

Modello dei dati

DAM.Solutions è l'applicativo per la gestione e la descrizione degli archivi storici e di deposito sviluppato dalla società Regesta.exe nell'ambito del progetto DAMS (Digital Archives and Memory Storage), sostenuto dal programma TEN-TELECOM (Trans European Networks for Telecommunications) dell'Unione europea. La caratteristica di spicco del prodotto è senz'altro l'elevata componente di ricerca e sperimentazione, che connota le potenzialità e i limiti dell'attuale struttura: non è stata ancora rilasciata, infatti, la versione definitiva dell'applicativo che, inserito all'interno di un progetto tuttora in corso, sta gradualmente raggiungendo gli obiettivi proposti, anche con il supporto degli enti sperimentatori.

La natura sperimentale del progetto ha spinto l'Istituto di studi per la tutela dei beni archivistici e librari (ISTBAL) dell'Università di Urbino 'Carlo Bo' ad assumere dal 2002 il ruolo di ente sperimentatore. Attualmente l'applicativo è utilizzato nell'ambito delle attività collegate al Laboratorio di descrizione archivistica attivato presso il corso di laurea in Scienze dei beni culturali, con l'obiettivo di valutare la coerenza scientifica del prodotto, sperimentare attraverso di esso le potenzialità descrittive della grammatica EAD e testare il modello dati proposto mediante il suo impiego nella descrizione di documentazione proveniente da archivi storici e di deposito locali oggetto di studio nell'ambito del Laboratorio stesso.

DAM nasce con l'obiettivo di realizzare la massima integrazione, interoperabilità e accesso alle risorse archivistiche gestite, e per tale motivo adotta una struttura tripartita basata su soluzioni tecnologiche versatili: un sistema di archiviazione e indicizzazione dei dati che poggia sul motore XML Extraway sviluppato dalla società 3D Informatica di Bologna, un applicativo sviluppato in ambiente Java, e un formato di rappresentazione dei dati in HTML.

Il recente sviluppo del motore di ricerca Extraway ha permesso di aderire alla filosofia XML senza mediazioni: i documenti nativi XML non sono importati all'interno di una base dati proprietaria, ma gestiti mediante l'impiego di un motore di ricerca che accede direttamente ai file. Ciò consente la piena interoperabilità della base dati e una sua elevata resistenza all'obsolescenza tecnologica, limitando il rischio di perdita di informazione insito in qualunque processo di migrazione dei dati da una piattaforma all'altra.

Il modello dei dati utilizzato è costituito da una selezione di elementi descrittivi tratti dalla *Encoded Archival Description* (EAD), lo standard di descrizione archivistica promosso dalla Society of American Archivists in collaborazione con il Network Development and MARC Standards Office

della Library of Congress e applicato ormai in molti paesi (USA, Canada, Australia, Africa del Sud, Taiwan, Gran Bretagna, Svezia, Norvegia, Germania, Portogallo, Spagna, Francia, Svizzera, Italia)¹.

EAD è tecnicamente una *document type definition* di XML per il dominio archivistico: come tale essa definisce gli elementi del modello dati, le reciproche relazioni, la loro ripetibilità ed i loro attributi, con il rigore caratteristico dei linguaggi formali. Tale standard è quindi estremamente adeguato come strategia di *comunicazione e scambio* di informazioni archivistiche, data la facilità di impiego in progetti che prevedano la diffusione dei contenuti sul web. Inoltre, la sua compatibilità con lo standard ISAD² ne fa uno strumento capace di descrivere e rappresentare adeguatamente strutture complesse di oggetti archivistici coerentemente con i principi enunciati dall'ICA, ma senza rinunciare alla ricchezza strutturale che EAD propone.

L'adozione di EAD nell'elaborazione di un applicativo per la descrizione archivistica rappresenta nel panorama italiano una scelta estremamente innovativa, condivisa al momento soltanto dai progetti di descrizione archivistica e di recupero retrospettivo di strumenti di corredo sviluppati dal Centro MAAS del Consorzio Roma Ricerche con l'impiego del medesimo motore di ricerca³.

Il modello dei dati estratto da EAD non è tuttavia un oggetto inalterabile: il nucleo di base può essere arricchito con ulteriori elementi della DTD al fine di realizzare una variante personalizzata e coerente con il progetto di descrizione cui si riferisce. La tecnologia XML facilita la gestione di schemi astratti dei dati che possono essere adattati con relativa facilità alle esigenze dell'utente e tale caratteristica sembra favorire le politiche di personalizzazione adottate dal produttore dell'applicativo, che non propone la commercializzazione di soluzioni predefinite.

