

MAGTOOL:

creazione dei metadati dell'oggetto digitale.

INDICE

Creazione di un progetto con il MAGTOOL.....	3
Upload del “pacchetto di ingestion” sul server di INERA.....	7
Creazione del file .XML con i metadati.....	8
Apertura dell’oggetto da metadattare.....	8
Elementi fondamentali dell’interfaccia del MAGTOOL.....	10
Sezioni (schede) del MAGTOOL.....	11
Sezione DATI GENERALI (corrispondente all’elemento GEN del MAG).....	11
Sezione DATI BIBLIOGRAFICI (corrispondente all’elemento BIB del MAG).....	12
Sezioni di NOMENCLATURA IMMAGINI, TESTI e TESTI OCR.....	15
IMMAGINI: import.....	16
IMMAGINI: nomenclatura.....	18
PDF: import.....	20
PDF: nomenclatura.....	20
OCR: import.....	21
OCR: nomenclatura.....	21
Sezione STRUTTURA DOCUMENTO (che corrisponde all’elemento STRU del MAG)...	21
NODI PRINCIPALI (voci di primo livello dell’indice).....	22
SOTTO NODI (voci dal secondo livello in poi dell’indice).....	25
Fasi finali della metadattazione.....	27
Elaborare i DATI TECNICI.....	27
Impostare lo USAGE per immagini e documenti.....	28
Impostare il VALORE <SOURCE> per gli OCR.....	29
Salvataggio.....	30
Pre-pubblicazione dell’oggetto digitale.....	33

I metadati di un oggetto digitale sono contenuti all'interno di un file in formato .xml e seguono (almeno per il momento) lo standard MAG (Metadati Amministrativo Gestionali).

Il MAGTOOL è uno strumento che permette di creare in modo “amichevole” questo file .xml presentando all'utente una interfaccia grafica con tutti i campi da riempire per creare il file dei metadati.

Per un eventuale riferimento sui vari campi dello standard MAG (e sapere quindi a cosa ogni campo si riferisce e quali sono i metadati minimi obbligatori secondo questo standard) è possibile consultare il manuale ufficiale a questo link:

<https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html>

Per la Digital Library di Unito sono state redatte delle linee guida cui attenersi nel nostro caso specifico disponibili a questo link:

<https://drive.google.com/file/d/1uq1eCHlbHBR7ZBgbBDbgy1q7j4-f7tRU/view>

Mediante il MAGTOOL:

1. **si crea un progetto** avente come base l'oggetto digitale per il quale si vogliono registrare i metadati (operazione da fare solo se necessaria, ovvero se non è mai stata creata prima - NOTA: per il Fondo Prato è già stato creato)
2. si **produce il file finale** (che sarà un file .xml contenente tutti i metadati registrati)
3. si può **inviare alla TECA il pacchetto** che comprende sia le immagini dell'oggetto digitale che il relativo file xml per poi pubblicarlo con lo strumento TECA

Creazione di un progetto con il MAGTOOL

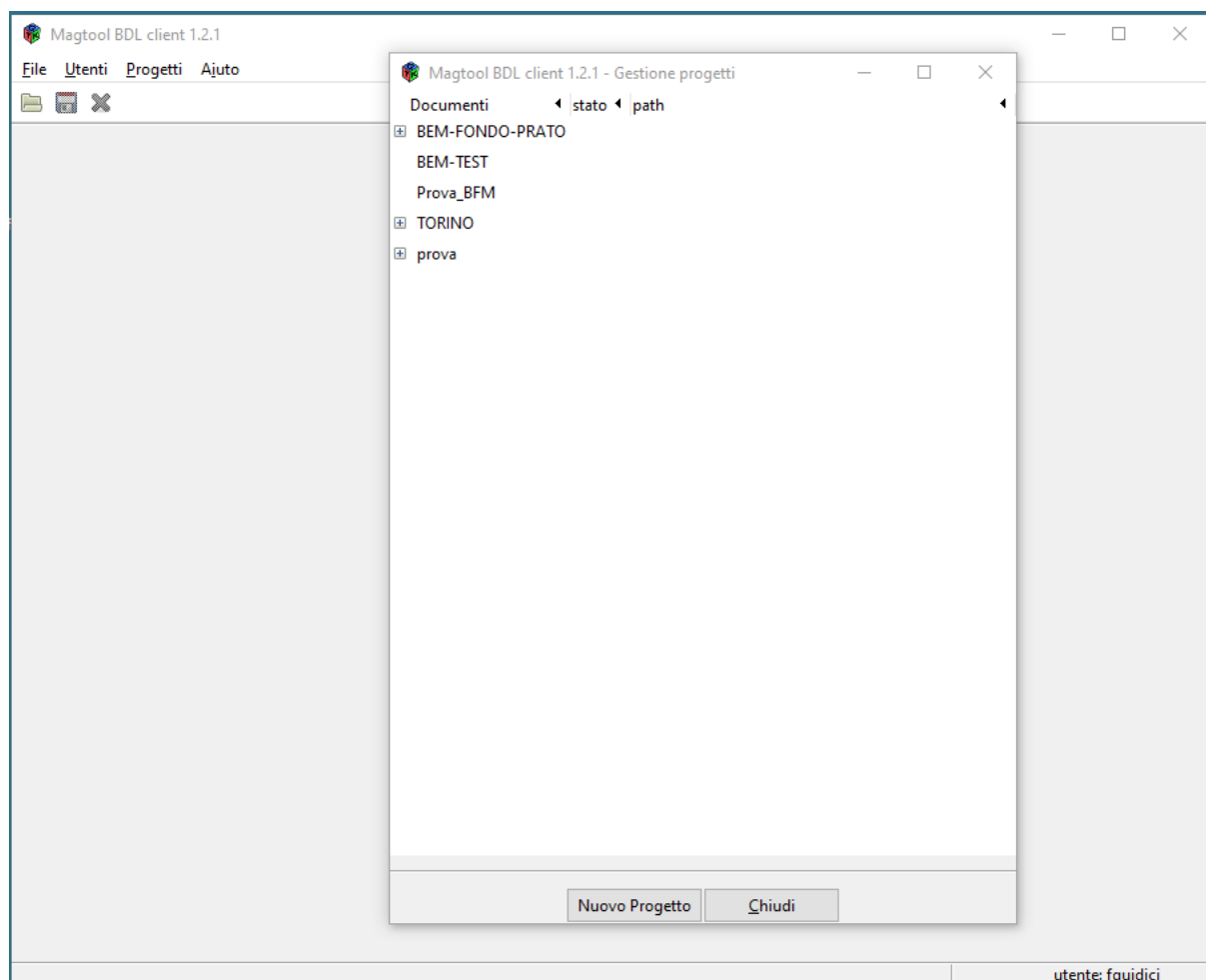
Prima di procedere con il MAGTOOL per la propria biblioteca è necessario creare un progetto che conterrà tutti gli oggetti ad esso relativi.

Attenzione: per il Fondo Prato della BEM il progetto è già stato creato e si chiama BEM-FONDO-PRATO. Quindi si può saltare questa sezione.

Per creare il progetto contenitore procedere come segue.

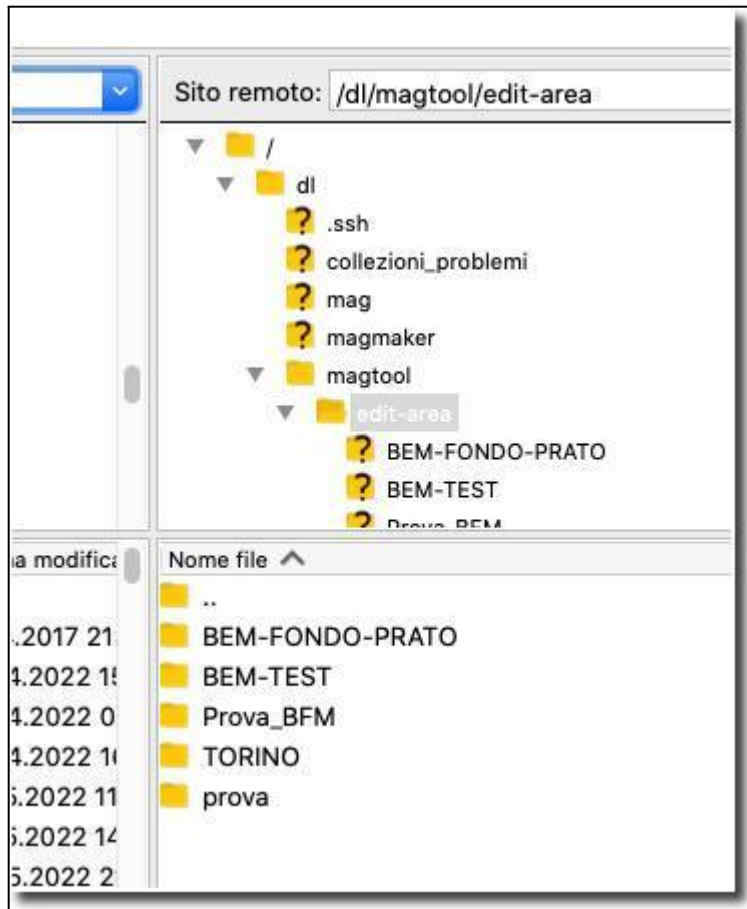
1. Aprire il MAGTOOL
2. Selezionare “*Progetti* → *Nuovo*”
3. Dare un nome al progetto (nell’esempio a seguire BEM-TEST) e confermare.
4. Si aprirà una pagina che elenca tutti i progetti, incluso quello appena creato.

Nell’immagine qui sotto si vedono il progetto appena creato (BEM-TEST) e altri, tra cui quello del Fondo Prato (BEM-FONDO-PRATO):



5. La creazione del progetto crea automaticamente sul server di Inera nel percorso *magtool* → *edit-area* una cartella con lo stesso nome del progetto, dove andremo a

depositare i file dei nostri oggetti digitali.



6. Il progetto è stato creato.

Prima di chiudere la finestra di progetto è necessario andare ad impostare correttamente:

1. il **server dell'OPAC Sebina** (che mediante la funzione di ricerca Z39.50 consentirà di recuperare automaticamente i metadati dall'OPAC nazionale);
2. le **impostazioni generali di progetto** (che consentiranno di velocizzare in parte la compilazione delle schede dei metadati andando a pre-popolare automaticamente la scheda "Dati generali");

Per queste due operazioni procedere come segue.

