

AVIndexer v. 0.9 (beta)

Guida all'utilizzo

Prerequisiti

- Sistema operativo Microsoft Windows XP/Vista/7/8
- .NET Framework 4.0

Installazione

AVIndexer è disponibile come applicazione portabile¹ e quindi non necessita di una installazione vera e propria. Per utilizzare il software è sufficiente estrarre in una cartella del file system il contenuto del pacchetto compresso.

Avvio

Per avviare il programma basta avviare il file eseguibile (estensione “.exe”) presente nella cartella di installazione.

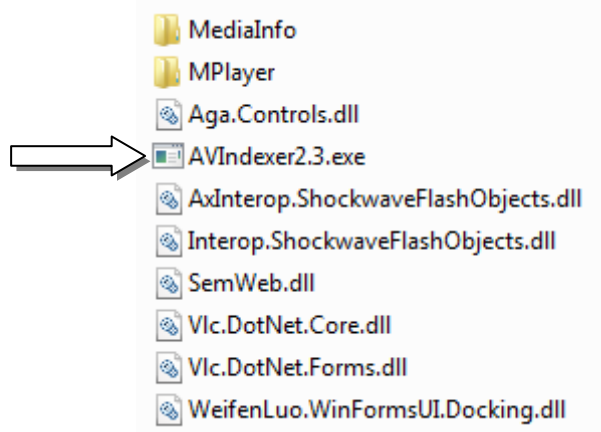


Fig. 1

Impostazioni

Per funzionare correttamente AVIndexer necessita di due programmi esterni disponibili con licenza open source:

- MediaInfo (disponibile su <http://mediaarea.net/it/MediaInfo>)
- MPlayer (disponibile su <http://www.mplayerhq.hu>)

¹ http://it.wikipedia.org/wiki/Applicazione_portabile

Al primo avvio, se nella cartella di AVindexer non sono presenti le cartelle relative ai due programmi, viene chiesto all'utente di impostare i percorsi dei file eseguibili di MediaInfo e di MPlayer.

Tali programmi sono indispensabili per ricavare le informazioni sui file audio/video (MediaInfo) durante l'aggiunta di un media ad un oggetto digitale e per riprodurli (MPlayer) durante la fase di annotazione.

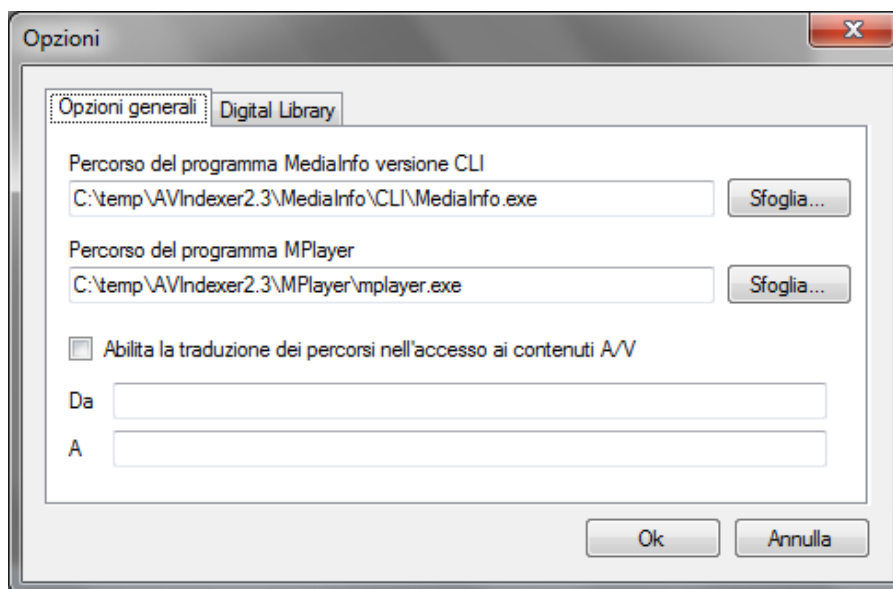


Fig. 2

Informazioni sul programma

"AVIndexer è stato realizzato presso la Scuola Normale Superiore di Pisa nell'ambito del progetto "Ti Racconto la Storia" della Direzione Generale per gli Archivi del Ministero per i Beni e le Attività Culturali. Questo programma può essere utilizzato e distribuito liberamente per qualsiasi scopo, anche commerciale. L'autore non ne garantisce il perfetto funzionamento e declina ogni responsabilità per eventuali danni che dovessero derivare dal suo utilizzo."

