

Il nuovo database di DFP

Il gruppo di lavoro dell'AIB sulla documentazione di fonte pubblica nel 2004 ha trasformato la conoscenza accumulata in anni di lavoro sui siti degli enti pubblici italiani in una banca dati nella quale da allora sono state segnalate circa 1.300 risorse documentarie di varia natura. In parte le segnalazioni sono state derivate da quanto contenuto nel volume *Documenti e dati pubblici sul web*¹ pubblicato lo stesso anno.

La banca dati è stata, a suo tempo, realizzata utilizzando il software Scout Portal Toolkit (SPT) prodotto dal Internet Scout Project e messo a disposizione con licenza GPL² (GNU General Public License). La realizzazione della banca dati DFP si è costantemente avvalsa della preziosa collaborazione di CILEA che ha messo a disposizione non solo le competenze informatiche per l'installazione dei programmi, ma anche la traduzione in italiano dell'interfaccia utente di SPT.

Dopo cinque anni di onorato servizio il software SPT ha cominciato a denunciare problemi sia per il calo delle prestazioni sia per l'evidenziarsi di alcune falle nella sicurezza.

La necessità di reinstallare il software ha spinto i responsabili tecnici di AIB-WEB e i membri della redazione di DFP a verificare la possibilità di utilizzare le versioni successive di SPT.

Il progetto Internet Scout in questi anni si è evoluto in varie direzioni. In particolare il software per la gestione di directory, SPT, è stato completamente rivisto, tanto che

ha ricevuto un nuovo nome: CWIS.³

Quello che andremo qui a descrivere è il progetto per l'installazione, la traduzione e la personalizzazione di CWIS per la banca dati DFP.

Inoltre descriveremo il lavoro svolto per trasferire i dati dalla vecchia alla nuova banca dati e per conservare gli identificativi delle risorse e delle pagine della classificazione.

Internet Scout Project e CWIS

Avviato nel 1994 presso il campus di Madison della University of Wisconsin, lo Scout Project ha come scopo quello di produrre strumenti sempre migliori per la ricerca, la selezione e la presentazione online di informazioni e metadata.

Scout Project, che è parte dell'University's College of Letters and Sciences, può avvalersi di ottimi specialisti e di una disponibilità di hardware e di risorse bibliotecarie di primo livello.

Lo staff che collabora a Scout Project rappresenta un ori-

ginale mix di docenti, ricercatori e professionisti provenienti dalle discipline bibliotecarie e informatiche.

Al progetto collaborano anche studenti e studenti laureati impegnati in studi scientifici, di scienze sociali e di scienze umane.

La principale pubblicazione del progetto è lo "Scout report" che dal settembre 1994 segnala le "new and newly-discovered online resources, network news, and bits of net culture".

Il lavoro in questi quindici anni è proseguito sulla base di progetti di ricerca finanziati da organismi pubblici e fondazioni, in particolare la National Science Foundation. Attualmente sono attive queste ricerche: AMSER, Access NSDL, ATE Central, Learning-Languages.net, CWIS.

Le prime quattro ricerche hanno come scopo la produzione di directory specializzate rispettivamente in: supporti educativi per la matematica applicata e le scienze, accessibilità dei siti web, strumenti educativi avanzati e risorse per l'apprendimento delle lingue.

L'ultimo progetto, CWIS - Collection Workflow Integration System, si occupa di sviluppare il software per il mantenimento di directory ed è utilizzato negli altri tre progetti. CWIS è quindi il cuore dell'attività dello Internet

Scout Project ed è l'erede del già citato SPT.

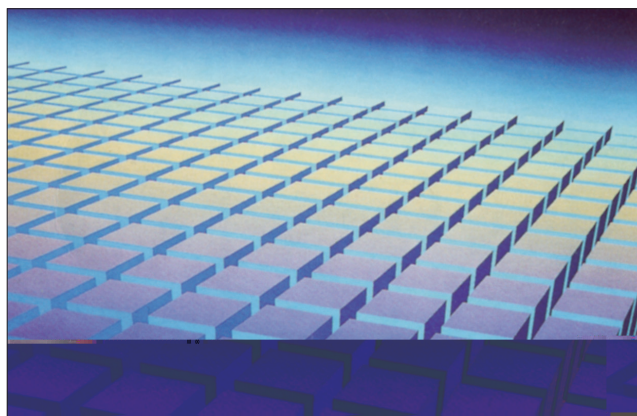
CWIS

CWIS (da pronunciarsi *si-uìs*) è un software per la raccolta, l'organizzazione, il mantenimento e la pubblicazione di collezioni di descrizioni di risorse documentali distribuite gratuitamente in base alla licenza GPL (GNU General Public License).⁴

