



Università degli Studi di Siena
Area Sistema delle Biblioteche
Servizio Automazione Biblioteche
SBS- Servizio Bibliotecario Senese

Via Banchi di Sotto, 55 53100 Siena Tel. Fax.+39 0577 232430 e-mail sbs@unisi.it

IL PROGETTO ITALE-ACNP
a cura di Stefano Bonanni
Servizio Automazione Biblioteche
AREA SISTEMA DELLE BIBLIOTECHE

INTRODUZIONE

Questa relazione si colloca a conclusione di un percorso iniziale – seppure dilatato nel tempo ben oltre le nostre previsioni – che fortunatamente ci ha consentito di conseguire alcuni risultati concreti. E' pertanto occasione per una sorta di riepilogo generale su quello che è stato fatto, su come è stato fatto e su quello – ancora tanto – che rimane da fare. La relazione è stata concepita per essere sintetica, anche perché gran parte degli argomenti (analisi del tracciato ACNP, flusso di lavoro, normalizzazione delle stringhe di possesso e lacune, etc.) sono già stati ampiamente e singolarmente trattati in precedenti documenti e il rischio – che in tutta onestà non so neppure quanto sia riuscito a scongiurare – era obiettivamente quello di ripetersi.

OBIETTIVI DEL PROGETTO

L'obiettivo del progetto consiste nel caricamento *batch* del possesso (consistenze, lacune, collocazioni, fondi) dei seriali cartacei in ACNP, per rendere questi dati disponibili all'interno di uno dei cataloghi collettivi più conosciuti e utilizzati in Italia senza costringere i bibliotecari a gestirli (catalogazione, gestione amministrativa dei fascicoli, etc.) su due piattaforme diverse, Aleph e ACNP.

In concreto l'operazione si propone di dare maggiore impulso ai servizi di documentazione delle Università, alla cooperazione interbibliotecaria, alle funzionalità del sistema Nilde (Network Inter-Library Document Exchange) al quale ad es. partecipano la maggior parte delle biblioteche aderenti

ad ITALE. Oltretutto, proprio per questo, esse sono tenute entro dicembre 2008 a far conoscere il proprio catalogo dei periodici aderendo ad uno dei cataloghi collettivi nazionali (SBN, ACNP).

ORGANIZZAZIONE

Il progetto, la cui genesi risale ormai alla fine del 2002, è stato tecnicamente ed organizzativamente coordinato dal Sistema Bibliotecario dell'Università di Siena, anche quando dal novembre 2004 altri Atenei aderenti alla associazione ITALE hanno manifestato il loro interesse al caricamento *batch*. E' stato così costituito uno specifico gruppo di lavoro ITALE, il *feedback* all'interno del quale si è svolto prevalentemente per posta elettronica. E' stato fatto sostanzialmente un lavoro di analisi dei vari cataloghi, mirato soprattutto a quantificare i record privi di ISSN e a comparare le strutture delle consistenze dei periodici. E' emerso un quadro eterogeneo – e la cosa non ci sorprende affatto se consideriamo il livello di disomogeneità riscontrato negli stessi singoli cataloghi –, ma è anche vero che in alcune realtà è stato approntato un protocollo per l'inserimento standardizzato delle consistenze. Una soluzione, questa, apprezzabile, che in prospettiva potrebbe semplificare e migliorare notevolmente tutto il nostro lavoro.

In sintesi l'Università di Siena, forte sia dei suggerimenti e delle proposte avanzate in seno al gruppo, sia dell'esperienza e del lavoro preparatorio accumulati negli anni precedenti, ha svolto una delicata e difficile funzione di apripista, definendo modalità e strumenti flessibili per la migrazione dei dati. Durante questa attività ha coordinato il rapporto con i gestori di ACNP in modo da evitare sovrapposizioni ed equivoci, cosa che consentirà poi agli altri Atenei – completate le indispensabili operazioni preliminari – di essere nelle condizioni di potersi avvalere di questi stessi strumenti adeguatamente parametrizzati.