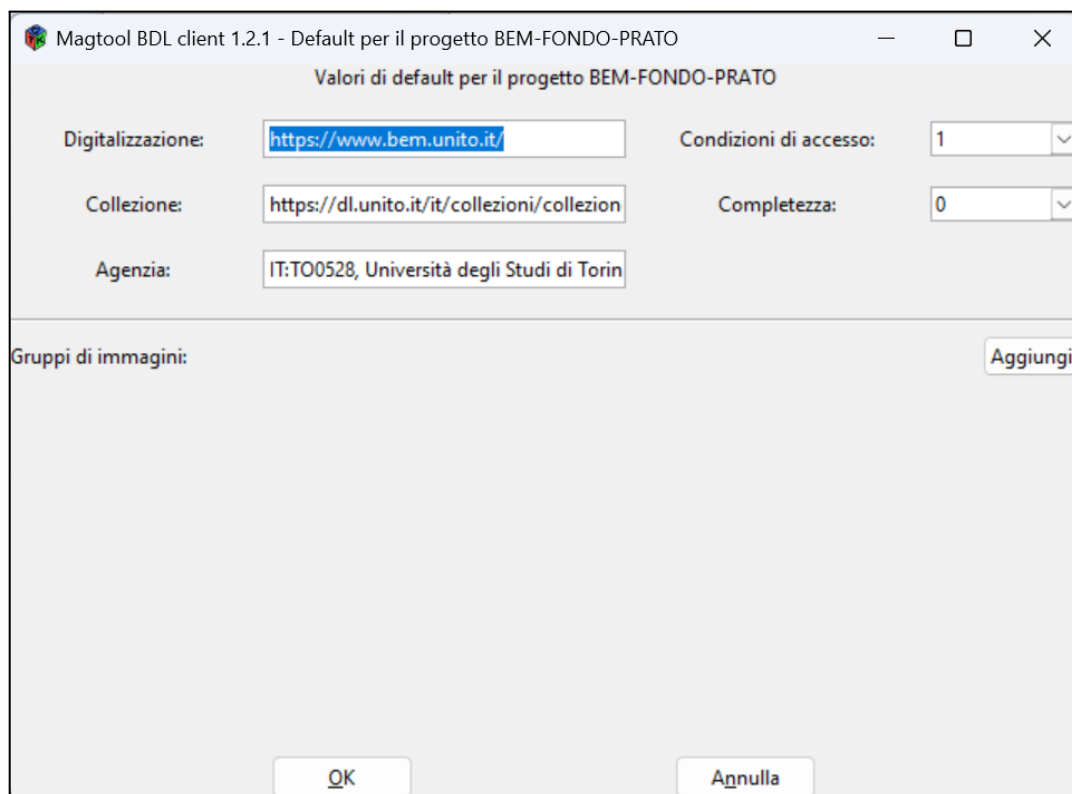
Con la finestra "Gestione progetti" ancora aperta (se non lo è andare in Progetti → Gestisci):

1. cliccare con il pulsante destro del mouse sul proprio progetto e scegliere "Modifica progetto"

2. si apre un box dove è possibile indicare i **parametri di collegamento con l'OPAC Sebina** i cui campi (se non lo sono già) andranno valorizzati come nell'immagine seguente:



3. Nella stessa finestra, cliccare poi sul pulsante in alto a destra nella “**Imposta default di progetto**” e si aprirà un box dove è possibile specificare per il progetto in corso i valori di default della scheda “Dati generali” in modo da non doverli compilare ogni volta e che siano consistenti per ogni oggetto incluso nel progetto:



In particolare i campi vanno valorizzati come segue:

- a. **Digitalizzazione:** qui va messa l'URL della pagina del sito web di presentazione del progetto. In sua assenza è possibile valorizzare il campo con l'URL principale del sito web della biblioteca.
- b. **Collezione:** qui va messa l'URI della collezione come impostata sul CMS (vedere le istruzioni operative sulla creazione di una collezione).

Per il Fondo Prato questa è:

<https://dl.unito.it/it/collezioni/collezione/fondo-giuseppe-prato/>

(**attenzione alla precisione della url** - inclusa la barra terminale ed eventuali spazi prima e dopo - altrimenti l'oggetto non comparirà nella collezione).

- c. **Agenzia:** valorizzare secondo quanto indicato nelle [Linee Guida per la DL Unito](#) nella “**Appendice A: Valori per il tag <agency> relativi alle biblioteche dello SBA dell'Università di Torino**”. Nel caso della BEM inserire la stringa:
IT:TO0528,
- d. **Condizioni di accesso** (uso riservato: 0 o pubblico: 1): scegliere *1*
- e. **Completezza** (digitalizzazione completa: 0, non completa: 1): scegliere *0*

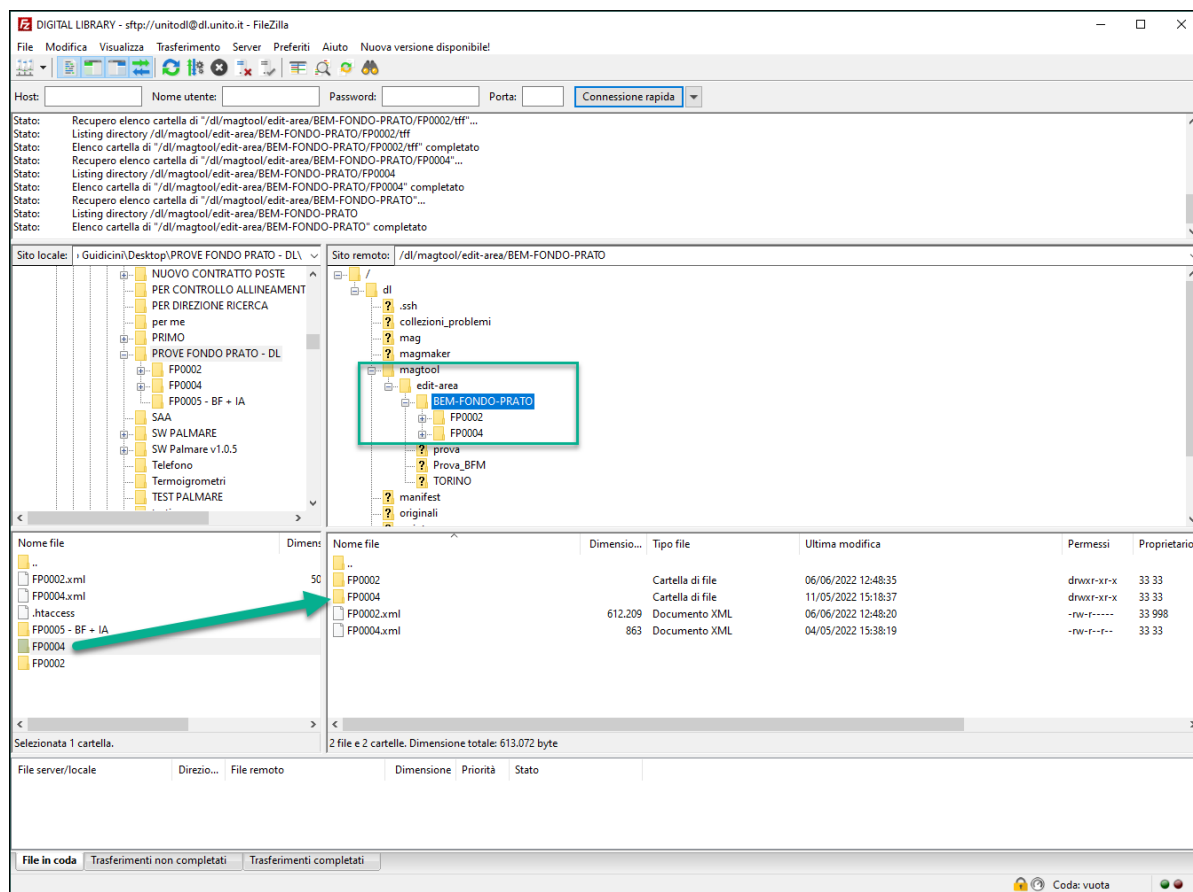
Upload del “pacchetto di ingestion” sul server di INERA

Per questa operazione bisogna, nell'ordine

1. Aprire il programma Filezilla (per informazioni su Filezilla vedere la sezione dedicata)
2. Navigare nella colonna di sinistra, fino a raggiungere la cartella sul pc locale che contiene i file dell'oggetto digitale
3. Navigare nella colonna di destra fino a raggiungere la cartella del server remoto dove copiare i file dal pc locale:

Nel caso del Fondo Prato: `dl/magtool/edit-area/BEM-FONDO-PRATO`

4. Caricare mediante Filezilla l'intera cartella dal disco locale sul server trascinandola dal PC locale al PC remoto nella cartella di destinazione:



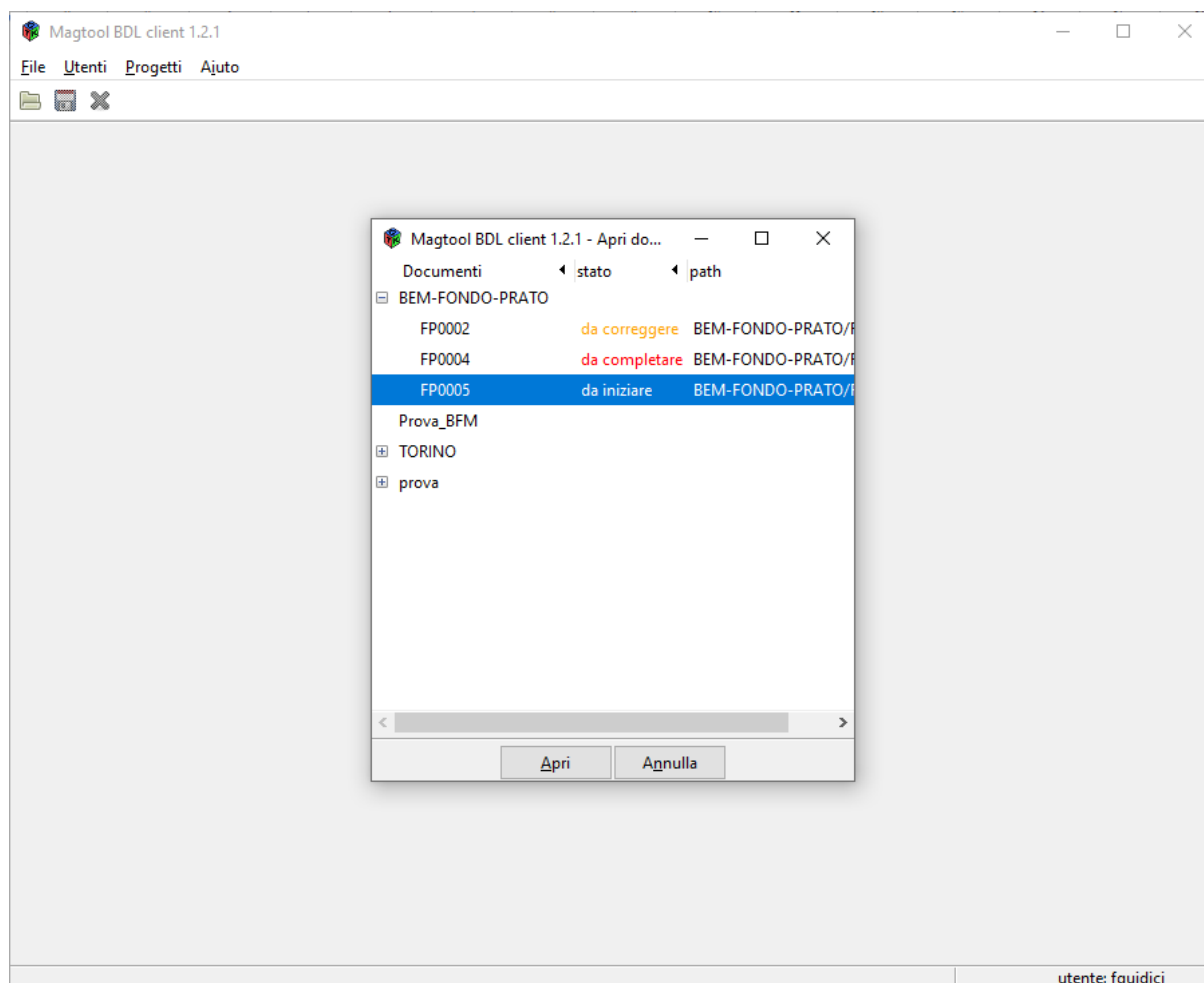
5. Attendere il caricamento dell'intera cartella (che potrebbe richiedere un po' di tempo a seconda della quantità di file).