L'interfaccia utente non mostra sensibili differenze rispetto ai comuni software per la descrizione archivistica: l'inserimento, la modifica e il recupero delle informazioni avvengono mediante la compilazione dei campi di una o più schede; tuttavia, mentre l'utente compila la schermata, l'applicativo traduce i dati inseriti in una pagina XML, associando ai dati i marcatori previsti dalla DTD EAD.

¹ Sulle origini e lo sviluppo di EAD si vedano le pagine curate dalla Library of Congress <<http://www.loc.gov/ead/>>.

² Cfr. il *mapping* tra EAD e ISAD elaborato nel 1996 da Sharon Thibodeau, membro della Commissione responsabile dell'elaborazione dello standard ISAD; in appendice alle *EAD Tag library* (2002) è consultabile una sua versione aggiornata.

³ Il lavoro di digitalizzazione della *Guida generale degli Archivi di Stato italiani*, effettuato dal centro MAAS mediante marcatura SGML/XML, è il primo e più significativo progetto italiano di applicazione dei linguaggi di marcatura ad oggetti archivistici. L'esperienza maturata ha consentito di utilizzare successivamente le stesse metodologie, unitamente all'adozione di EAD, all'interno di altri rilevanti progetti: la redazione della *Guida ai fondi dell'Istituto per la Ricostruzione Industriale*; il recupero retrospettivo della *Guida delle Camere di Commercio italiane* e della *Guida dell'Archivio Gramsci* per conto del Servizio V dell'Amministrazione archivistica, e il recupero degli inventari degli archivi storici comunali della provincia di Latina per conto della Regione Lazio.

L'applicativo offre la possibilità di descrivere oggetti archivistici di livello gerarchico diverso, proponendo un'articolazione in *fonds* (fondo e collezione), *subfonds* (subfondo), *series* (serie), *subseries* (sottoserie), *file* (unità archivistica) e *item* (unità documentaria), ma consente all'utente di selezionare e definire anche articolazioni diverse, mediante l'uso dell'attributo *otherlevel*. Tale scelta appare opportuna soprattutto in relazione alle esigenze descrittive che sorgono nel caso di fondi per i quali si voglia proporre, accanto alla strutturazione logica, anche una partizione di altro tipo (si pensi alle partizioni storico-istituzionali presenti in strumenti quali la *Guida generale degli archivi di Stato italiani*, o a quelle tipiche dei fondi archivistici comunali). Non sono previste limitazioni quantitative alla possibilità di annidamento di oggetti archivistici all'interno di altri. Il livello di descrizione è il primo elemento che il sistema richiede all'utente di selezionare.

I curatori del modello dati hanno ritenuto opportuno distinguere due diversi *set* descrittivi in base alla natura degli oggetti da trattare: i livelli fondo, subfondo, serie, sottoserie e *otherlevel* presentano una struttura più ricca, composta dagli elementi

1. *<did>* (informazioni relative all'identificazione e al contesto dell'unità di descrizione), strutturato nei sottoelementi:
 - *<unitid>* (segnatura/e o codice/i identificativo/i)
 - *<unititle>* (denominazione o titolo)
 - *<unitdate>* (data/e), elemento all'interno del quale è possibile esprimere la data cronica, topica ed eventuali note alla data
 - *<origination>* (denominazione del/dei soggetto/i produttori, descritti secondo lo standard ISAAR; le informazioni inserite incrementano l'*authority file*)⁴
 - *<repository>* (istituzione o ente responsabile dell'accesso intellettuale)⁵
 - *<abstract>* (sommara descrizione dell'unità e del suo produttore)
 - *<physdesc>* (descrizione fisica)
 - *<container>* (numero di corda)
 - *<note>* per l'indicazione degli allegati
2. *<descgrp>* (insieme di elementi descrittivi) strutturato in:
 - *<accessrestrict>* (condizioni di accesso)
 - *<acqinfo>* (modalità di acquisizione e versamento)
 - *<bibliography>* (bibliografia)
 - *<custodhist>* (storia della custodia del materiale)
 - *<otherfindaid>* (riferimenti ad altri strumenti di ricerca)
 - *<phystech>* (caratteristiche materiali e requisiti tecnici)⁶

⁴ Attualmente presente nella versione ISTBAL.