LA CORNICE “TECNICA”

La complessità delle operazioni richieste dalla migrazione automatica si scontra paradossalmente con la semplicità del contesto tecnico e concettuale entro il quale ci siamo dovuti muovere. Esso è ben esemplificato dal formato ACNP che, trasmessoci dai gestori del catalogo, stabilisce rigorosamente i dati che devono essere caricati, la loro successione, la lunghezza dei campi, la punteggiatura, etc. Anche ad un suo esame superficiale è evidente che il caricamento *batch* avrebbe riguardato solo ed esclusivamente record di possesso, non le informazioni bibliografiche e che il “gancio” tra gli uni e le altre sarebbe stato il codice ISSN o in mancanza di questo il codice CNR. L'oggetto del caricamento poi sarebbero stati tanti files – creati secondo questo tracciato – quanti

sono le biblioteche dell'Università coinvolte nel progetto, ciascuna con un proprio codice identificativo ACNP.

I sistemisti di ACNP – per completare la descrizione del quadro di riferimento – hanno poi esplicitamente richiesto che la frequenza di questo caricamento *batch* fosse annuale e hanno sottolineato poi come sarebbe stato opportuno che le biblioteche gestissero *on line*, preferibilmente in forma centralizzata, almeno l'inserimento dei titoli privi di identificativo nonché gli errori, incongruenze, etc. prodotti dalla reversione *batch*. Non solo, prima di ogni nuova importazione, i sistemisti di ACNP provvedono ad eliminare per intero i dati preesistenti, processano di nuovo i dati da caricare secondo due procedure distinte; una destinata al controllo della correttezza formale dei record (codifica del tracciato, *match* tra gli identificativi) che, in caso di errori, costituisce motivo di blocco, l'altra alla correttezza dei dati stessi. In quest'ultimo caso i record con dati incongruenti (ad es. la mancata corrispondenza tra l'indicazione del posseduto e gli anni di pubblicazione di una rivista) sono comunque caricati e visualizzati in ACNP; questi errori sono recuperati in automatico e opportunamente inviati ai singoli referenti o – come più probabile – all'unico referente delle biblioteche di ciascun Ateneo. In base a queste indicazioni saranno poi impostati ed organizzati i lavori di correzione e di allineamento. La raccomandazione – considerato il flusso di lavoro concordato – è quella di intervenire fin dove possibile sui dati alla fonte, ossia sui cataloghi Aleph, a patto ovviamente di non snaturarne le caratteristiche di fondo, in particolare l'apparato reticolare creatosi negli anni con il passaggio da uno *standard* descrittivo ad un altro, da un programma di automazione ad un altro. Del resto – ma si tratta di una questione che merita un adeguato approfondimento – sarebbe quanto meno bizzarro procedere in modo sistematico con tutta una serie di interventi tramite gestionale ACNP, se poi la logica sottostante al caricamento automatico impone la cancellazione totale di tutti i nostri dati prima di ogni importazione.

IL PROGRAMMA DI CONVERSIONE

Sono diversi gli elementi di criticità legati al progetto. Di alcuni possiamo limitarci a prenderne atto, nel senso che i margini operativi sono nulli, o quasi. Ad esempio il fatto che non possano essere caricati in automatico i seriali, per ragioni storiche, privi di ISSN, o lo scarso livello di sovrapponibilità tra i record Aleph e quelli in ACNP specialmente nella ricostruzione – fatta con i legami orizzontali e gerarchici tra titoli – delle vicende editoriali di una testata. Questo dipende semplicemente dal fatto che i due sistemi adottano *standard* di catalogazione diversi e che, a differenza del nostro catalogo, ACNP cerca di mantenere una rigorosa – tranne forse qualche rara eccezione – relazione biunivoca tra un titolo e il corrispondente ISSN. Le ricadute negative

originate da queste difformità descrittive si faranno sentire soprattutto quando – una volta trasmessoci il file delle incongruenze dai gestori – ci sarà bisogno di allineare i dati.

