un sistema a modello distribuito, completamente gratuito e funzionante sul protocollo FTP (File Transport Protocol).

All'interno della struttura a modello distribuito chiunque può descrivere risorse (proprie o di altri enti) attraverso metadati descrittivi standard, collocarli su un server FTP anonymous o web e comunicare al server centrale (nel caso di DoIS a Manchester) dove reperire le informazioni (metadati). Attraverso un motore che funziona con un meccanismo simile a quello utilizzato dai *webcrawlers*, e posto sul server centrale, vengono raccolti dai vari server, tutte le notti, i metadati presenti, di modo che al mattino successivo DoIS sia in grado di

elaborare le informazioni e presentarle in modo aggiornato all'utenza finale.

In questo modo è possibile lavorare in modo cooperativo ma autonomo, alla creazione "partecipata" di metadati, coinvolgendo direttamente gli enti produttori di informazione (biblioteche, università, associazioni...), distribuendo il lavoro all'interno di un sistema che consente la massima visibilità dei documenti, la loro disseminazione, senza costi aggiuntivi tranne quelli del personale coinvolto nella descrizione delle risorse (creazione dei metadati).

Il trasporto dei metadati dai server distribuiti al server centrale avviene grazie a una struttura gerarchica che organizza le informazioni in metadati di archivio e di serie (metadati di tipo MAG amministrativi gestionali) e metadati descrittivi, pertinenti alle risorse da descrivere. Il server centrale conserva la lista con gli indirizzi di tutti i server distribuiti. È attraverso questa lista che il motore è in grado ogni notte di "spazzolare" tra i folder locali negli URL segnalati, recuperando i file (contenenti uno o più metadati), copiandone i contenuti, per poi trasportare le copie dei file al server centrale che provvederà a creare un indice collettivo, mantenendo al contempo anche "copia della struttura locale".

Il formato dei file è ReDIF, da non confondere con RDF (Resource Description Framework), che non è uno standard di metadati ma un modello di metadati, o meglio una piattaforma per il trattamento dei metadati, in quanto costituisce un'ipotesi di soluzione individuata dal W3C al problema della rappresentazione dei metadati.

In DoIS tutti i file presentano un'estensione ".rdf."; sono formatati in modo testuale per una loro visibilità in formato puro testo ASCII, sul lato per gli "addetti ai lavori".

Obiettivo di ReDIF è la descrizione non solo dei documenti scientifici, ma anche degli autori che li producono, dei canali attraverso i quali tali documenti vengono resi pubblici, dei processi di controllo su questi canali, delle persone e delle istituzioni che gestiscono e organizzano queste attività. Per il momento DoIS non applica appieno tutte le funzionalità offerte da ReDIF come sistema relazionale. Questo ovviamente per il fatto che, se per ambiti disciplinari come l'economia, la matematica, la fisica (prima fra tutte), esistono numerosi archivi elettronici contenenti documenti a testo pieno (e-print server), per il contesto LIS non abbiamo, al momento, tale configurazione.

DoIS, per questa ragione, utilizza solo una parte del modello ReDIF, pur cogliendone appieno la filosofia portante. A livello di organizzazione dell'informazione, DoIS utilizza i manuali tecnici per la creazione dei metadati che avviene con template o modelli differenziati.

I metadati creati attraverso dei template ispirati ai modelli IAFA sono ricercabili dall'utenza tramite un motore di ricerca messo a punto allo scopo attraverso strumenti come ROADS,<sup>7</sup> possono risiedere su server differenti e comunque sono "esterni" al documento a testo pieno (risorsa elettronica).

Dopo la rielaborazione dei metadati che in questo modo sono presenti in due luoghi distinti, presso il server locale e presso il server centrale, che ne copia la struttura organizzativa, avviene la rielaborazione in formato HTML.

Tale elaborazione consente la creazione della vista utente con la possibilità di ricercare attraverso una semplice stringa nel contenuto di tutti i metadati presenti in DoIS, compreso il campo abstract, se presente e disponibile nei metadati creati.

Il servizio per ogni metadato ➤

assegna un identificatore permanente attraverso il sistema noto con il nome di *handle*. All'interno del servizio, ogni metadato presenta quindi un aggancio alla risorsa, utile alla sua localizzazione, che consente una certa garanzia di stabilità ai fini del recupero effettivo del documento nel tempo.

Il formato ReDIF per la rete di economia e, di conseguenza anche per DoIS, si basa sul protocollo Guildford.<sup>8</sup>

Il *Guildford protocol* specifica l'architettura e la gestione di un database distribuito su un insieme di siti locali. Ogni sito risiede ed è gestito da un singolo sistema informatico, in modo del tutto indipendente e autonomo, in un'area di memoria accessibile via anonymous FTP e/o http (*webserver*).