Creazione del file .XML con i metadati

Apertura dell'oggetto da metadattare

1. Aprire il MAGTOOL. A seconda di come lo si è installato è possibile:
 - a. cliccare sull'icona relativa presente sul desktop, o
 - b. se l'icona non è presente aprire la cartella in cui ci sono i file dell'applicazione e lanciare il file magtool.exe (quella classificata come "applicazione").

2. Cliccare su *File* → *Apri* e nella finestra che si apre navigare nelle cartelle fino a trovare la cartella di progetto appena caricata (nell'esempio dell'immagine seguente consideriamo FP0005) che sarà classificato come in stato *“da iniziare”* (*)
3. Evidenziare l'oggetto su cui si vuole lavorare e cliccare sul pulsante *“Apri”*



(*) - Nota sugli stati di un oggetto digitale

Notare che sotto il progetto BEM-FONDO-PRATO nella finestra indicata sopra compaiono 3 oggetti digitali FP0002, FP0004 e FP0005. Ognuno di essi accanto presenta lo stato in cui si trova. Gli stati hanno il seguente significato:

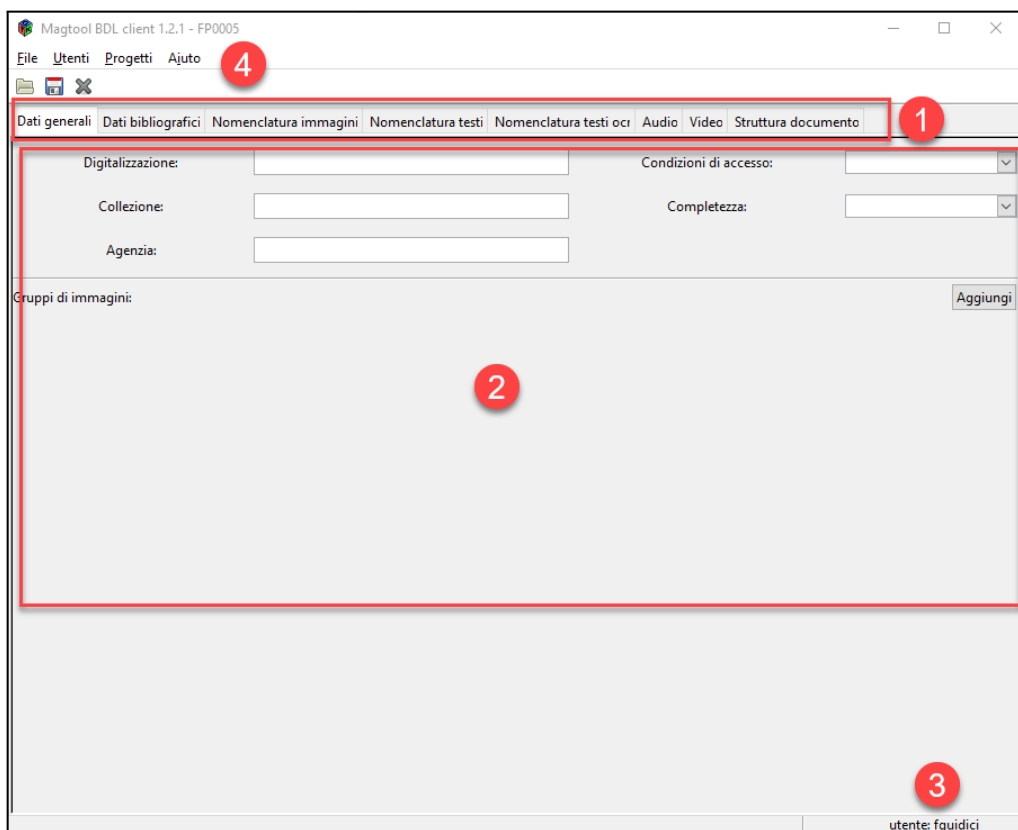
- *“da correggere”*: significa che l'oggetto era stato pubblicato in precedenza, ma è stato riportato qui nuovamente per correggere alcuni

metadati che erano scorretti (vedere la sezione “Modifica di un oggetto già pubblicato sulla DL”)

- “**da completare**”: questo stato compare quando si è iniziato a compilare i metadati di un oggetto, ma non si è riusciti a terminare. Si è dunque salvato il file dei metadati per poi riprenderlo successivamente.
- “**da iniziare**”: questo è lo stato che hanno tutti i nuovi oggetti che non hanno ancora un file di metadati cominciato.

Elementi fondamentali dell'interfaccia del MAGTOOL

Per seguire l'esempio apriremo l'oggetto FP0005. Una volta cliccato su “Apri” avremo di fronte l'interfaccia del programma che sarà così composta (vedere i numeri sull'immagine seguente):



1. Le schede in cui vengono raggruppati i metadati da compilare. Le schede seguono la denominazione delle sezioni dello standard MAG.

2. L'area centrale che a seconda della scheda scelta presenterà i campi dei metadati da compilare

1. Le schede in cui vengono raggruppati i metadati da compilare. Le schede seguono la denominazione delle sezioni dello standard MAG.
2. L'area centrale che a seconda della scheda scelta presenterà i campi dei metadati da compilare

3. il proprio username in basso a destra che cliccato permette di cambiare la password del proprio account MAGTOOL
4. la barra dei menu dove sono importanti in particolare:
 - a. il **pulsante di salvataggio** a forma di dischetto per salvare man mano quello che si sta facendo
 - b. e il **menu File** che vedremo più avanti e offre una serie di funzioni utili sia durante la compilazione dei campi che al termine per la verifica del file stesso e l'esportazione nella TECA.

Sezioni (schede) del MAGTOOL

L'applicativo MAGTOOL come abbiamo visto sopra si compone di diverse schede ciascuna delle quali rappresenta una sezione dello schema di metadati MAG adottato da Unito.

Nel seguito vediamo quali campi valorizzare (e come) nelle varie schede.

Sezione DATI GENERALI (corrispondente all'elemento GEN del MAG)

Riferimento ufficiale: capitolo “4.2 Sezione GEN” delle [linee guida per la DL Unito](#)

La scheda “**Dati generali**” è quella che contiene i metadati comuni a tutti gli oggetti di una determinata collezione / progetto.

Se è stato impostato il riempimento di default della scheda come da [istruzioni indicate più sopra](#) ogni qual volta si apre un oggetto digitale per la sua metadattazione questa scheda non ha bisogno di essere riempita perchè il Magtool provvede automaticamente.

E' comunque possibile modificarne i dati qualora siano scorretti (anche se si consiglia di farlo nelle [impostazioni di default](#)).

Ripetiamo qui i campi e come vanno valorizzati:

CAMPO	VALORE / COMMENTO
Digitalizzazione	Qui va messa l'URL della pagina del sito web di presentazione del progetto. In sua assenza valorizzare con l'URL principale del sito web della biblioteca.
Collezione	Qui va messa l'URL della collezione come impostata sul CMS (vedere il manuale operativo dedicato). Per il Fondo Prato questa è: <i>https://dl.unito.it/it/collezioni/collezione/fondo-giuseppe-prato/</i> (attenzione alla precisione della url - inclusa la barra terminale ed eventuali spazi - altrimenti l'oggetto non comparirà nella collezione).
Agenzia	valorizzare secondo quanto indicato nelle Linee Guida per la DL Unito nella "Appendice A: Valori per il tag <agency> relativi alle biblioteche dello SBA dell'Università di Torino". Nel caso della BEM inserire la stringa: <i>IT:TO0528</i>
Condizioni di accesso (uso riservato: 0 o pubblico: 1)	scegliere 1
Completezza (digitalizzazione completa: 0, non completa: 1)	scegliere 0

Sezione DATI BIBLIOGRAFICI (corrispondente all'elemento BIB del MAG)

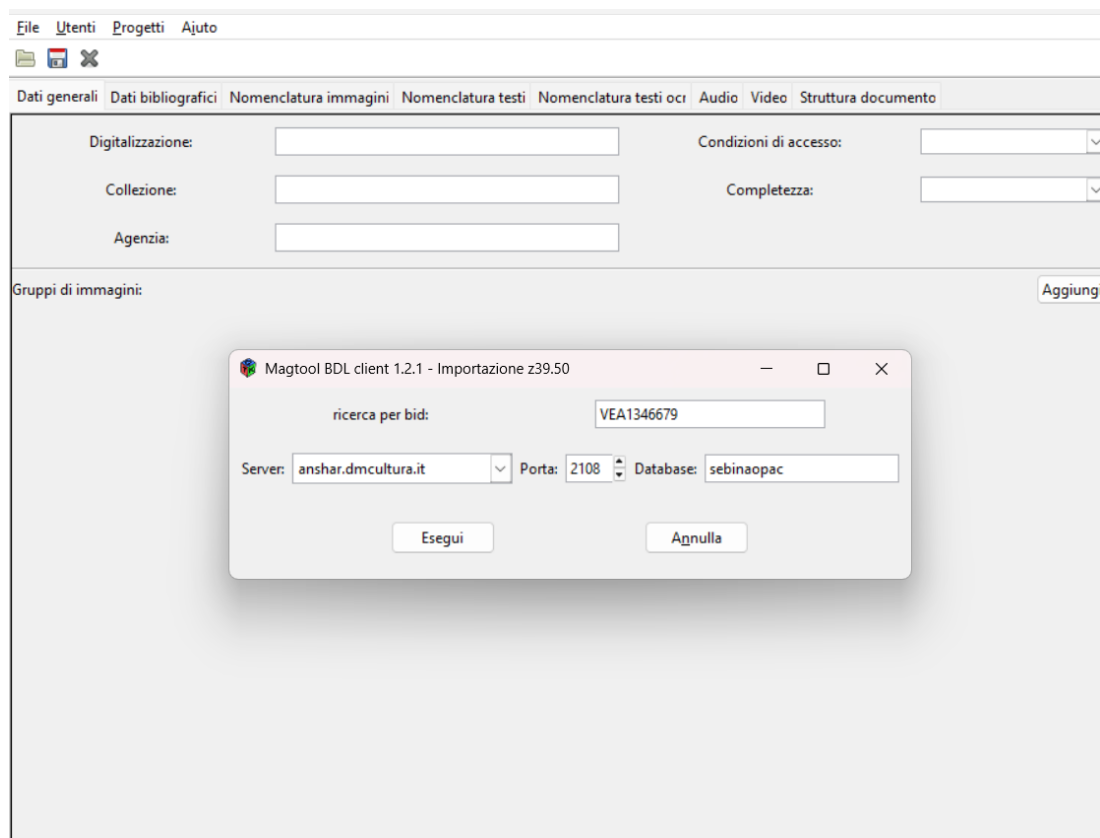
Riferimento ufficiale: capitolo “4.3 Sezione BIB” delle [linee guida per la DL Unito](#)

Questa sezione, come dice il nome, è quella che raccoglie i metadati bibliografici del volume, incluse alcune informazioni sul posseduto fisico.