Caratteristica specifica di CWIS è che rispetta gli standard più apprezzati dai professionisti della gestione delle informazioni per quanto riguarda i metadata. Le caratteristiche di CWIS sono:

- annotazione e votazione delle risorse (tipo Amazon);
- ricerca per parole chiave;
- ricerca su campi specifici funzionali per ottenere raccomandazioni in base al proprio profilo;
- esportazione dei metadata in base al protocollo OAI 2.0 (uso degli schema *oai_dc* e *nsdl_dc*);
- supporto per i feed RSS;
- funzione integrata per la modifica e l'aggiunta dei metadata;
- schema dei metadata definibile dall'amministratore (il sistema offre di base l'intero schema Dublin Core⁵ con i qualificatori);
- schemi terminologici pre-strutturati (include il GEM Subject⁶);
- template diversi per l'interfaccia utente;
- funzione di installazione "chiavi in mano".

Strutturalmente CWIS si presenta con la parte funzionale, scritta in PHP, ben separata dai template, scritti in HTML, utilizzati per la visualizzazione. Questo rende più semplice la personalizzazione delle singole installazioni e la manutenzione del software.



L'installazione di CWIS per DFP

Per scaricare CWIS basta collegarsi con il sito di Internet Scout Project, <<http://scout.wisc.edu/index.php>>, scegliere l'apposita voce nel menù al centro della pagina e poi, ancora, scegliere download nella barra delle attività offerta sulla pagine del sotto-progetto.

Installare CWIS è relativamente semplice. Basta disporre di un sito che possa eseguire programmi in PHP e da cui sia possibile accedere a un database MySQL.

La scelta dello staff di AIB-WEB è stata di utilizzare l'ultima versione stabile, la 1.5.2 rilasciata il 29 maggio del 2009. Ora è disponibile anche la release Alpha della versione in sviluppo, 1.9.2. Per utilizzare CWIS è necessario disporre del seguente software:

- Linux web server
- PHP 4.0.6
- o versioni successive
- MySQL 3.23.29
- o versioni successive

Per l'installazione è necessario avere i diritti di accesso attraverso command-line.

CWIS può essere utilizzato anche su piattaforme diverse da Linux, per esempio server Windows, ma il gruppo di sviluppo non supporta tali installazioni.

Dopo aver copiato i file nella directory del server prescelta, assegnati i grant necessari e creato il database in cui si dovranno collocare le tabelle, l'installazione vera e propria avviene attraverso un programma cui si accede via web.

Il programma provvede a creare tutti i file e le tabelle del database senza bisogno di alcun ulteriore intervento. Il risultato dell'installazione è

una banca dati esemplificativa contenente 31 risorse e una classificazione con 170 voci.

L'eliminazione di questa banca dati di prova può essere effettuata dall'amministratore della banca dati via web. L'installazione è stata seguita dai tecnici CILEA (Daniela Dragone e Anna Marini) insieme ad Andrea Marchitelli, che ha seguito l'attività come membro del gruppo di lavoro CILEA e come coordinatore di AIB-WEB.

La personalizzazione

Una volta installato il software, la Redazione DFP si è confrontata con il gruppo di AIB-WEB su che cosa fosse opportuno fare riguardo alla personalizzazione e alla traduzione in italiano dell'interfaccia utente. La decisione è stata di modificare l'interfaccia di CWIS per renderla specifica di DFP e di tradurre l'interfaccia utente. CWIS è strutturato in modo tale da separare i file necessari per la visualizzazione da quelli che interagiscono con i dati e che determinano il comportamento del programma. Questa caratteristica ha reso più semplici entrambe le operazioni.

In concreto si è trattato di procedere alla modifica dei file che vengono inclusi in tutte le pagine in modo da inserire il logo DFP e spostare il menù principale a sinistra. Inoltre è stata creata una nuova pagina in cui compaiono le ultime risorse inserite, mentre nell'homepage della banca dati sono state inserite le segnalazioni delle tabelle sinottiche delle risorse legislative e parlamentari italiane online.

La modifica di altre caratteristiche è stata ottenuta agendo sul CSS.

La traduzione si è svolta in due fasi. Nella prima fase due persone hanno provveduto ad individuare i file coinvolti e a tradurre i testi che in questi file erano inseriti tra il codice HTML e alcuni indispensabili comandi PHP.