⁵ Attualmente presente nella versione ISTBAL.

- *<relatedmaterial>* (riferimenti a materiale correlato)
 - *<separatedmaterial>* (riferimenti a materiale avente la stessa provenienza ma conservato altrove)⁷
 - *<appraisal>* (modalità di selezione e scarto)⁸
 - *<altformavail>* (esistenza e localizzazione di copie)⁹
3. *<bioghist>* (storia istituzionale/amministrativa, nota biografica)
 4. *<scopecontent>* (ambiti e contenuto)
 5. *<controlaccess>* (chiavi d'accesso controllate che implementano l'*authority file* e offrono tra l'altro la possibilità di descrivere gli enti che compaiono nella descrizione), in cui si distinguono i sottoelementi:
 - *<corpname>*, per le forme normalizzate dei nomi di enti
 - *<persname>*, per le forme normalizzate dei nomi di persona
 - *<geogname>*, per i toponimi.

Dalla sezione dedicata all'elemento *<controlaccess>* si accede alla scheda descrittiva per enti e persone, in fase di sviluppo al fine di renderla coerente con lo standard ISAAR(CPF).

6. *<dao>* (oggetto archivistico digitale), con il quale è possibile definire: la destinazione ad un'utenza interna o esterna all'ente (*audience*), il titolo (*title*) e l'indirizzo del collegamento ipertestuale (*href*); e *<daogrp>* (gruppo di oggetti archivistici digitali), contenente informazioni relative all'identificativo del gruppo (*id*), all'utenza (*audience*), alla descrizione degli oggetti archivistici digitali (*<daodesc>*) e ai collegamenti ipertestuali (*<daoloc>*). I due elementi *<dao>* e *<daogrp>* consentono la gestione di oggetti digitali associati alla descrizione archivistica, come ad esempio immagini digitalizzate dei documenti (attualmente il sistema supporta file con estensione jpg).

Per gli altri livelli descrittivi (unità archivistica e unità documentaria) è previsto un set di elementi ridotto, costituito da *<did>*, *<scopecontent>*, *<controlaccess>* e *<dao>*. Tale scelta, ancorché giustificabile dalla volontà di alleggerire la fase di inserimento dei dati, non convince per la specifica soluzione adottata che impedisce di descrivere adeguatamente alcune tipologie di documenti.

⁶ Si noti che *<physdesc>* è utilizzato per descrivere in generale le caratteristiche fisiche dei materiali (consistenza, supporto, dimensioni, colore, aspetto fisico, ecc.) mentre *<phystech>* si riferisce alle caratteristiche rilevanti ai fini dell'utilizzo dei materiali descritti.

⁷ Si noti che *<relatedmaterial>* si riferisce a fonti collegate a quelle descritte sulla base di un qualunque tipo di associazione logica, mentre l'elemento *<separatedmaterial>* si riferisce a fonti aventi la medesima provenienza ma conservate altrove.

⁸ Attualmente presente nella versione ISTBAL.

⁹ Attualmente presente nella versione ISTBAL.

A ciascuno dei sei elementi di primo livello sopra descritti corrisponde una schermata, la cui compilazione può essere effettuata inserendo i dati nei campi opportuni.

La funzione di inserimento di una nuova unità di descrizione (*component*) prevede anche l'esplicitazione dei suoi legami con gli altri elementi dell'albero gerarchico: l'utente può definire il livello, di elemento paritario o subordinato, che esso assume rispetto all'oggetto descrittivo al quale lo si sta collegando, ovvero può riservarsi di stabilire successivamente le relazioni gerarchiche tra gli elementi.

La descrizione viene completata al momento del salvataggio dei dati con la registrazione in automatico, mediante la rilevazione della password, del nome del compilatore della scheda descrittiva, della data e del tipo di compilazione (inserimento o modifica) effettuata.

L'applicativo non è stato progettato per una specifica tipologia di documentazione ed è attualmente impiegato per la descrizione di fondi archivistici di natura diversa, conservati in archivi di Stato, di organi istituzionali, di enti pubblici e privati.