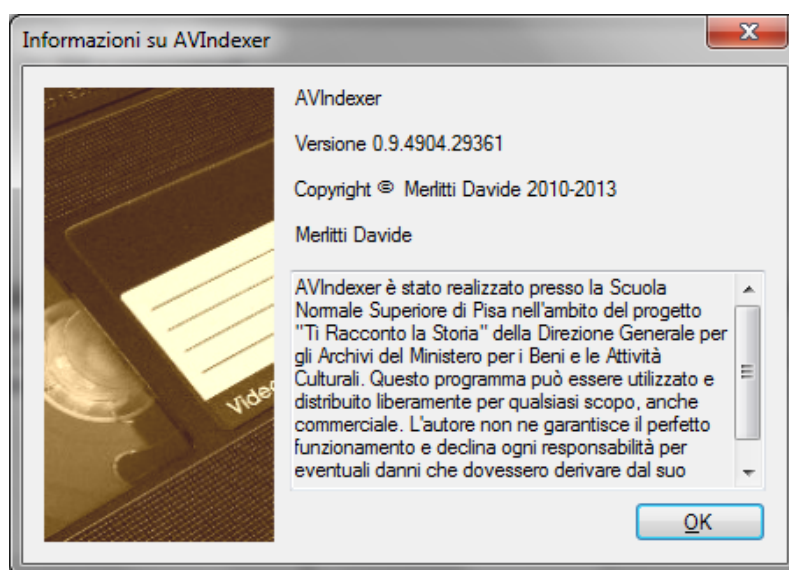


Fig. 3

Creazione di un nuovo progetto

Viene chiesto di selezionare il tipo di progetto:

- “Local File System”
- “Fedora Commons Digital Library”

Normalmente si seleziona il primo tipo a meno che non si disponga di un server *Fedora Commons* accessibile tramite protocollo http. Un progetto di tipo “Local File System” memorizza gli oggetti digitali sul file system del computer locale e consente di lavorare scollegati dalla rete locale e da Internet.

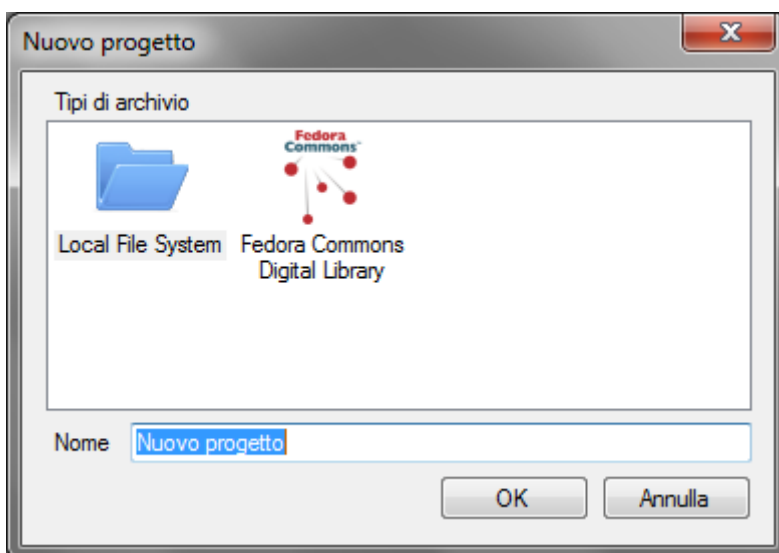


Fig. 4

Aggiunta di un thesauro

Se il progetto prevede l'utilizzo di un thesauro predefinito (esistente o generato ad hoc), si può selezionare un file in formato xml SKOS e inserirlo una volta per tutte. Riaprendo il progetto, ovviamente, il thesauro sarà subito disponibile per l'aggiunta dei termini ai segmenti.

