



Linee guida per la deduplicazione dei periodici elettronici

OBIETTIVO

Obiettivo della deduplicazione è **individuare il numero di periodici elettronici disponibili**

- 1) acquisiti o in abbonamento con gli editori (esclusi i titoli free),
- 2) sia correnti che dismessi (purché per questi ultimi si possiedano i diritti di accesso),
- 3) inclusi quelli contenuti in banche dati a testo pieno,
- 4) contando ogni periodico una sola volta anche se disponibile su più piattaforme o database.

DEDUPLICAZIONE IN SFX¹

Da SFX Admin Center:

- selezionare KBTools -> Export Tool -> Advanced Export Queries.
- selezionare i gruppi e/o gli institutes che nel loro insieme costituiscono l'intero posseduto dell'Ateneo, utilizzando il comando "Restrict to the following institutes/groups (optional)" (in alto a destra)

Nota Bene: se si ha un'istanza locale ed una su server remoto (es. istanza consortile CIPE) i due dati vanno sommati, quindi la procedura di export illustrata sotto (fino al salvataggio del file .txt) va ripetuta sul server remoto e i dati vanno integrati in un unico file excel.

- selezionare le seguenti opzioni di Advanced Export:
 1. Select Output Format = TXT,
 2. Export which Object Type = Serials
 4. Export **active** portfolios with the following services = getFullTxt (verranno selezionati per l'esportazione i soli target attivi),
 5. Export from "SPECIFIC"
- selezionare i soli target in abbonamento escludendo quelli free; cliccare su "Add" e spostare i target selezionati nel riquadro a destra
- cliccare il pulsante "Submit".
- cliccare su "Download Result" e salvare il file .txt.

Importare il file .txt in un file Excel. Seguire una delle due procedure di seguito illustrate, a seconda che i titoli siano o meno tutti contenuti in SFX.

1) Opzione Object ID² (da utilizzare se tutti i titoli disponibili sono contenuti in SFX):

- rimuovere tutte le colonne ad esclusione dell'Object ID (di norma è la colonna "E", la quinta dopo quella dell'ISSN)
- filtrare per "record unico" Object ID
- il numero di titoli filtrati fornirà i titoli unici.

2) Opzione ISSN (da utilizzare solo se non tutti i titoli sono contenuti in SFX):

¹ La procedura è valida sia per SFX3 che per SFX4

² L' "object ID" è il numero unico utilizzato nel database di SFX per identificare un oggetto (SFX General Users Guide)

- rimuovere tutte le colonne ad esclusione di titolo, ISSN e ObjectID
- filtrare per "record unico" Object ID
- selezionare e spostare i titoli filtrati in un foglio specifico (con questa operazione si ottiene una prima riduzione dei record e quindi una più agevole gestione delle procedure successive)
- aggiungere al foglio prodotto i titoli ed eventuali ISSN delle riviste aggiuntive non contenute in SFX
- filtrare i titoli con ISSN e spostarli su di un altro foglio; rimarranno quindi su foglio precedente i soli titoli senza ISSN
- nel foglio di titoli senza ISSN filtrare per "record unico" titolo
- nel foglio di titoli con ISSN filtrare per "record unico" ISSN
- la somma del numero dei titoli filtrati di questi due ultimi fogli fornirà i titoli unici.

Se il numero di target attivi in SFX, e conseguentemente il numero dei titoli da deduplicare, è tale rendere poco agevole l'utilizzo da subito di un unico file excel si consiglia di frazionare preliminarmente in più file il lavoro lanciando le volte necessarie la procedura di selezione, importazione e filtro. I file andranno quindi combinati per ottenere un unico file sul quale operare la procedura di filtro finale.

DEDUPLICAZIONE MEDIANTE CATALOGO/OPAC

Se si sono catalogati tutti i propri periodici elettronici in un catalogo:

- verificare se il gestionale consente l'estrazione dei titoli unici, oppure
- se è possibile farne richiesta specifica agli amministratori.

È possibile ad esempio richiedere agli amministratori di ACNP l'estrazione sul proprio posseduto di vari dati statistici, fra i quali il dato "Testate digitali" corrisponde al numero di titoli unici di periodici.

Se non è possibile nessuna di queste alternative, estrarre dal catalogo o dall'opac i periodici elettronici posseduti rispondenti alle caratteristiche indicate nell'"Obiettivo" di cui sopra e procedere come indicato per il punto "Deduplicazione mediante Database locale, File excel, o altre soluzioni di gestione locale".

DEDUPLICAZIONE MEDIANTE DATABASE LOCALE, FILE EXCEL, O ALTRE SOLUZIONI DI GESTIONE LOCALE

Se si gestiscono i periodici elettronici mediante un database locale, un file Excel, una lista A-Z o altri strumenti simili, ricavarne un file importabile in un foglio Excel che contenga le due colonne "Titolo" e "ISSN" e procedere come segue:

- filtrare i titoli con ISSN e spostarli su di un altro foglio; rimarranno quindi sul primo foglio i soli titoli senza ISSN
- nel foglio di titoli senza ISSN filtrare per "record unico" Titolo
- nel foglio di titoli con ISSN filtrare per "record unico" ISSN
- la somma del numero dei titoli filtrati dei due fogli fornirà i titoli unici.

Il presente documento si ispira a: ARL Statistics Best Practices for Deduplicating Serial Titles - December 4, 2008 <www.arl.org/bm~doc/dedupe_best_practices.doc>.