Funzioni per l'utente/operatore

Sono previsti quattro profili di utente, con un grado crescente di interazione con l'applicativo: il *guest*, ovvero l'utente occasionale che dispone unicamente della possibilità di consultare il sistema; l'*operatore*, responsabile dell'inserimento, dell'eliminazione e della modifica dei dati; il *gestore*, incaricato del mantenimento e della gestione della banca dati archivistica, al quale è consentita anche la rinumerazione delle schede nel caso in cui sia necessario cancellare o spostare una o più unità di descrizione; e l'*amministratore*, che dispone di tutte le funzionalità previste dal sistema e associa agli utenti i rispettivi profili di accesso.

L'utente che accede al sistema come operatore può modificare in ogni momento l'ordinamento di tutto o di una parte dell'albero gerarchico elaborato, utilizzando se opportuno la procedura di riordinamento automatico, applicabile a tutti i campi; al termine delle operazioni di riordinamento l'utente con profilo di gestore può attivare la funzione automatica di rinumerazione sequenziale delle schede e conferire alla struttura archivistica l'assetto definitivo.

E' consentita una gamma diversificata di formati di stampa, che possono essere personalizzati selezionando gli elementi e le caratteristiche tipografiche desiderate; l'opzione offre la possibilità di generare differenti strumenti di corredo, caratterizzati da un diverso grado di analiticità e da un diverso formato grafico. Inoltre, per tutti gli elementi del modello, è prevista la stampa di indici, personalizzabili su richiesta del committente.

La funzione di ricerca prevede sia la modalità a testo libero (con ricerca su tutti gli elementi del modello) sia quella avanzata, che consente di compiere interrogazioni su alcuni elementi combinabili tra di loro (la denominazione, l'ambito e contenuto, i nomi di enti e di persone, gli

oggetti digitali collegati); nel caso in cui si abbiano a disposizione più banche dati è inoltre possibile selezionare quelle su cui effettuare la ricerca impostata.

Le procedure di import-export sono agevolate dall'uso di XML per la costituzione del *repository* documentale: la scelta garantisce il recupero completo di informazione sia nel caso di procedure di aggiornamento legate all'evoluzione della piattaforma sia nell'ipotesi di una migrazione verso un altro sistema.

Come si dirà più diffusamente in seguito, la diffusione su web costituisce uno degli elementi fondanti dell'applicativo, la cui presentazione dei dati poggia sul protocollo HTTP.

La compilazione e la consultazione della base dati sono agevolate da un manuale di istruzioni attualmente disponibile in lingua inglese e, a breve, anche in italiano, personalizzato in base alla versione dell'applicativo concordata con il committente; è inoltre disponibile una funzione di *help* multilingue compilato in XML, attivo durante la consultazione, l'inserimento e la modifica dei dati, che coadiuva le attività dell'operatore grazie alla visualizzazione di *balloon* esplicativi per i singoli elementi descrittivi.

In merito alla qualità dell'assistenza e della manutenzione del software si è attualmente in una fase di piena sperimentazione e dunque l'attenzione del personale tecnico che cura lo sviluppo e l'ottimizzazione del prodotto verso le osservazioni e le richieste degli enti che testano l'applicativo è estremamente sollecita.

Infine, un paio di osservazioni per completare l'illustrazione dell'applicativo.

La prima riguarda le limitazioni nel grado di dettaglio e nella libertà di scelta degli elementi mediante i quali descrivere una fonte archivistica: il modello dati DAM rappresenta un *subset* di elementi EAD e dunque non offre né l'ampiezza del vocabolario che è alla base di tale grammatica, né le sue infinite combinatorie. Come già detto, le personalizzazioni del prodotto consentono sì di scegliere il modello più adeguato in ragione di specifiche esigenze descrittive, ma tale strategia impone una ridefinizione e un nuovo intervento sull'applicativo qualora il modello adottato non risulti soddisfacente per la descrizione di altri materiali. Occorre sottolineare che un approfondimento dell'analisi del prodotto può consentire in breve termine l'adozione di un modello dati caratterizzato da una *granularità* adeguata per un ampio ventaglio di situazioni, ma è evidente che la possibilità di disporre *tout court* dell'intera e complessa struttura di EAD costituirebbe un potente e vantaggioso strumento nelle mani di un utente esperto: non bisogna del resto trascurare il fatto che la scelta di adottare schermate con le "classiche" schede a campi predefiniti contempera da una parte la necessità di far evolvere l'applicativo verso soluzioni tecnologicamente innovative e dall'altra le esigenze di una comunità di utenti caratterizzata da una conoscenza non ancora matura dello standard EAD e dotata di un patrimonio tecnico informatico non approfondito. E' comunque allo studio l'ipotesi di creare una versione di DAM destinata ad un'utenza avanzata che consenta di

fruire di tutti gli elementi della grammatica EAD e permetta di compilare direttamente una pagina XML secondo modalità simili a quelle di un *editor XML*.