La **valorizzazione dei campi di questa scheda deve essere fatta automaticamente** prelevandoli dall'**OPAC Sebina**.

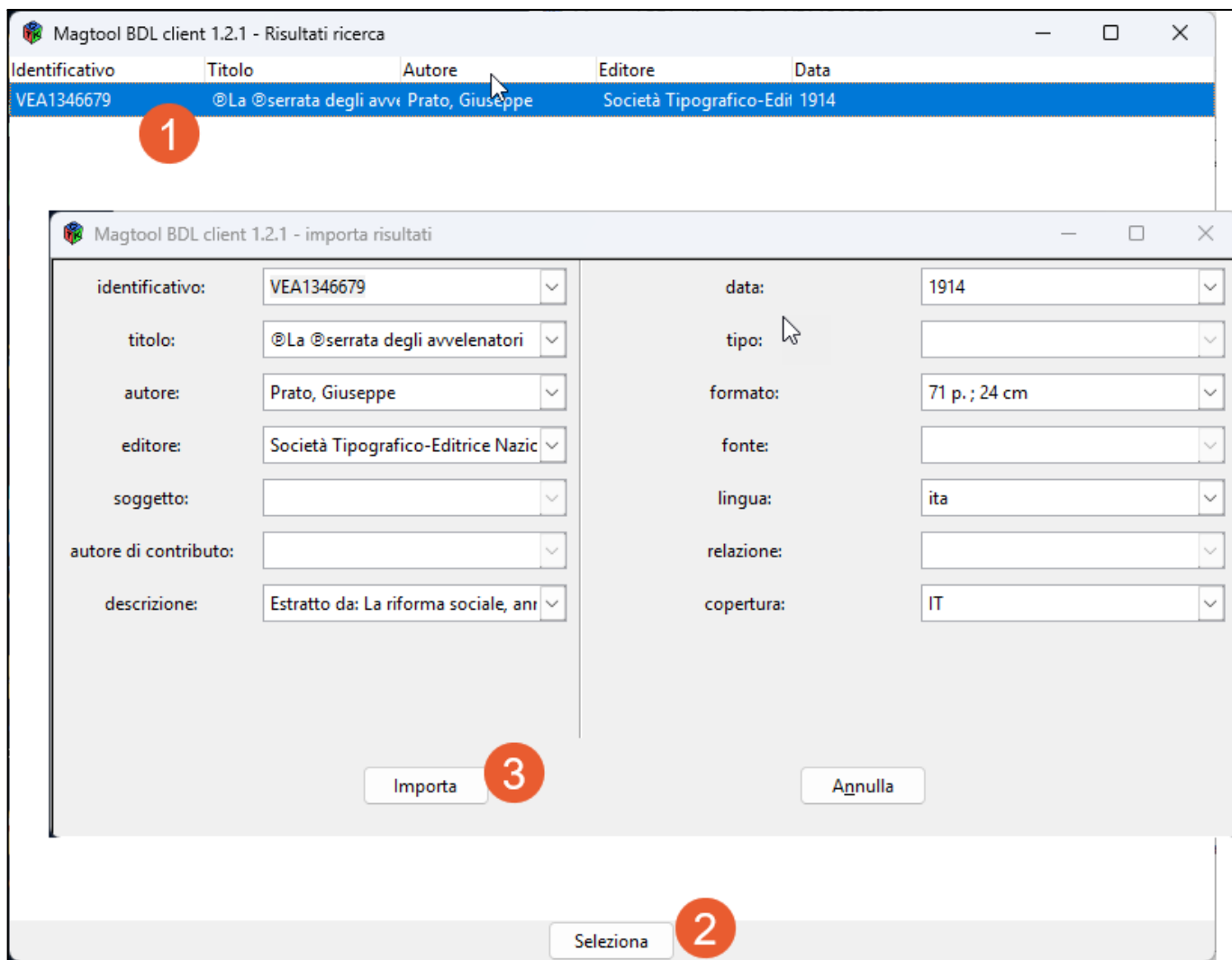
Per avvalersi di questa funzione, bisogna avere **configurato correttamente l'OPAC** nella sezione apposita ([vedere le istruzioni dedicate](#)).

Per recuperare dall'**OPAC Sebina** i metadati di un volume selezionare dal menu “**File**” del Magtool la funzione “**Ricerca su server Z39.50**” e inserire nel campo “Identificativo” della finestra che compare il BID del volume che si desidera cercare:



The screenshot shows the Magtool BDL client interface. The main window has a menu bar with 'File', 'Utenti', 'Progetti', and 'Ajuto'. Below the menu bar is a toolbar with icons for file operations. The main area is divided into tabs: 'Dati generali', 'Dati bibliografici', 'Nomenclatura immagini', 'Nomenclatura testi', 'Nomenclatura testi ocr', 'Audio', 'Video', and 'Struttura documento'. The 'Dati bibliografici' tab is active, showing fields for 'Digitalizzazione:', 'Collezione:', 'Agenzia:', 'Condizioni di accesso:', and 'Completezza:'. A dialog box titled 'Magtool BDL client 1.2.1 - Importazione z39.50' is open, showing a search form with the following fields: 'ricerca per bid:' (containing 'VEA1346679'), 'Server:' (containing 'anshar.dmcultura.it'), 'Porta:' (containing '2108'), and 'Database:' (containing 'sebinaopac'). The dialog box has 'Esegui' and 'Annulla' buttons.

Cliccando su esegui, il Magtool cercherà sull'OPAC il volume corrispondente e se lo trova proporrà un elenco di risultati (solitamente uno solo: il volume cercato!). Selezionare il volume (1 nell'immagine seguente) e cliccare su "Seleziona" (2 nell'immagine seguente) e successivamente nella finestra di riepilogo dei dati che stanno per essere importati su "Importa" (3 nell'immagine seguente).



The screenshot shows two windows from the Magtool BDL client 1.2.1. The top window, titled "Risultati ricerca", displays a table of search results. The first result is highlighted in blue and marked with a red circle containing the number 1. The bottom window, titled "importa risultati", shows the details of the selected record with fields for various metadata, marked with a red circle containing the number 3. At the bottom of the main interface, a button labeled "Seleziona" is marked with a red circle containing the number 2.

Identificativo	Titolo	Autore	Editore	Data
VEA1346679	@La @serrata degli avve	Prato, Giuseppe	Società Tipografico-Edi	1914

identificativo:	VEA1346679	data:	1914
titolo:	@La @serrata degli avvelenatori	tipo:	
autore:	Prato, Giuseppe	formato:	71 p.; 24 cm
editore:	Società Tipografico-Editrice Nazic	fonte:	
soggetto:		lingua:	ita
autore di contributo:		relazione:	
descrizione:	Estratto da: La riforma sociale, anr	copertura:	IT

Buttons: Importa, Annulla, Seleziona

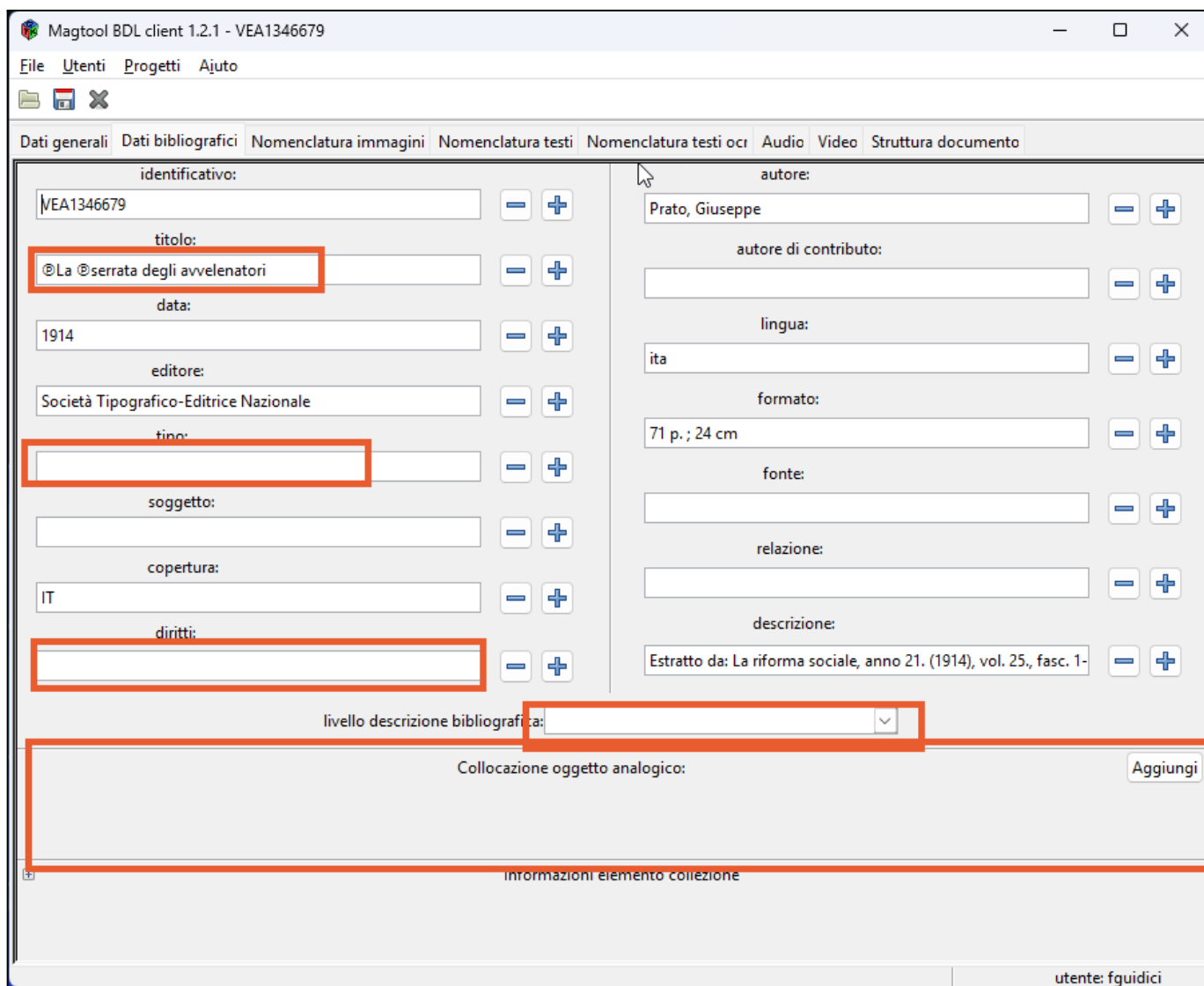
I dati bibliografici prelevati dall'**OPAC Sebina** popoleranno la scheda Dati Bibliografici del Magtool (vedere immagine seguente).