Di sicuro uno degli aspetti più problematici è rappresentato dalla differenza tra la struttura delle consistenze e delle lacune in ACNP e quella, anzi quelle, adottate nei nostri cataloghi Aleph. In ACNP sono inserite – secondo procedure standard – le sole annate, sempre di quattro cifre, con una punteggiatura essenziale, nei cataloghi Aleph invece – con alcune eccezioni – non viene seguita una struttura precisa, pertanto le soluzioni utilizzate sono le più svariate. Eppure era proprio lì che era necessario intervenire attraverso un applicativo capace di convertire l'insieme – o quanto meno la parte principale – di queste casistiche nella sintassi *standard* di ACNP. E' stato utilizzato il software database Filemaker developer 8, la cui potenza e flessibilità ci ha consentito di ottenere da un file in formato Aleph sequenziale – estratto con una semplice print-03 dai servizi di Aleph e poi opportunamente manipolato dagli *scripts* del programma – i *files* nel formato richiesto destinati al caricamento. Gli *scripts*, che ho sopra menzionato, sono semplici istruzioni, progettati per finalità specifiche. Ad esempio – senza cioè alcuna pretesa di esaustività considerata la complessità della architettura finale del programma – uno *script* serve per la validazione dei campi ISSN, altri devono controllare la lunghezza dei campi fissi, altri eseguono una semplice mappatura di campi, altri ancora – la maggior parte – devono normalizzare le consistenze e le lacune nella sintassi richiesta. In quest'ultimo caso è stata fondamentale una preliminare attività di analisi e capire se l'eterogeneità delle soluzioni adottate in Aleph fosse in qualche modo riconducibile ad un comune schema logico, indispensabile per gestire correttamente la consecutività delle annate nelle stringhe (nelle consistenze e nelle lacune di ACNP non sono ammesse infatti annate consecutive divise dal “punto e virgola”). Eppure mancava ancora una componente fondamentale, e cioè la predisposizione di quegli strumenti necessari per poter visualizzare – e quindi monitorare, estrarre, etc. – direttamente all'interno del catalogo Aleph l'esito della procedura di conversione. Non restava che intervenire di nuovo sul programma attraverso *routines* aggiuntive che agissero sugli indicatori del tag proprietario 958, dove sono inseriti i dati da processare (la sottobiblioteca, la consistenza, le lacune). La procedura deve pertanto poter generare due file di *output*, non solo quello da inviare al server ACNP, ma anche quello da caricare nel catalogo Aleph con le modifiche apportate agli indicatori, fino ad ora non utilizzati. La combinazione dei due indicatori fornisce tutta una serie di informazioni sul funzionamento – o sul mancato funzionamento – della procedura applicata a ciascun 958. Essa indica ad es. se eventuali motivi di blocco siano riconducibili alla errata codifica del nome della biblioteca, alla mancanza degli identificativi, al mancato *match* con gli identificativi inseriti in ACNP o alla particolare sintassi delle consistenze e delle lacune che il programma non è riuscito ad interpretare. Questo file sovrascrive quello importato originariamente

in Filemaker; si tratta dunque di operazioni delicate che richiedono cautela, nonché accurate misure di *backup*.