Creazione di un oggetto digitale

La prima cosa da fare per iniziare il lavoro di indicizzazione è creare un oggetto digitale corrispondente al documento (ad esempio una intervista). L'oggetto digitale è la rappresentazione quindi di un documento e serve per tenere insieme la versione digitalizzata del documento stesso, eventuali file derivati e tutte le informazioni su di esso (metadati).

Nuovo oggetto digitale

Archivio

Tipo: File System

URL: file:///C:/Users/davide/AppData/Roaming/AVIndexer/Archivio%20Digitale

Sfogli... Apri

ID:

Etichetta: Nuovo Oggetto Digitale 2

Tipo:

Collezioni

PID	Etichetta
-----	-----------

Nuova collezione... OK Annulla

Fig. 5

E' possibile assegnare all'oggetto digitale un identificatore (ID), un'etichetta e un tipo. L'identificatore, che deve essere univoco, è composto di due parti separate dal carattere ':' (due punti). Esempio: "prefisso:oggetto1".

Aggiunta di un media

Altra operazione fondamentale è la selezione di un file in formato audio o video (avi, mpeg, flv, mp3 etc.) che serva come base per il lavoro di indicizzazione. Il file non deve essere necessariamente il file master prodotto da un progetto di digitalizzazione. Può essere anche un file derivato a bassa risoluzione purché sia conservata la durata complessiva.

Oltre a selezionare un file dal file system del computer è possibile aggiungere anche un contenuto accessibile via Internet/Intranet specificando l'URL (ad esempio è possibile aggiungere contenuti presenti su YouTube o altre digitali library che rendono accessibili i propri contenuti ma è necessario che questi possano essere oggetto di operazioni di seek casuali e di estrazione di metadati).

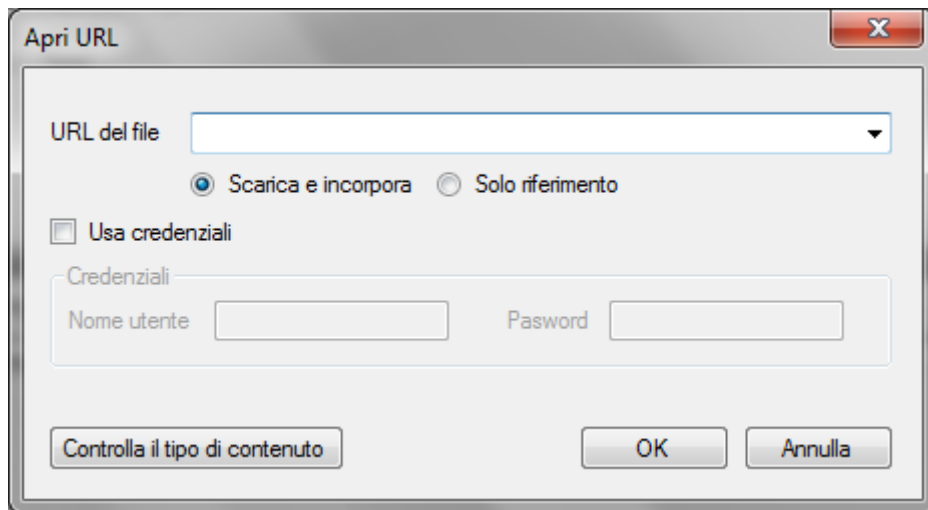


Fig. 6

Elementi dell'interfaccia utente

La finestra è suddivisa in tre aree principali: l'area del contenuto, dove viene visualizzato l'oggetto digitale corrente; l'area di destra che consente di esplorare il progetto nella sua struttura e di visualizzare/modificare informazioni contestuali; l'area dell'annotazione corrente. L'area del contenuto contiene a sua volta il media player, la timeline, i segmenti, il thesaurus, le annotazioni etc. L'elenco degli elementi è il seguente:

1. Media Player (visualizza il media corrente e consente di metterlo in pausa e di farlo ripartire)
2. Timeline (visualizza il tempo corrente del media e i segmenti)
3. Segmenti (elenco dei segmenti presenti nel media corrente)
4. Thesaurus (il thesauro selezionato nel progetto)
5. Annotazioni (l'elenco delle annotazioni presenti nel segmento corrente)
6. Esplora soluzione (visualizzazione del progetto nella sua struttura)
7. Proprietà oggetto digitale (visualizzazione e modifica dei metadati)
8. Esplora oggetto digitale (visualizzazione dei media presenti nel documento)
9. Finestra annotazione (visualizzazione e modifica degli attributi dell'annotazione corrente)

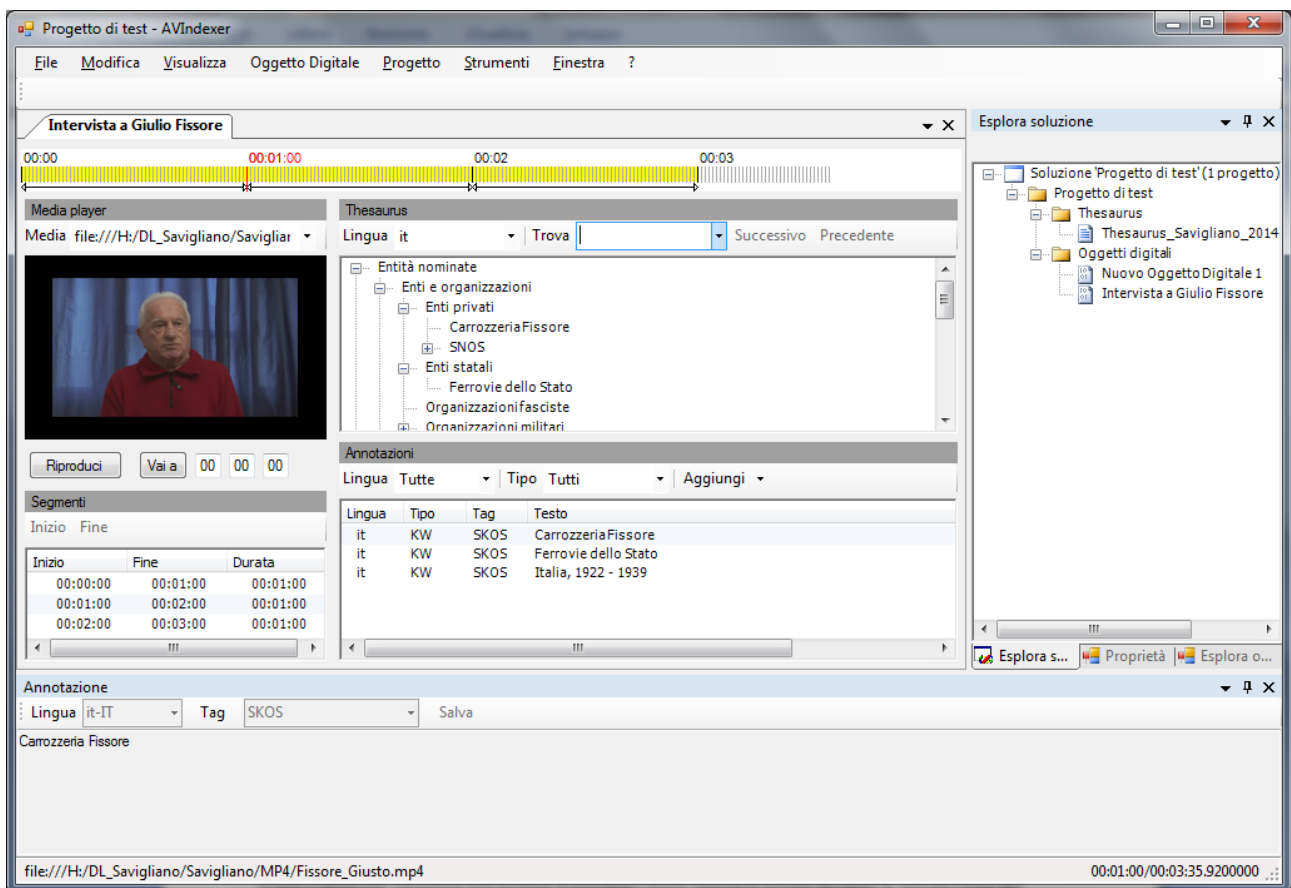


Fig. 7

Aggiunta di segmenti

Prima di iniziare il lavoro di aggiunta di annotazioni (parole chiave o testo libero) è possibile creare un certo numero di segmenti di durata prefissata a partire da un punto del flusso audio/video (modello SHOA).

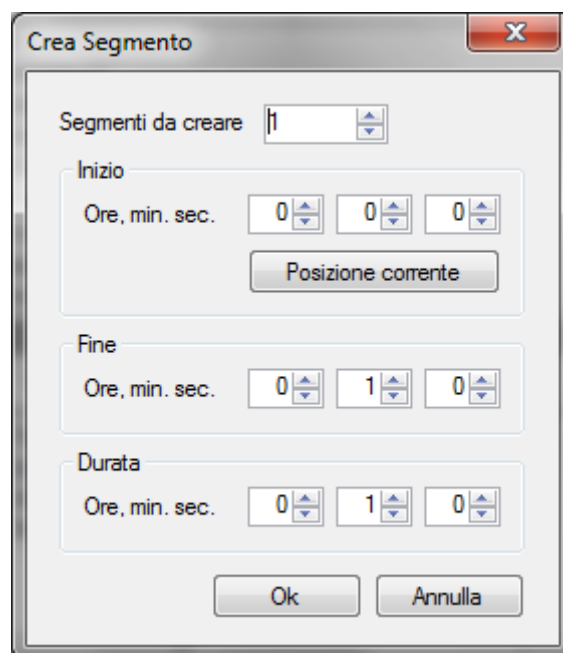


Fig. 8

In alternativa alla creazione preventiva, è possibile creare segmenti al volo durante la riproduzione del contenuto tramite i pulsanti “Inizio” e “Fine” presenti nella barra dell’area “Segmenti” della finestra principale.

Eliminazione di un segmento

Un segmento precedentemente creato può essere eliminato (ad esempio se non è più necessario o se è stato creato per sbaglio) in qualsiasi momento tramite il comando “Elimina segmento” del menu “Oggetto Digitale” (attenzione perché non viene chiesta conferma).

Aggiunta di annotazioni

Tramite la apposita funzione (da menu principale o pulsante nella barra dell’area “Annotazioni” della finestra principale) è possibile aggiungere un’annotazione di tipo “Parola chiave” o “Testo libero”. Le parole chiave sono frasi molto brevi che esprimono concetti utili in fase di recupero tramite ricerche.

L’annotazione a testo libero consente di esprimere periodi più complessi utilizzabili ad esempio per trascrizioni parziali, regesti e annotazioni di lavoro. Le annotazioni possono essere ulteriormente arricchite con tag liberi il cui scopo è lasciato alle particolari implementazioni (ad esempio i tag, nel caso di trascrizioni o regesti, possono corrispondere con i nomi e cognomi o le sigle degli intervistati e degli intervistatori. Nel caso di annotazioni di lavoro possono servire come filtro di ricerca etc.).