La modifica dei metadati recuperati automaticamente dall'**OPAC Sebina** è consentita in misura minimale qualora si debba correggere la presenza di "caratteri strani" (dovuti ad una scorretta mappatura di caratteri particolari presenti nelle stringhe di metadati prelevati dall'OPAC) oppure per l'introduzione di spazi mancanti al loro interno.

Aggiungere o correggere l'editore se manca o presenta degli errori, o correggere il titolo

(qualora ad esempio vi siano dei caratteri “strani” o venga recuperato il solo numero del volume, etc..). Qui sotto **un esempio** di campi da modificare nell’esempio specifico:

- il titolo riporta dei caratteri “strani”
- il campo “Tipo” va valorizzato a “Testo”
- il campo “Diritti” va valorizzato con l'url della licenza da assegnare al volume ([vedere linee guida](#))
- Il livello bibliografico va selezionato in modo opportuno ([vedere linee guida](#))
- i campi che compariranno scegliendo il livello bibliografico andranno compilati ([vedere linee guida](#))



Magtool BDL client 1.2.1 - VEA1346679

File Utenti Progetti Aiuto

Dati generali Dati bibliografici Nomenclatura immagini Nomenclatura testi Nomenclatura testi ocr Audio Video Struttura documento

identificativo: VEA1346679

titolo: ©La ©serrata degli avvelenatori

data: 1914

editore: Società Tipografico-Editrice Nazionale

tipo:

soggetto:

copertura: IT

diritti:

autore: Prato, Giuseppe

autore di contributo:

lingua: ita

formato: 71 p.; 24 cm

fonte:

relazione:

descrizione: Estratto da: La riforma sociale, anno 21. (1914), vol. 25., fasc. 1-

livello descrizione bibliografica:

Collocazione oggetto analogico: Aggiungi

Informazioni elemento collezione

utente: fguidici

Di seguito una tabella con l'indicazione dei campi principali e come sono valorizzati:

CAMPO	VALORE / COMMENTO
Identificativi	Precompilati da ricerca z39.50. Riportano il BID e l'ID Sebina del volume digitalizzato. Se verrà modificato uno dei due identificativi, non verrà creato un oggetto nuovo (fatta eccezione per i periodici e per gli oggetti nei cui mag viene valorizzata la sezione <piece>)
Titolo	Precompilato da ricerca z39.50. Riporta il titolo della pubblicazione.
Data	Precompilato da ricerca z39.50. Riporta la data di edizione.
Editore	Precompilato da ricerca z39.50. Riporta l'editore della pubblicazione.
Tipo	Valorizzare con il termine "Testo"
Soggetto	Precompilato da ricerca z39.50. Riporta il soggetto o i soggetti come appaiono sull'OPAC SBN. Controllare che i soggetti siano correttamente spaziati. In caso negativo provvedere ad inserire gli spazi o i separatori necessari. NOTA: i record con soggetto in OPAC SBN sono pochissimi e quindi è possibile che diversi oggetti non abbiano soggetti correlati.
Copertura	Precompilato da ricerca z39.50.
Diritti	Va obbligatoriamente valorizzato a mano riportando i diritti associati all'uso della risorsa. L'elemento

	dovrà essere popolato con la URL della licenza con cui è rilasciato l'oggetto digitale. Se l'opera è in Pubblico Dominio: https://creativecommons.org/publicdomain/mark/1.0/ Se l'opera è coperta da diritto d'autore si dovrà optare per una licenza d'uso più opportuna e rispettosa del copyright in vigore, da determinare caso per caso. Ad esempio: http://creativecommons.org/licenses/by-nc-sa/4.0/
Autore	Precompilato da ricerca z39.50. Riporta nome e cognome dell'autore principale nella forma: Cognome, Nome
Autore di contributo	Precompilato da ricerca z39.50. Riporta nome e cognome dell'autore secondario formattato, se presente, come per l'autore.
Lingua	Precompilato da ricerca z39.50.
Formato	Precompilato da ricerca z39.50.
Fonte	Precompilato da ricerca z39.50.
Relazione	Precompilato da ricerca z39.50.
Descrizione	Precompilato da ricerca z39.50.
Livello descrizione bibliografica	Selezionare il valore "m" (monografia) oppure il valore "s" (periodici). NOTA: nel caso dei periodici si rimanda alle Linee Guida per la DL Unito al capitolo "Sezione BIB per i fascicoli di periodico" per informazioni su come compilare la parte sottostante.

Nella parte sottostante della finestra denominata “**Collocazione oggetto analogico**”, bisognerà cliccare su “**Aggiungi**” e compariranno dei campi da valorizzare come segue:

CAMPO	VALORE / COMMENTO
Istituzione proprietaria	Inserire “ <i>Biblioteca di Economia e Management, Università di Torino, Italia</i> ”
Collocazione oggetto	Inserire la collocazione del volume (es. FP.0002)
Numero inventario	Inserire l'inventario del volume (es. IFP 2859) - notare lo spazio
Identificativo istituzione	Lasciare vuoto
Tipo di collocazione	Lasciare vuoto

FASCICOLI DI PERIODICI: nel caso in cui l’oggetto digitale che si sta trattando fosse un periodico, si rimanda alla [linee guida per la DL Unito](#) al capitolo “**Sezione BIB per i fascicoli di periodico**” per la specificazione su come valorizzare i campi di questa sezione.

Sezioni di NOMENCLATURA IMMAGINI, TESTI e TESTI OCR

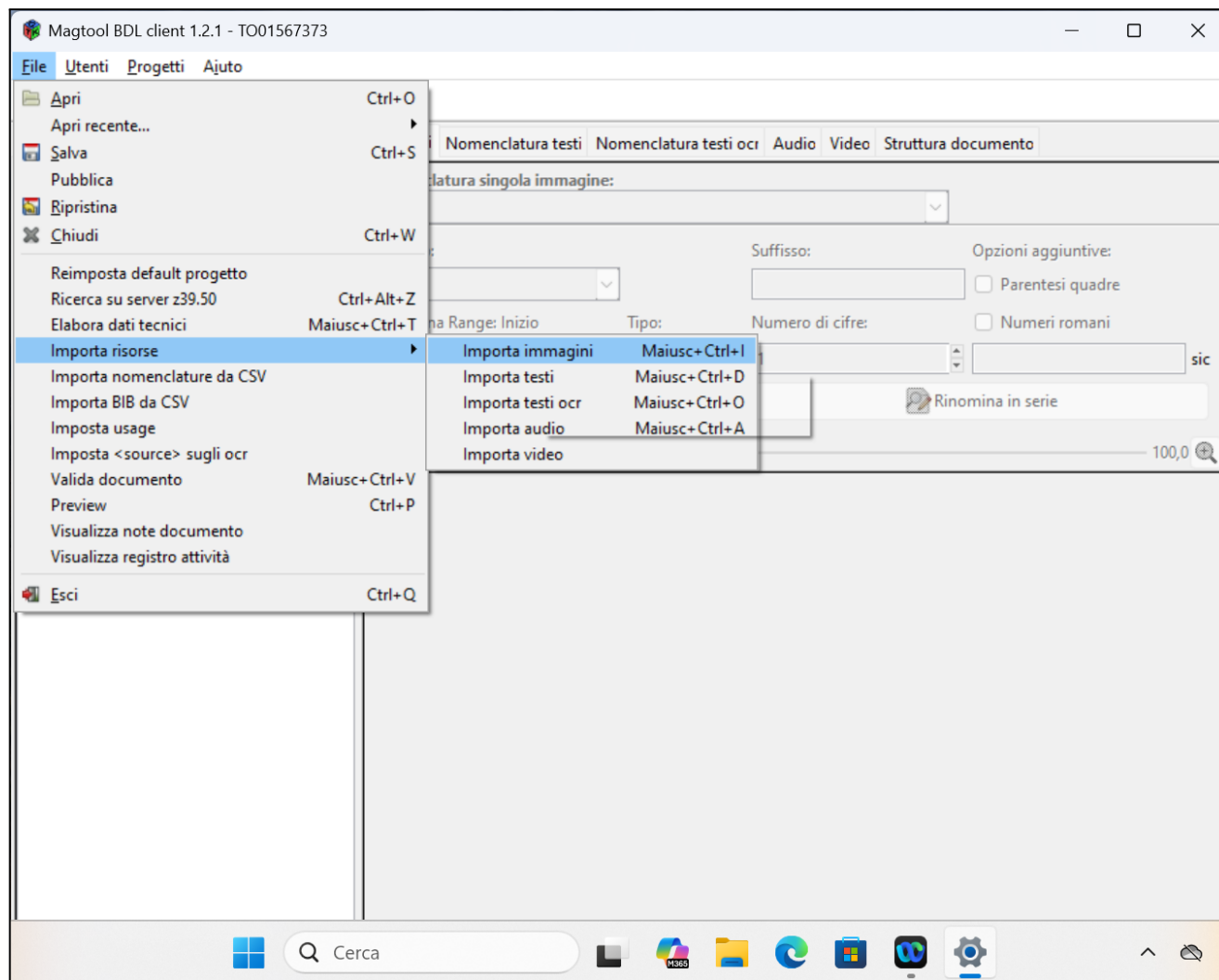
Riferimento ufficiale: capitolo “**4.4 Sezione IMG**” delle [linee guida per la DL Unito](#)

Nelle 3 schede seguenti “Nomenclatura immagini”, “Nomenclatura testi” e “Nomenclatura testi ocr” andranno importati rispettivamente le immagini dell’oggetto digitale, il suo pdf completo e i files dell’OCR dell’oggetto.