È attraverso l'utilizzo di un protocollo "comune" che avvengono gli scambi informativi e che si instaura un modello comunicativo tra le parti che si scambiano/comunicano informazioni. Nello scambio e trasporto di metadati l'adozione di un protocollo adeguato allo scopo del servizio che si vuole implementare risulta di fondamentale importanza. Il protocollo Guildford prevede una serie di azioni atte a regolare due processi comunicativi distinti ma correlati.

Il primo processo consente di raccogliere e rendere disponibili all'utente finale uno o più archivi, che possono contenere documenti a testo completo (*e-print server*), file di metadati (come nel caso di DoIS) e/o software di gestione, inseriti e mantenuti localmente.

Il secondo processo rende disponibili copie totali o parziali di archivi remoti, prevedendo regolare aggiornamento.

Sono stati sviluppati diversi pacchetti di software in linguaggio Perl per implementare questi processi, e anche DoIS ne utilizza alcuni. Uno dei fronti di prossimo sviluppo ed evoluzione di DoIS

sarà appunto in relazione allo sviluppo e implementazione di software confezionati per DoIS, piuttosto che riadattati sul modello RePEc per l'economia che presenta delle diversità in base alla diversa configurazione degli archivi.

Altro fronte assai interessante è una partecipazione attiva ed estesa a livello italiano che dovrebbe prevedere il coinvolgimento dei produttori di letteratura LIS, editori di riviste elettroniche del nostro settore, enti promotori di eventi (congressi, seminari, conferenze) al fine di rendere agevolmente disponibile quanto prodotto in Italia.

La disponibilità su web dei documenti a testo pieno, liberamente accessibili da un unico punto d'aggregazione, il servizio DoIS, potrebbe essere un'ottima opportunità per i bibliotecari italiani per più ragioni. Prima tra tutte rendere visibile la nostra produzione intellettuale all'estero, facendo confluire i metadati in DoIS internazionale. In secondo luogo offrirebbe ai bibliotecari uno strumento di raccolta dei lavori italiani, un contenitore LIS italiano con motore di ricerca che agisce sui metadati consentendo un immediato recupero dei documenti.

Senza troppo correre, va comunque detto che operare in un contesto di servizio esteso porta anche a migliorare rapporti e metodologie tra attori diversi del mercato dell'editoria elettronica (gratuita o commerciale). Per esempio DoIS potrebbe essere un'ottima opportunità per collaborazioni tra editori e biblioteche nella creazione e produzione di metadati. Altro aspetto importante, seppur marginale, la considerazione che il confronto può portare a migliorare la qualità dell'offerta informativa in generale, per esempio laddove al fine della creazione dei metadati per DoIS si rendesse auspicabile dotare gli articoli e i paper

di un abstract fornito dall'autore (come già avviene per alcune riviste italiane, per esempio il "Bollettino AIB").

L'aspetto a mio avviso più rilevante, e di diretto interesse per i bibliotecari, è quello di partecipare a un progetto che consentirebbe loro di toccare con mano i metadati, crearli, modificarli, interagire direttamente e comprendere, al di là di tante parole, il loro vero significato.

#### Note

<sup>1</sup> Il server si trova a Manchester: <<http://dois.mimas.ac.uk/>>.

<sup>2</sup> <<http://www.dlib.org/dlib/september00/09clips.html>>.

<sup>3</sup> <[ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/redif\\_1.html](ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/redif_1.html)>.

<sup>4</sup> Per la documentazione sul modello RePEc e sullo standard ReDIF: <<http://ideas.uqam.ca/ideas/data/rpcrdoc.html>>.

<sup>5</sup> RePEc fa parte del gruppo NetEc finanziato dal Joint Information Systems Committee (JISC) Higher Education Funding Councils, come parte dell'eLib-Electronic Libraries Programme, per gli anni 1996-1999, <<http://netec.mcc.ac.uk/>>.

<sup>6</sup> *The UPS prototype project: exploring the obstacles in creating a cross e-print archive end-user service*, with Herbert Van de Sompel, Michael L. Nelson and others, Old Dominion University Computer Science TR 2000-01, February, <[http://cs-tr.cs.cornell.edu:80/Dienst/UI/1.0/Display/ncstrl.odu\\_cs/TR\\_2000\\_01](http://cs-tr.cs.cornell.edu:80/Dienst/UI/1.0/Display/ncstrl.odu_cs/TR_2000_01)>.

<sup>7</sup> Resource organisation and discovery in subject-based services, <<http://www.roads.lut.ac.uk/>>.

<sup>8</sup> THOMAS KRICHEL, *Guildford protocol*, <<http://ideas.uqam.ca/ideas/data/Papers/rpcrdocGuildP.html>>.