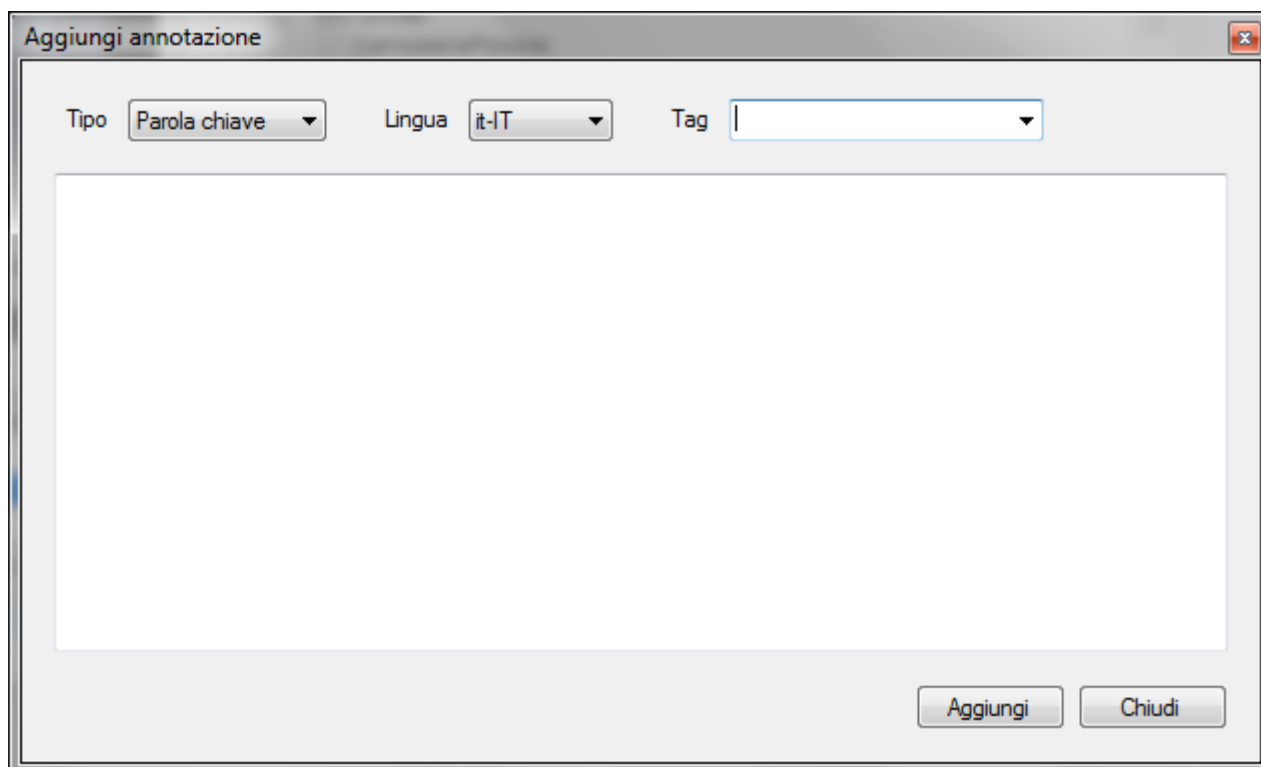


Fig. 9

Inserimento dei metadati principali

Luogo e data dell’intervista

Intervistati

Intervistatori

Produttori

Creatori

Titolo (eventualmente in diverse lingue)