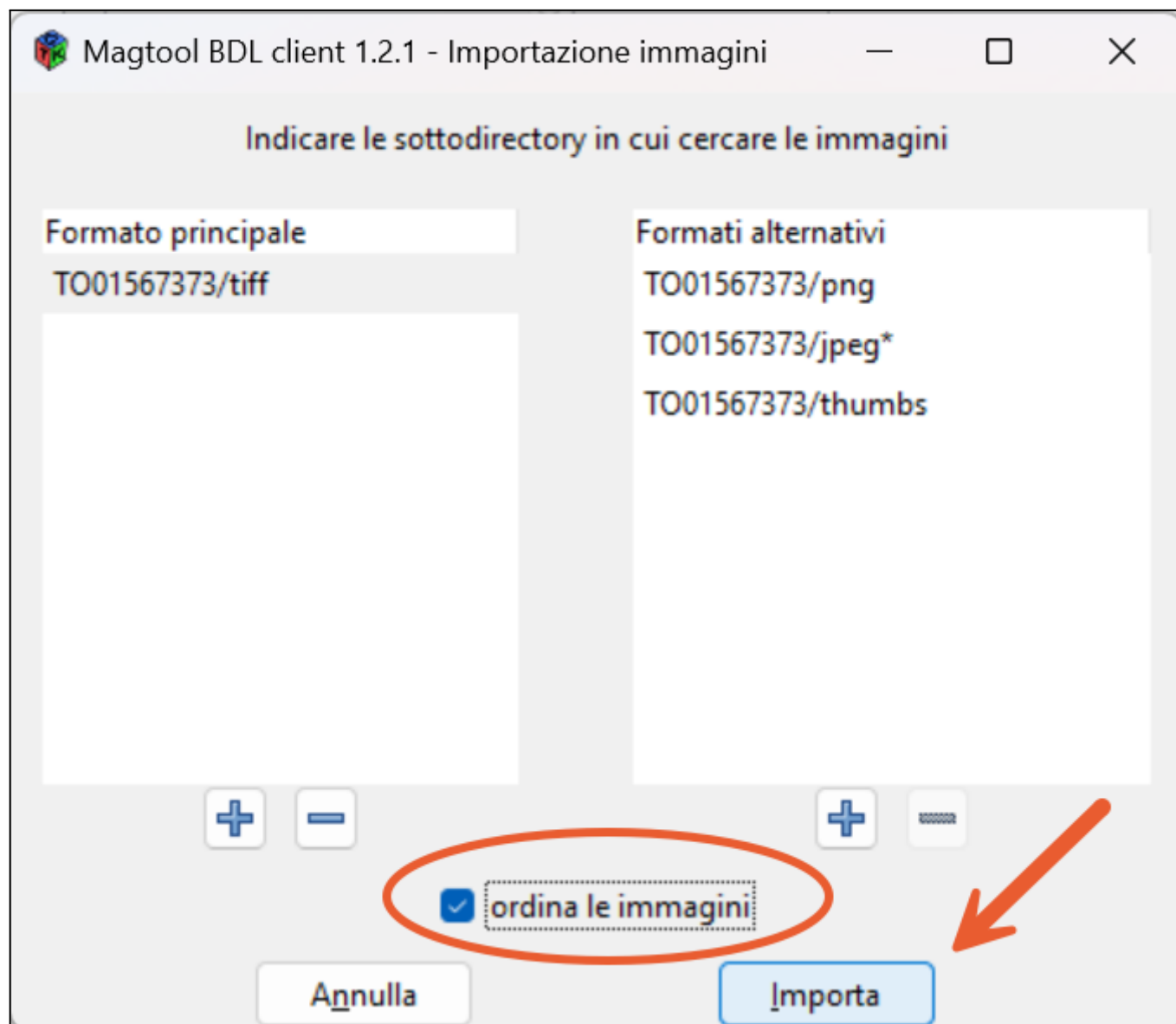
Vediamo **come importare tutti e 3 questi elementi** e successivamente come “nomenclarli” ai fini della metadatazione.

IMMAGINI: import

Selezionare dal menu del Magtool *File* → *Importa risorse* → *Importa immagini*.



Nella finestra che compare selezionare la casella “*ordina le immagini*” e cliccare su *Importa*.

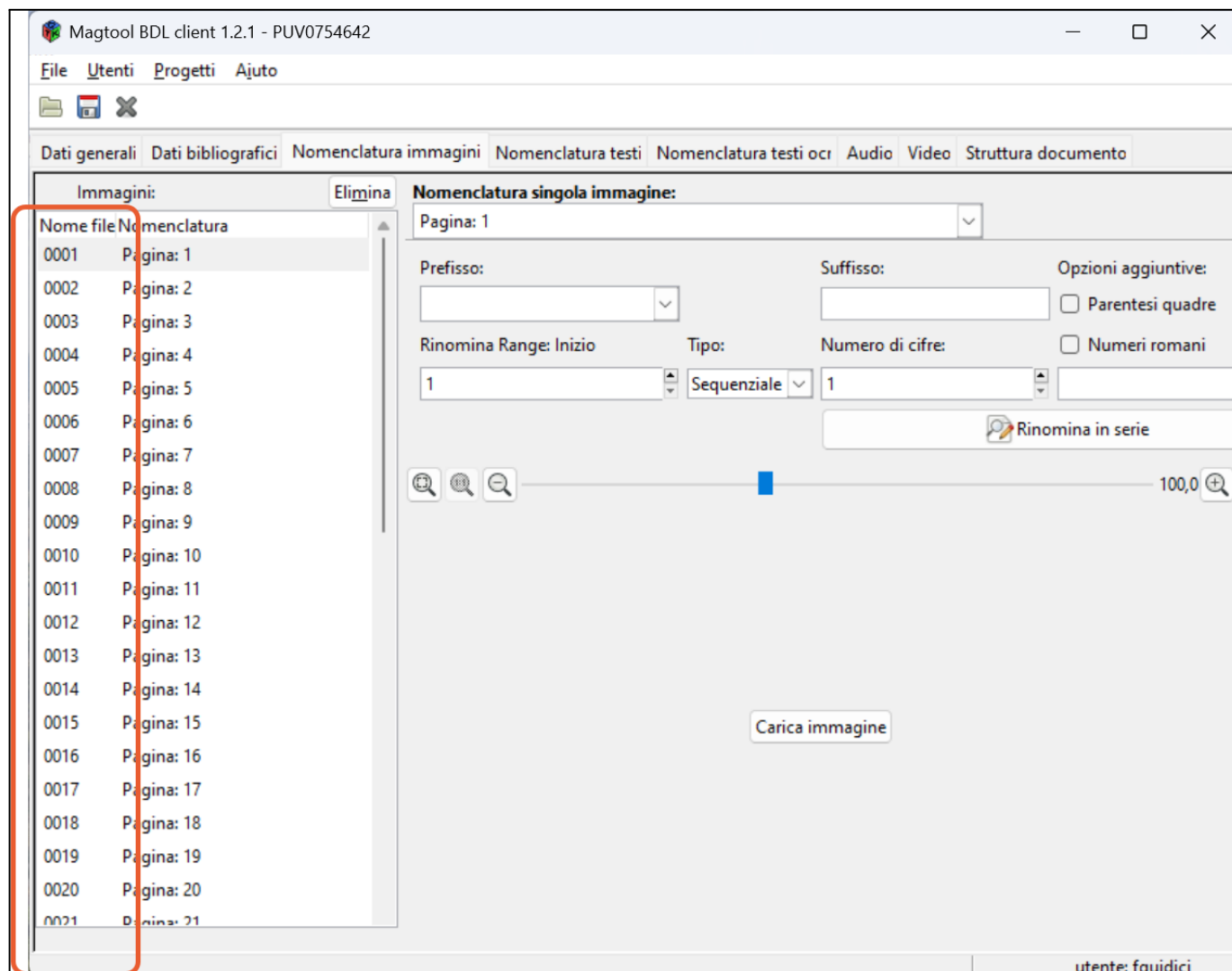


Se compare il messaggio che conferma l'importazione di N immagini l'operazione è andata a buon fine altrimenti bisogna risolvere il problema.

Il valore N delle immagini deve corrispondere esattamente al numero di file immagine creati durante la scansione inclusi Color Chart e Dorso del libro.

Esempio: se il volume digitalizzato è costituito da 20 pagine (comprehensive di piatti e contropiatti anteriori e posteriori) + Dorso + Color Chart il totale delle immagini sarà 22

Una volta importate le immagini esse compaiono elencate nella colonna di sinistra della scheda ciascuna con il proprio nome del file:



IMMAGINI: nomenclatura

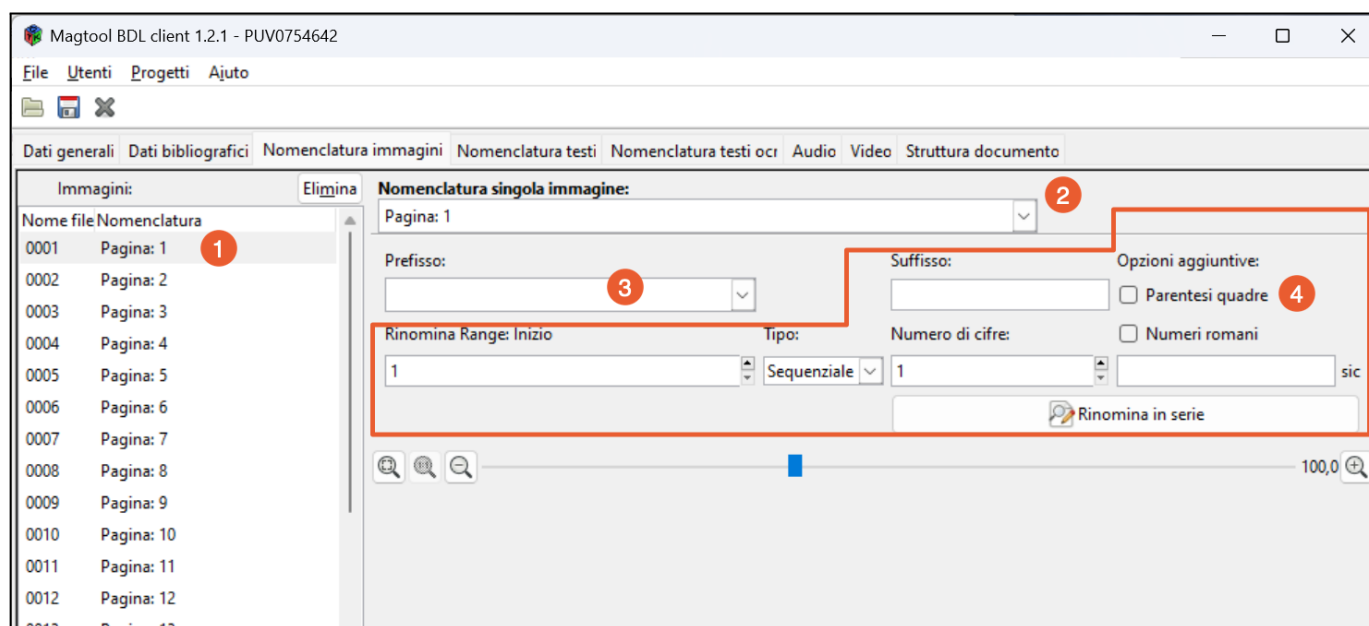
Per dare un nome alle immagini bisogna prima selezionarle nella colonna di destra (numero 1 nell'immagine seguente). Possono essere selezionate una ad una oppure può essere selezionato un range di immagini che devono avere un nome simile (per farlo selezionare la prima immagine del range e, tenendo premuto il pulsante SHIFT, selezionare l'ultima del range).