La seconda osservazione scaturisce dall'impiego dell'applicativo nelle operazioni di schedatura seriale: la possibilità di evitare la digitazione dei medesimi dati riferiti a unità archivistiche diverse costituirebbe un innegabile vantaggio per lo snellimento del lavoro ma, nonostante esista una funzione di 'taglia-incolla' per porzioni di testo, al momento è impossibile duplicare le schede descrittive e modificarne successivamente alcuni dati.

Considerazioni complessive

L'introduzione di software per la descrizione archivistica ha di fatto modificato nel corso dell'ultimo decennio il lavoro degli archivisti, inducendo la sovrapposizione delle distinte attività di schedatura, riordinamento e descrizione del fondo: i vincoli posti dagli applicativi spesso inducono ad operare un riordinamento di massima prima di procedere alla schedatura del materiale, ovvero ad effettuare una schedatura cartacea prima di provvedere alla descrizione su supporto informatico per ovviare a lunghe procedure di riorganizzazione dei dati o di rinumerazione delle unità archivistiche. In generale, la flessibilità e la versatilità dell'applicativo agiscono in misura inversamente proporzionale sulla necessità dell'archivista di alterare la metodologia di lavoro consolidata: DAM appare, all'analisi e all'uso diretto, un prodotto maneggevole e duttile. All'archivista è consentito inoltre di lavorare in una condizione di estrema libertà di movimento, potendo gestire in momenti diversi l'attività di schedatura del fondo, grazie alla possibilità di abbozzare la struttura descrittiva e di incrementarla e modificarla *in itinere*, e quella di riordinamento logico, agevolato da operazioni di riordinamento automatiche. Il sistema non prevede alcun tipo di prescrittività nella compilazione della descrizione archivistica (fatta eccezione per la denominazione dell'oggetto descritto), garantisce all'utente la possibilità di intervento sulle singole schede descrittive in ogni momento del lavoro e permette la redazione e la stampa di molteplici strumenti di corredo. A questo proposito appare estremamente interessante la gamma di scelte a disposizione: da una descrizione analitica si possono infatti generare in fase di stampa strumenti diversi, più o meno dettagliati, mediante l'adozione di formati diversificati.

Un'ultima osservazione a proposito delle due modalità di acquisto e di utilizzo di DAM: il prodotto è disponibile mediante la sua installazione su una piattaforma di rete oppure mediante la concessione di una licenza d'uso via web. La seconda opzione offre all'utente la possibilità di implementare la propria base dati da remoto, mediante la compilazione di pagine *asp*, con un notevole risparmio sui costi di hardware e software che si riducono in tal modo alla sola postazione abilitata all'accesso internet, mediante la quale si interviene sul contenuto della base dati ospitata sul server della società Regesta. L'introduzione di procedure di redazione e manutenzione di strumenti

descrittivi da remoto segna senza dubbio una rivoluzione nelle abituali logiche di gestione delle attività di riordinamento e descrizione archivistica. Da una parte, infatti, rende estremamente flessibile il lavoro dell'archivista, svincolandolo dalla variabile del tempo e del luogo e consentendo a più operatori di intervenire sulla stessa base dati contemporaneamente da luoghi diversi, oppure permettendo ad un operatore di effettuare alcune attività (riordinamento logico, perfezionamento della descrizione, redazione dell'*authority file*) anche da un luogo diverso da quello in cui è fisicamente conservato l'archivio; dall'altra implica la necessità di confrontarsi con il tema della esternalizzazione completa dell'accesso e della gestione delle *metarisorse* archivistiche, tema in parte già affrontato in ambito biblioteconomico con l'introduzione di analoghi strumenti per la descrizione e la gestione di risorse bibliografiche in rete, che comporta inevitabilmente il ripensamento e l'individuazione di nuove modalità di condivisione di responsabilità e competenze finora interamente affidate al ruolo dell'archivista.