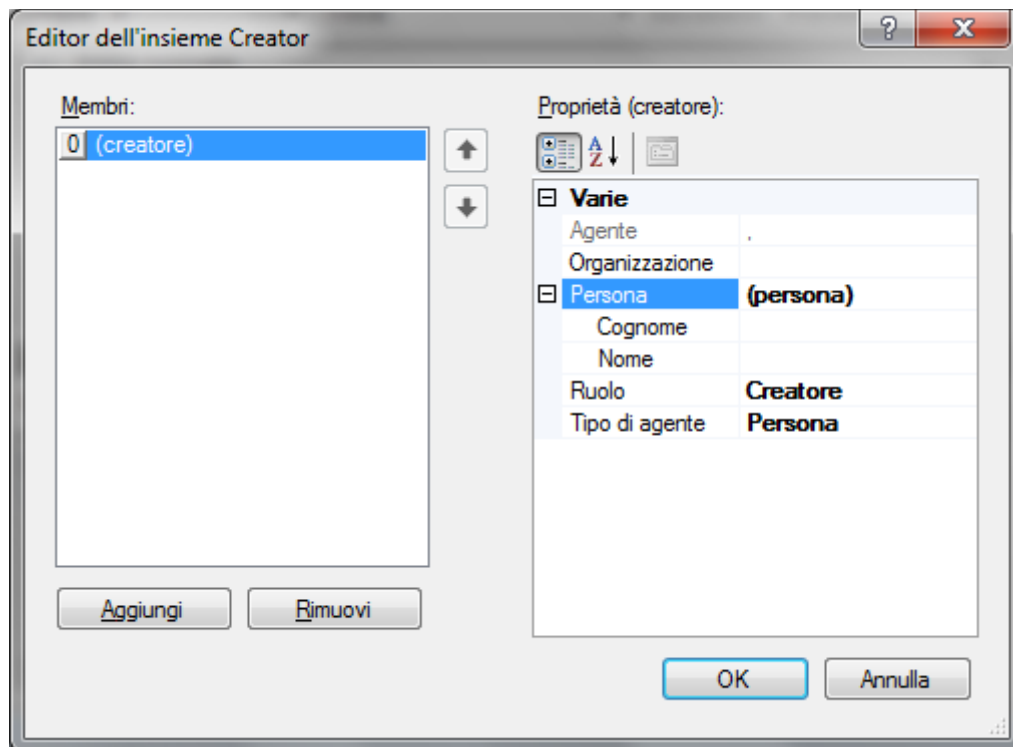


Fig. 10

Salvataggio

Salvataggio dell'oggetto digitale

L'oggetto digitale corrente può essere salvato tramite il comando "Salva oggetto digitale" del menu "File".

Salvataggio del progetto

Il progetto corrente può essere salvato tramite il comando "Salva progetto" del menu "File".

Chiusura del progetto

Una volta terminato il lavoro di indicizzazione ed effettuato il salvataggio, è possibile chiudere il progetto (File > Chiudi progetto).

Apertura di un progetto esistente

All'avvio del programma, per tornare a lavorare sul progetto di indicizzazione si apre il progetto e ci si ritrova nell'esatta situazione in cui lo si è salvato.

Aggiunta di un oggetto digitale

Dopo aver aperto un progetto si può continuare ad esempio ad aggiungere altri oggetti digitali composti da uno o più media e suddivisi in segmenti.

Approfondimenti

Thesauri in formato XML SKOS

SKOS è uno standard del W3C ed è l'acronimo di *Simple Knowledge Organization System* (semplice sistema di organizzazione della conoscenza) . E' basato su altri standard del W3C (RDF e OWL) e fornisce un metodo per la rappresentazione di vocabolari controllati, tassonomie e thesauri.

Vocabolario controllato: lista (elenco) di termini

Esempio

- Italia
- Francia
- Spagna
- Grecia

Tassonomia: liste di termini strutturati gerarchicamente

Esempio

- Europa
- Italia
 - Abruzzo
 - Basilicata
 - [...]
- Francia
- Spagna
- Grecia

Thesaurus: una tassonomia dove ogni termine ha informazioni supplementari come etichette in varie lingue, etichette supplementari, note d'ambito, note storiche, esempi di utilizzo, relazioni con altri termini per esprimere correlazioni, attinenze e corrispondenze con altri termini.

L'elemento fondamentale di un thesauro SKOS è il *Concept* che corrisponde alle unità di pensiero, le idee, i significati, gli oggetti, le categorie, gli eventi. Ogni concept è identificato da un URI.