Una volta completate le operazioni, la Redazione DFP ha provveduto a rivedere l'interfaccia utente per individuare eventuali dimenticanze, la coerenza della traduzione con il contesto e, soprattutto, per migliorare lo stile della traduzione stessa. In contemporanea i tecnici di CILEA hanno provveduto a riportare sul nuovo programma le modifiche migliorative che erano state apportate a quello precedente: visualizzazione dell'intero albero della classificazione, sistema per la verifica dei link e invio di una e-mail agli utenti registrati per segnalare l'inserimento di nuove risorse. Il completamento della modifica del layout, della traduzione e delle aggiunte è avvenuto all'inizio di settembre ed è coinciso con la chiusura del vecchio database.

Il trasferimento dei dati e la messa in linea

Il mese di settembre è stato dedicato al trasferimento dei dati delle risorse e degli utenti registrati.

Si è scelto di fare in modo che il trasferimento avvenisse senza cambiamenti degli identificativi delle singole risorse, in modo che le URL delle descrizioni potessero rimanere permanenti.

Il contributo di Eugenio Gatto ha permesso di individuare tutte le lacune nella sequenza degli identificativi delle risorse, così che nel file da caricare nel nuovo database si potessero inserire nei punti appropriati dei record dum-

my che in seguito sono stati eliminati.

Un intervento analogo è stato fatto per quanto riguarda gli identificativi delle singole classi della classificazione. Questo intervento, visto che le classi vengono create in fase di caricamento delle risorse, ha richiesto delle modifiche dirette dei record delle tabelle del database.

Il riversamento dei dati degli utenti registrati è avvenuto in modo da conservare tutte le informazioni inserite dagli stessi, comprese le loro password.

I dati relativi agli utenti registrati hanno richiesto un intervento di pulizia perché negli anni si erano accumulati molti inserimenti effettuati da hacker che in alcuni casi interferivano con il buon funzionamento del programma.

Complessivamente sono state trasferite 1.292 risorse, 185 classi e 457 utenti.

Il 1° ottobre 2009 è stata resa pubblica la nuova banca dati di DFP, che mantiene lo stesso indirizzo: <<http://dfp.aib.it/>>.

Dal momento del rilascio si sono registrati 351 nuovi utenti.

L'installazione di CWIS per DFP utilizza alcune delle nuove funzioni disponibili, mentre si è ritenuto di non abilitare i forum perché al momento i redattori non sono in grado di gestirli.

Le principali novità rispetto alla precedente banca dati sono le funzionalità per l'assegnazione di punteggi alle risorse da parte degli utenti, l'inserimento di commenti e il suggerimento di risorse.

Queste tre funzioni sono collegate tra loro e introducono DFP nella realtà del web 2.0, o meglio noi speriamo che consentano di avviare DFP alla creazione di una comunità di utenti che con i

loro voti e i loro commenti consentano un arricchimento della banca dati che si tradurrà nella possibilità per gli altri utenti di ottenere utili indicazioni su quali risorse possano risolvere i loro problemi informativi.

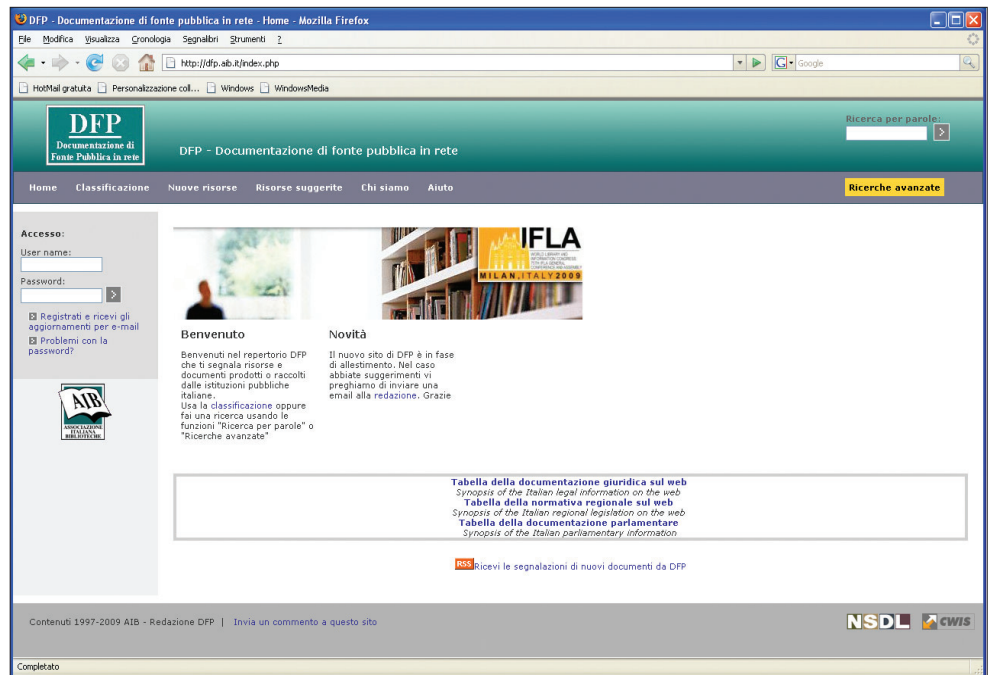
Non nascondiamo che la scelta di attivare queste funzioni vuole anche rappresentare un esperimento per verificare quanto sia reale la disponibilità degli utenti a registrarsi, votare e commentare le risorse.