Con riferimento all'immagine seguente è possibile assegnare un nome alle immagini, operando con le funzioni della parte centrale della scheda:

- Il menu 2 serve per dare all'immagine il nome corrispondente alla parte fisica del volume che è stata digitalizzata (ad esempio per nominare piatti e contropiatti, dorso, color chart, etc...), qualora questa non sia numerabile/numerata;

- Il menu 3 ha la stessa funzione del menu 2, ma si usa per le parti fisiche numerabili/numerate, ad esempio “Pagina”, “Carta”; l’etichetta scelta viene trattata come prefisso di una stringa che si completa con ciò che si definisce nella sezione 4;
- La sezione 4 serve a completare l’etichetta scelta nel menu 3 con un numero e a indicare se deve essere un numero romano, posto o meno tra parentesi quadre; se si seleziona un range di immagini è possibile applicare a tutte i criteri scelti e numerarle consecutivamente.

Esempio: se devo numerare le pagine con nome file da 4 a 10 con i numeri pagina da 1 a n consecutivi, selezionerò le pagine come indicato sopra nella colonna di sinistra, sceglierò la voce nel campo “Prefisso”, imposterò il numero 1 nel campo “Rinomina Range: Inizio”, lascerò selezionato il termine “Sequenziale” nel campo “Tipo” e cliccherò sul pulsante “Rinomina in serie”



Quando si ha un'immagine selezionata nella colonna di sinistra, nel centro della scheda è disponibile il pulsante “Carica immagine” che consente di vedere una anteprima (zoomabile) della pagina digitalizzata.

Questa funzione è un ottimo modo di controllare che la nomenclatura che viene data corrisponda effettivamente alla pagina in questione. Se così non fosse bisogna verificare dove stia l'errore che diversamente si rifletterà sull'oggetto pubblicato nella DL.

I valori che da [linee guida Unito](#) (capitolo “4.3 Sezione IMG”) vengono adottati per le nostre digitalizzazioni è costituito dal seguente sottoinsieme di quelle disponibili:

• Pagina • Pagina bianca • Pagina non numerata • Pagina bianca non numerata • Carta • Carta bianca • Carta non numerata • Carta bianca non numerata	• Tavola • Piatto anteriore • Contropiatto anteriore • Piatto posteriore • Contropiatto posteriore • Dorso • Color Chart
--	--

PDF: import

Selezionare dal menu del Magtool *File* → *Importa risorse* → *Importa testi*.

Nella finestra che compare (analoga a quella delle immagini) non è necessario selezionare la casella “*ordina i testi*” (ne andremo ad importare uno solo: l’unico pdf disponibile).

Prima di procedere **cancellare dalla finestra i due percorsi doc e txt e lasciare presente solo il percorso pdf** altrimenti anziché importare solo il pdf si importeranno anche i file txt dell’OCR ed è un errore.

Cliccare su *Importa*

Se compare il messaggio che conferma l’importazione di 1 testo l’operazione è andata a buon fine altrimenti bisogna risolvere il problema.

PDF: nomenclatura

Selezionare l’unico file PDF elencato nella colonna a sinistra e lasciare tutti i campi invariati tranne il campo “Nomenclatura singolo testo” che va valorizzato inserendo come nome il **BID dell’oggetto**.

Anche per il PDF è disponibile come per le immagini il pulsante che consente di caricare un’anteprima del file.

OCR: import

Selezionare dal menu del Magtool *File* → *Importa risorse* → *Importa testi ocr*.

Nella finestra che compare selezionare la casella “*ordina le immagini*”.

Prima di procedere **cancellare dalla finestra i due percorsi doc e pdf e lasciare presente solo il percorso txt.**

Cliccare su *Importa*.

Se compare il messaggio che conferma l'importazione di N testi ocr l'operazione è andata a buon fine altrimenti bisogna risolvere il problema. Gli N testi ocr devono essere uguali alle immagini della scheda ad esse relativa.

OCR: nomenclatura

La nomenclatura dei testi OCR è speculare a quella delle immagini.

Procedere quindi nello stesso modo fatto per le immagini. Usando la stessa nomenclatura in quanto ogni ocr è associato alla relativa immagine e quindi ne avrà lo stesso nome.

Sezione STRUTTURA DOCUMENTO (che corrisponde all'elemento STRU del MAG)

Riferimento ufficiale: capitolo “4.5 Sezione STRU” delle [linee guida per la DL Unito](#).

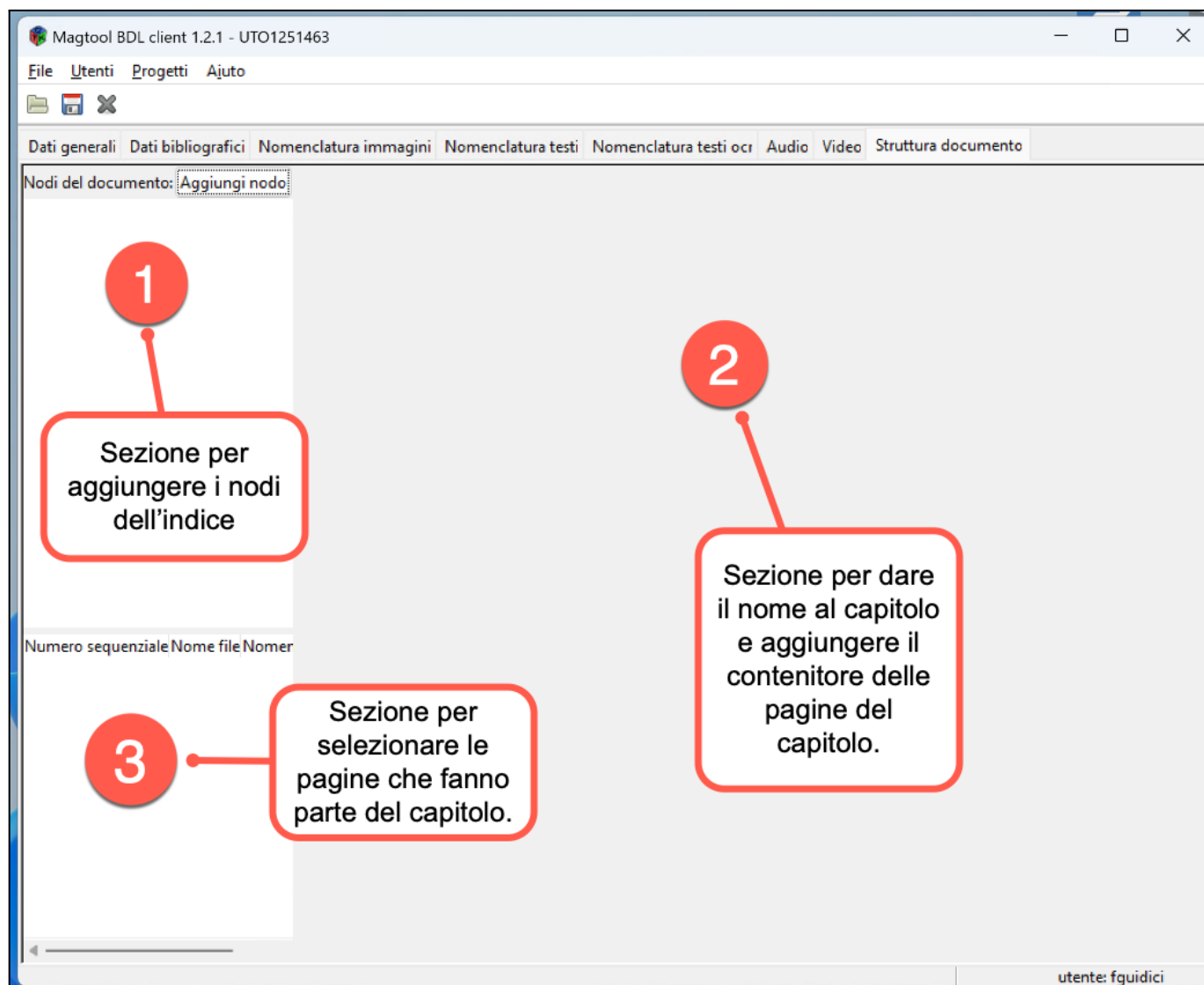
Questa scheda serve a creare l'indice del volume digitalizzato.

Fermo restando il mantenimento delle voci di primo livello, il livello di profondità cui addentrarsi per la creazione dell'indice può variare da testo e testo a va valutata di volta in volta.

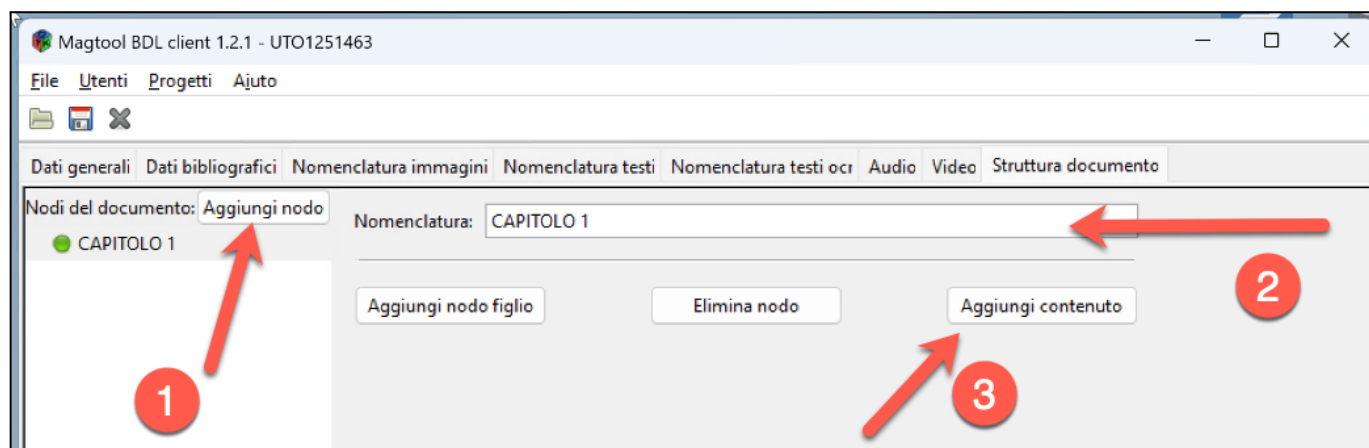
L'indice se opportunamente creato secondo le istruzioni seguenti sarà disponibile anche per l'utente in formato cliccabile quando l'oggetto verrà pubblicato.

