

# INTRODUZIONE

Il permesso di fare copie digitali o fisiche di tutto o parte di questo lavoro per uso di ricerca o didattico è acconsentito senza corrispettivo in danaro, mentre per altri usi o per inviare a server, ridistribuire a liste di discussione o diffondere ulteriormente è necessario il permesso da parte dell'autore.

L'utilizzo per scopi di profitto non è consentito senza il permesso dell'autore.

Gli eventuali lavori derivanti dallo stesso dovranno contenere opportuna citazione.

Il presente elaborato di tesi sperimentale prende in considerazione una realtà presente da qualche anno nel processo della comunicazione scientifica: gli Open Archive e l'iniziativa OAI (Open Archive Iniziative).

L'interesse per tale argomento è scaturito dalla lettura di un articolo, pubblicato sul Notiziario del SIMAI (Società Italiana di Matematica Applicata e Industriale) della dott.ssa De Robbio Antonella, Responsabile della biblioteca del CAB (Centro di Ateneo per le Biblioteche) presso l'Università degli Studi di Padova, nel quale venivano evidenziati alcuni problemi aperti relativi alla comunicazione dei risultati di ricerca all'interno della comunità scientifica accademica. Una soluzione praticabile per tali problematiche risulta essere l'adozione, da parte di Università ed Istituzioni in genere, dei sistemi documentari basati sugli Open Archive, che consentirebbero ai ricercatori di autoarchiviare e disseminare i propri lavori ovviando alla necessità di ricorrere al tradizionale circuito editoriale, consentendo, da un lato la rapida diffusione dei lavori scientifici, dall'altro di liberare i ricercatori dalle restrizioni dettate dalle politiche editoriali.

Tale elaborato, insieme a quello di Amelotti Ercole dal titolo "Protocollo OAI-PMH negli Open Archive e applicazione CDSware per la rappresentazione dei relativi dati bibliografici", prende in considerazione la gran parte degli aspetti riguardanti gli Open Archive.

In particolare questo elaborato puntualizza l'attenzione sui metadati e la loro importanza quale mezzo per la realizzazione della comunicazione scientifica su cui l'OAI ripone i propri sforzi.

Per quanto riguarda i metadati vengono particolarmente approfonditi il formato

Dublin Core, adottato come requisito base dall'OAI, e il formato MARC 21, uno dei pochi formati in grado di gestire dati altamente formalizzati, come ad esempio quelli dei cataloghi bibliografici.

Altro aspetto trattato, riguardante i metadati all'interno di tale iniziativa, è il loro formato XML di rappresentazione.

Importanza viene anche data ad uno dei soggetti attivi nel processo di comunicazione tra Open Archive: il Service Provider.

La parte sperimentale dell'elaborato, svolta su proposta dello SBA (Servizio Bibliotecario di Ateneo) dell'Università di Messina, consiste nell'applicazione del software CDSware (Cern Document Server Software) per la conversione dei record bibliografici importati dal sistema gestionale di biblioteca Aleph 500 in formato UNIMARC al formato XML MARC 21.

In particolare, la realizzazione di tale parte è focalizzata sull'utilizzo del modulo interno a CDSware: BibConvert, che permette la conversione dei record bibliografici prelevati da una sorgente esterna nel formato interno del sistema CDSware.

In concreto, si è realizzato il file di configurazione di BibConvert necessario per la conversione dei record bibliografici in formato UNIMARC, importati da Aleph 500, nel formato XML MARC 21 di CDSware. Tale applicazione è stata creata appositamente per la migrazione dei record immagazzinati dallo SBA in Aleph nella nuova installazione di CDSware da essi utilizzata.

Contributo fondamentale alla conoscenza del software CDSware è stata l'attività di stage formativo svolta presso il CERN di Ginevra, durante la quale si è avuta la possibilità di un confronto diretto con i capi progetto ed i membri di rilievo dello staff di sviluppo di CDSware.

La struttura della tesi è la seguente:

Il primo capitolo introduce gli Open Archive e l'iniziativa OAI correlata, prendendo in considerazione le problematiche relative alla comunicazione scientifica e la possibile soluzione attraverso gli Open Archive. Viene inoltre tracciato un quadro

storico relativo alla loro nascita ed evoluzione.

Il secondo capitolo ha per oggetto i metadati e dà una descrizione generale di cosa essi rappresentano all'interno del processo di comunicazione scientifica, le caratteristiche per referenziare le varie risorse e l'importanza del loro utilizzo entro l'iniziativa OAI.

Il terzo capitolo tratta il formato Dublin Core di rappresentazione dei metadati, adottato come formato base di interscambio tra Open Archive OAI compatibili. Vengono descritti la sua struttura e gli elementi di cui si compone, prendendo in considerazione sia la forma "Semplice" che "Qualificata".

Il quarto capitolo tratta il formato MARC 21 di rappresentazione dei metadati, più complesso e descrittivo del Dublin Core, adottato come formato interno dell'applicazione CDSware per gestire dati altamente formalizzati. Vengono analizzati e descritti la sua struttura e gli elementi di cui è composto.

Il quinto capitolo specifica la conversione dei dati bibliografici dal formato MARC 21 al Dublin Core e viceversa.

Il sesto capitolo presenta brevemente una panoramica sul linguaggio XML utilizzato per la rappresentazione dei record di metadati utilizzati da CDSware e di quelli forniti come risposta da parte del Data Provider al Service Provider.

Il settimo capitolo espone le linee guida dettate dall'OAI per la realizzazione di un Service Provider conforme all'iniziativa, specificandone i dettagli tecnici utili all'implementazione.

L'ottavo capitolo prende in esame l'applicazione CDSware descrivendone le caratteristiche generali, il suo utilizzo al CERN e i moduli funzionali di cui si

compone.

L'ultimo capitolo tratta in maniera approfondita il modulo BibConvert di CDSware ed in particolare del compito svolto su proposta dello SBA relativo alla parte sperimentale dell'elaborato di tesi.

*Questa tesi è stata sviluppata con uno stage presso il CERN di Ginevra supportato dai Fondi Europei per il corso di laurea in Informatica dell'Università di Messina.*

*Per la realizzazione del presente elaborato di tesi si coglie l'occasione per porgere i seguenti ringraziamenti a:*

*Prof. Pizzimenti Giovanni, Coordinatore del corso di Laurea in Informatica della Facoltà di Scienze.*

*Dott.ssa De Robbio Antonella, Responsabile della biblioteca del Seminario Matematico presso l'Università degli Studi di Padova, nonché correlatrice della tesi, per tutto il materiale bibliografico, i consigli ed i suggerimenti forniti.*

*Dott. Pettenati Corrado, Direttore della Biblioteca del CERN di Ginevra e lo staff di CDSware nelle persone di Jean-Yves Le Meur, Martin Vesely, Tibor Simko, Tullio Basaglia e Jens Vigen, per l'accoglienza e l'assistenza durante il periodo di stage.*

*Il Sistema Bibliotecario di Ateneo (SBA) di Messina, in particolare nelle persone della dott.ssa Alosi Benedetta e del dott. Feminò Nunzio, per la collaborazione avuta relativamente alla parte sperimentale*