Esempi di thesauri SKOS usati per progetti di indicizzazione con AVIndexer

- http://www.shoah.acs.beniculturali.it/thesauri/thesaurus_shoah_it_1.0.3.skos.xml
- http://carta.sns.it/dl/lartte2/thesauri/thesaurus_museo_della_mente_1.2.1.skos.xml
- http://carta.sns.it/dl/lartte2/thesauri/thesaurus_termini_rurali_artigianali_toscana_1.0.skos.xml

Esempi di thesauri SKOS generici

- <http://skos.um.es/unescothes/CS000/rdfxml> (<http://skos.um.es/unescothes/>)
- <http://thes.bncf.firenze.sbn.it/thes-dati.htm> ()

Thesauri SKOS censiti dal W3C

<http://www.w3.org/2001/sw/wiki/SKOS/Datasets>

Riferimenti:

“W3C - SKOS Simple Knowledge Organization System Primer”

<http://www.w3.org/TR/skos-primer/>

Ontologie OWL

Classi, proprietà, individui (istanze), dati per la rappresentazione della conoscenza.

Esempio di ontologia che rappresenta un'intervista della SHOAH:

<http://lartte.sns.it/fedora34/objects/shoah%3A11536/datastreams/OWL/content>

Riferimenti:

- <http://www.w3.org/TR/owl2-primer/>
- <http://www.w3.org/TR/owl2-overview/>

MPEG-7

Standard MPEG per la rappresentazione di metadata relative a contenuti multimediali (immagini, audio, video, testi, composti)

Esempio di datastream MPEG-7 corrispondente a una intervista:

<http://carta.sns.it/fedora34/objects/shoah%3A11536/datastreams/MPEG-7/content>

Riferimenti:

<http://mpeg.chiariglione.org/standards/mpeg-7>

Fedora Commons

Oggetti digitali, datastream, RELS-EXT, RELS-INT, ontologia di base, modelli di contenuti

Esempi di oggetti digitali corrispondenti ad interviste:

- <http://carta.sns.it/fedora34/objects/savigliano:FissoreGiusto>
- http://carta.sns.it/fedora34/objects/carm:006_1991

Dump xml di un oggetto digitale:

<http://carta.sns.it/fedora34/objects/savigliano%3AFissoreGiusto/objectXML>

notare il datastream “AUDIT” che contiene tutte le operazioni subite dall'oggetto digitale durante tutta la sua vita iniziata dal momento della creazione in Fedora Commons.

Interfacce

Interfaccia per la ricerca

<http://carta.sns.it/fedora34/search>

Interfaccia di amministrazione web

<http://carta.sns.it/fedora34/admin>

Interfaccia di ricerca nell'indice delle risorse in formato RDF (triplestore)

<http://carta.sns.it/fedora34/risearch>

Interfacce applicative REST API-A e API-M

(usate dalle applicazioni software per dialogare con Fedora Commons. A=Access, M=Management)

Interfaccia OAI-PMH

(usata per fornire all'esterno i metadati via protocollo di harvesting OAI-PMH usato ad esempio da CulturalItalia, Europea e altri aggregatori di metadati)

Riferimenti:

- <http://www.fedora-commons.org/about>
- <https://wiki.duraspace.org/display/FEDORA37/Fedora+Digital+Object+Model>
- <https://wiki.duraspace.org/display/FEDORA37/Digital+Object+Relationships>
- <http://www.fedora.info/definitions/1/0/fedora-relsext-ontology.rdfs>
- <https://wiki.duraspace.org/display/FEDORA37/Content+Model+Architecture>
- <https://wiki.duraspace.org/display/FEDORA37/Resource+Index+Search>
- <https://wiki.duraspace.org/display/FEDORA37/REST+API>
- <https://wiki.duraspace.org/display/FEDORA37/Basic+OAI-PMH+Provider>
- <https://wiki.duraspace.org/display/FEDORA37/Fedora+Web+Administrator>

L'interfaccia di amministrazione della Digital Library Ti Racconto La Storia

Gestione dei contenuti (pagine statiche)

Collezioni e oggetti in evidenza

Modifica dei metadati DC di collezioni e oggetti

Gestione degli accessi a collezioni e oggetti

Gestione degli utenti

Gestione delle richieste di accesso

Gestione degli indici

Gestione dei thesauri