LA FASE ESECUTIVA PRESSO L'AREA SISTEMA BIBLIOTECHE DELL'UNIVERSITA' DI SIENA

Lo schema della fase esecutiva qui descritta è quella posta in essere presso l'Area sistema delle Biblioteche della Università di Siena, che però, anche per quanto sopra detto, finisce per essere esemplificativa di tutte le operazioni richieste dal progetto. Essa si è articolata in diverse attività, alcune delle quali “decentrate”, a carico cioè dei bibliotecari presso le singole strutture soprattutto nella fase preparatoria, altre invece “centralizzate” svolte dallo staff dell'Area. Ecco un sintetico schema delle attività pianificate:

- individuazione dei seriali cartacei privi di ISSN;
- assegnazione ai titoli delle riviste del codice ISSN o codice CNR attraverso un controllo incrociato sul catalogo ACNP;
- periodico monitoraggio sull'avanzamento del lavoro;
- relazione sui problemi riscontrati nel corso di tale attività;
- sulla base delle relazioni delle biblioteche analisi delle discrepanze tra i record di ACNP e quelli di Aleph in relazione alle consistenze, alla duplicazione dei codici ISSN, ai legami orizzontali e verticali;
- analisi delle “casistiche” delle consistenze e delle lacune, ossia delle diverse tipologie con cui sono state strutturate in Aleph;
- individuazione delle “casistiche” da processare e delle relative “istruzioni” per normalizzarle secondo lo schema previsto da ACNP (annate di 4 cifre, punteggiatura limitata al “punto”, “punto e virgola”, “trattino separatore”, mancanza di annate consecutive separate dal “punto e virgola”);

- implementazione del programma per convertire il file in formato Aleph sequenziale nel tracciato ACNP;
- elaborazione di una *routine* aggiuntiva per assegnare ai tag 958 gli indicatori legati all'esito della procedura;
- inserimento tramite *form* in linea, previa assegnazione dei codici ACNP, delle informazioni anagrafiche delle biblioteche: sede, orario, indirizzo, politica di *document delivery* adottata, etc.
- controllo del file per il caricamento in ACNP;
- invio file per il caricamento in ACNP;
- controllo del file in formato Aleph sequenziale per il caricamento in Aleph;
- caricamento file con modifiche agli indicatori del 958 in Aleph;
- analisi dei files di *feedback* prodotti dal sistema ACNP;

CONCLUSIONI

Come si ricava anche da quanto fino ad ora detto, la parte del progetto legata all'allineamento dei dati deve essere ancora verificata. In questo ambito il *mix* tra parziale gestione *on line* degli allineamenti e annuale caricamento *batch* è ancora tutto da sperimentare; sarà infatti necessario trovare un giusto equilibrio tra l'esigenza – per altro legittima – dei sistemisti di ACNP – che dobbiamo ringraziare per la loro costante disponibilità e cortesia – di avere dati aggiornati e facilmente bonificabili e l'altra, nostra, di procedere quasi unicamente con massicce importazioni automatiche, per evitare – proprio come dicevo all'inizio – una duplice, seppure parziale, gestione. Per avere un quadro più preciso sarà comunque interessante valutare nel *file* di *feedback* prodotto da ACNP dopo ogni importazione la quantità degli “errori” registrati rispetto alla totalità dei dati trasmessi, errori riconducibili per lo più ai diversi sistemi di catalogazione adottati nelle due piattaforme.

Rimane poi da definire esattamente il flusso di lavoro con il quale organizzare il caricamento dei dati di tutti gli Atenei interessati. Nella implementazione del programma si è cercato di garantire un

certo livello di flessibilità proprio in questa prospettiva; ora non resta che metterlo alla prova dei fatti.

Prima però il database deve essere integrato con una indispensabile interfaccia grafica. Deve essere cioè implementata quella parte di applicativo che consentirà a ciascuna sede di interagire in modo chiaro e razionale con l'intero sistema.

Un altro punto a cui tengo particolarmente riguarda proprio il formato delle consistenze in Aleph, che tanto ci ha fatto penare. Se la politica adottata sarà appunto quella di periodiche e massicce importazioni *batch* un elemento irrinunciabile dovrà essere – come già sperimentato presso alcuni Atenei – la definizione di un protocollo *standard* per l'inserimento dei dati di posseduto e delle lacune, magari elaborato proprio sulla base dei *pattern* individuati e interpretabili dal programma.