I suggerimenti che un utente riceve dipendono da quante risorse ha valutato (rating) e dai punteggi che ha assegnato. Il sistema compie delle inferenze partendo da queste valutazioni, dalle valutazioni di altri utenti e dalle parole contenute nei record delle risorse. Se un utente non fornisce al sistema informazioni attraverso l'attività di valutazione non potrà ottenere alcun suggerimento. La valutazione delle risorse è riservata agli utenti registrati ed è attivabile sulla scheda completa delle risorse stesse.

Conclusione

L'aggiornamento del software di DFP è stato sicuramente un processo relativamente semplice sia dal punto di vista tecnico, grazie alle solide competenze dei tecnici CILEA, sia per i membri della redazione, grazie all'esperienza ormai quinquennale con Scout, che non hanno dovuto modificare in modo sostanziale le proprie modalità di inserimento e aggiornamento dei dati.

Certamente la semplicità con cui queste operazioni sono state svolte non deve ingenerare l'idea che sia facile installare e configurare CWIS partendo da zero. Come tut-



ti software Open Source, richiede che l'ambiente di installazione sia perfettamente conosciuto da chi lo gestisce, che tutto il software di base e complementare sia perfettamente configurato, che si conoscano bene i linguaggi con cui è realizzato e la struttura dei dati del database che fa da background all'applicazione.

In sintesi, per usare CWIS come tutti i software di questo tipo bisogna poter contare su persone dotate di elevata professionalità.

Nel nostro caso, visto che le risorse professionali erano disponibili, si è rilevato più complesso il trasferimento dei dati, in particolare per la scelta del gruppo di lavoro di mantenere le URI sia delle risorse sia delle classi.

La collaborazione tra tutti i partecipanti al progetto – Redazione DFP, redazione AIB-WEB e tecnici del CILEA – è stata ottima, tanto che l'intera operazione si è potuta svolgere comunicando solo via e-mail.

L'essere riusciti a portare a termine una operazione co-

si significativa come la messa in linea di una nuova banca dati in un tempo relativamente breve, salvaguardando i dati e i loro identificativi, pensiamo dimostri la professionalità di tutti coloro che collaborano con AIB-WEB e DFP e, soprattutto, che i veri ostacoli per la realizzazione di progetti significativi in campo bibliografico non derivano sicuramente dalla mancanza di capacità tecnologiche e organizzative, ma dalla scarsità delle risorse disponibili per l'individuazione, la descrizione, il commento, l'indicizzazione delle risorse.

Questa scarsità di risorse per DFP si traduce nel doversi basare esclusivamente sul lavoro volontario di bibliotecari e documentalisti, a scapito della possibilità di inserire un numero più alto di risorse e di migliorare la qualità dei record prodotti.

Nonostante ciò la redazione DFP vuole continuare a mettere a disposizione una banca dati aggiornata che documenti al meglio le risorse informative e documentarie mes-

se a disposizione della collettività dalle pubbliche amministrazioni italiane.

Piero Cavaleri

Biblioteca "Rostoni"
Università Carlo Cattaneo LIUC
pcavaleri@liuc.it

Note

¹ Documenti e dati pubblici sul Web. Guida all'informazione di fonte pubblica in rete, a cura di Piero Cavaleri e Fernando Venturini, contributi di Laura Ballestra [et al.], Bologna, Il Mulino, 2004.

² GPL, <<http://www.gnu.org/licenses/gpl.html>>.

³ Altri progetti italiani che utilizzano CWIS, a nostra conoscenza, sono: Bibliostoria Web Resources (<http://users.unimi.it/storia/h/SPT--Home.php>) e Preferito (<http://library.unito.it/preferito/SPT--Home.php>).

⁴ GNU General Public License, <<http://www.gnu.org/licenses/gpl.html>>, consultato il 2 novembre 2009.

⁵ Dublin Core, <<http://www.dublincore.org/documents/dcmi-terms/>>.

⁶ GEM controlled vocabulary, <<http://www.thegateway.org/about/documentation2/gem-controlled-vocabularies/vocabulary-subject>>.