NODI PRINCIPALI (voci di primo livello dell'indice)

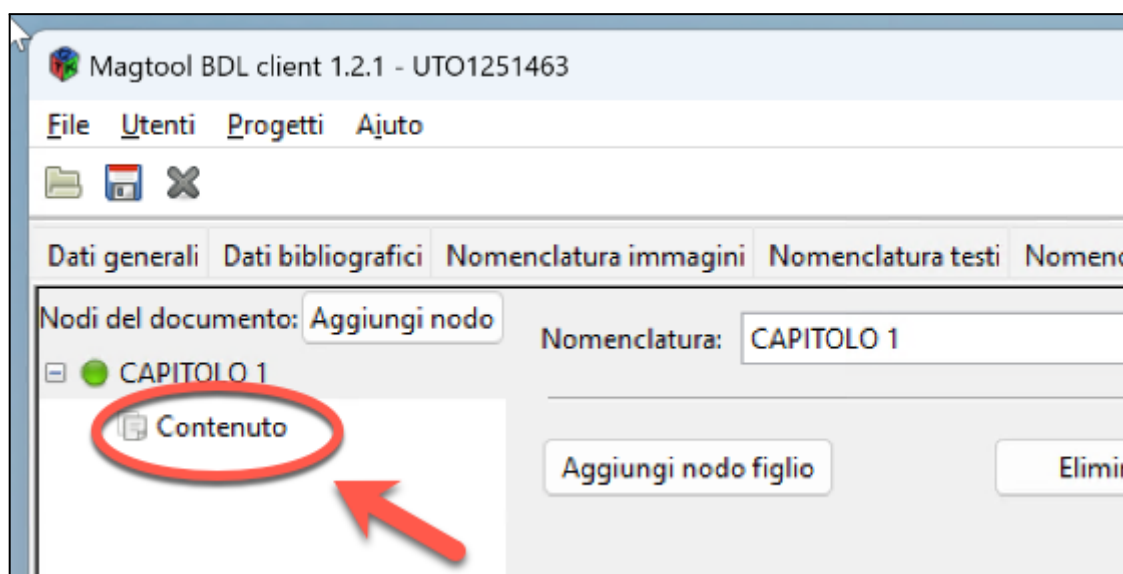
Per creare l'indice (o struttura) del documento aprire la **scheda “Struttura documento”** del Magtool. La scheda apparirà come segue (sono indicate le varie aree):



Per aggiungere i capitoli si procede come segue (fare riferimento all'immagine seguente):

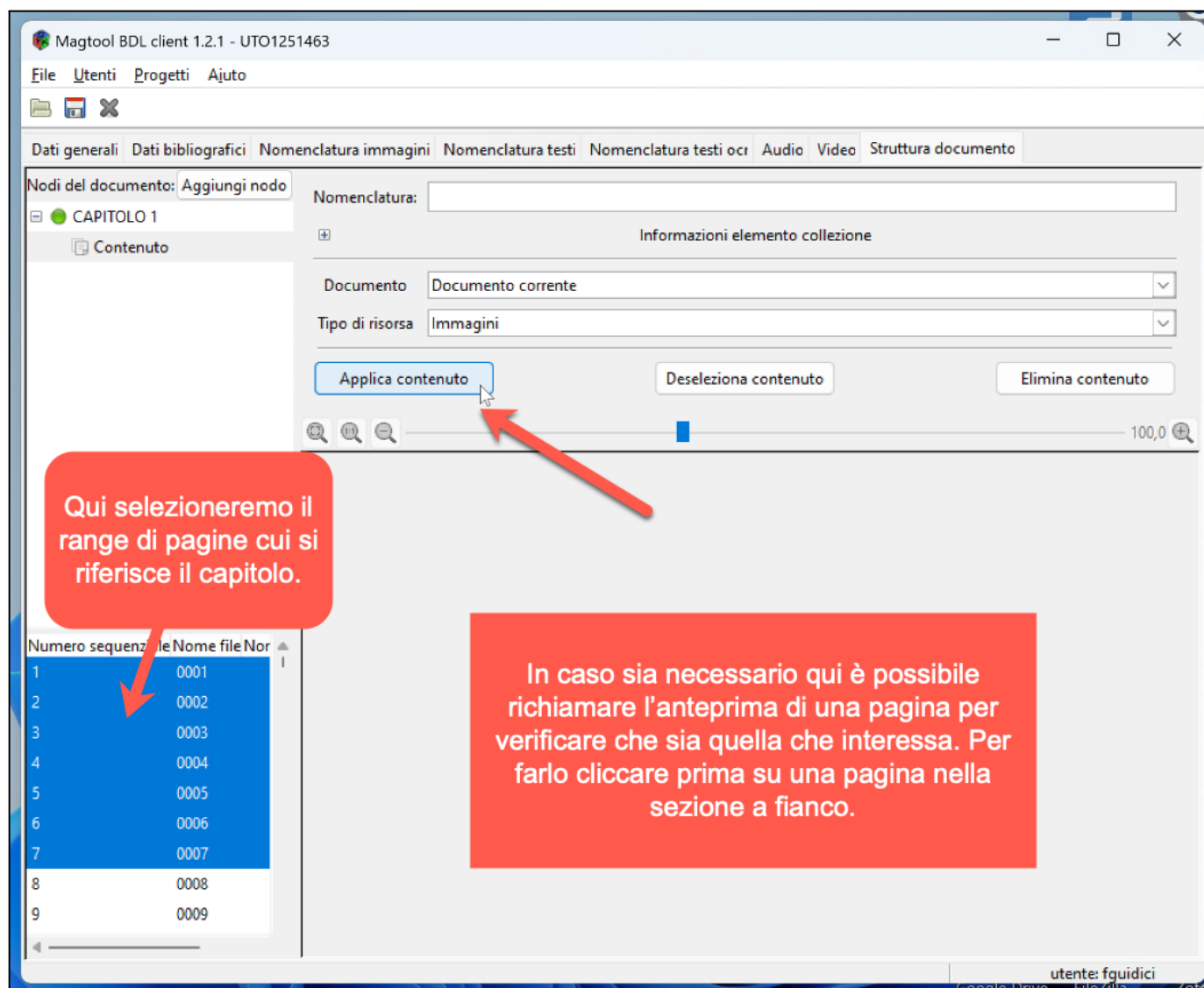


1. Cliccare nella colonna di sinistra sul pulsante “Aggiungi nodo” (1), verrà creata una voce vuota dell’indice che inizialmente sarà nominata genericamente “nodo strutturale”.
2. Selezionare il nodo appena creato e nel campo (2) della scheda centrale andare a dare un nome al capitolo nel campo "Nomenclatura". Il nome sarà il titolo del capitolo come rilevato dal libro fisico.
3. Cliccare sul pulsante “Aggiungi contenuto” (3) presente subito sotto al campo. Nella colonna di sinistra verrà creato un contenitore subito sotto al nodo/capitolo selezionato:



4. Selezionare la voce “Contenuto” appena creata e nel riquadro subito a quello dei nodi potrà essere selezionato il range di pagine a cui il capitolo si riferisce (dalla pagina iniziale a quella finale).
Per selezionare più pagine in un colpo solo cliccare sulla prima, tenere premuto il tasto SHIFT e selezionare l’ultima.
Tutto il range di interesse verrà selezionato. Cliccare a questo punto sul pulsante

“Applica contenuto” (vedere immagine seguente):



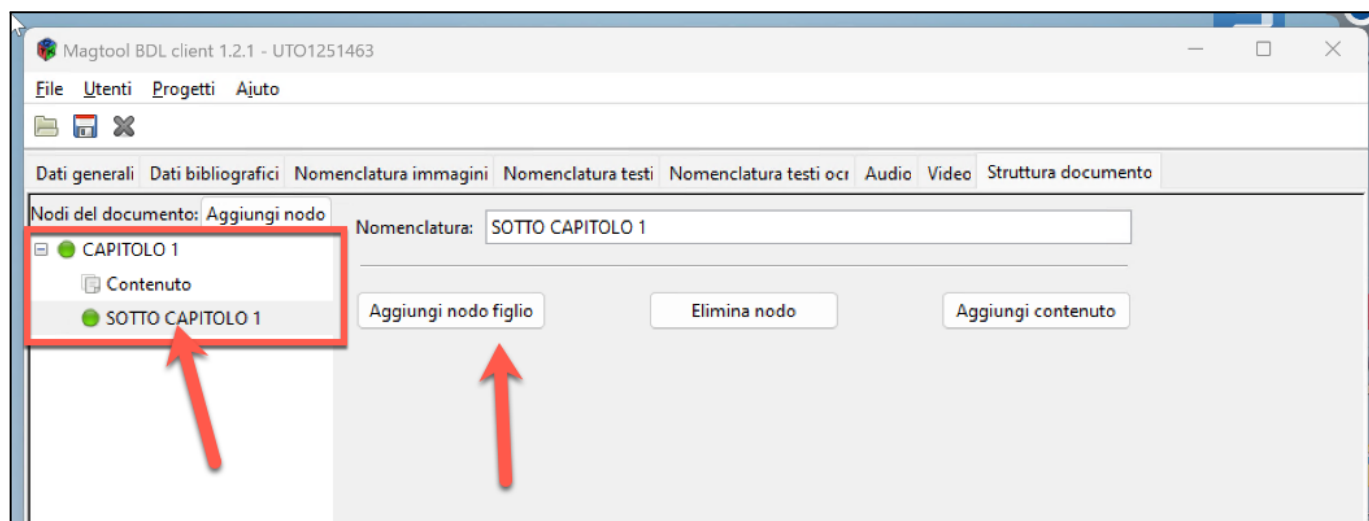
5. Ripetere i passaggi precedenti fino alla fine dell'indice.

NOTA: in caso di necessità è possibile caricare un'anteprima della pagina nella sezione di destra per verificare che si stia considerando la pagina iniziale e finale corrette.

SOTTO NODI (voci dal secondo livello in poi dell'indice)

È possibile annidare eventuali sotto capitoli all'interno di un capitolo seguendo gli stessi passaggi indicati sopra. Con l'unica accortezza che per aggiungere un sottonodo andrà

utilizzato il pulsante “Aggiungi nodo figlio” seguendo poi gli stessi passaggi indicati in precedenza per il nodo figlio così creato:



NOTA DI ATTENZIONE!!

Se viene creato l'indice del volume, eventuali errori commessi (ad esempio è stato creato un nodo ma non gli sono state assegnate delle pagine) non verrà segnalato dalla funzione “Valida documento” del Magtool (vedere più avanti).

Il documento sarà comunque valido per il Magtool ma quando con la TECA si pubblicherà l'oggetto sulla Digital Library verrà emesso un errore e l'oggetto non sarà pubblicato.

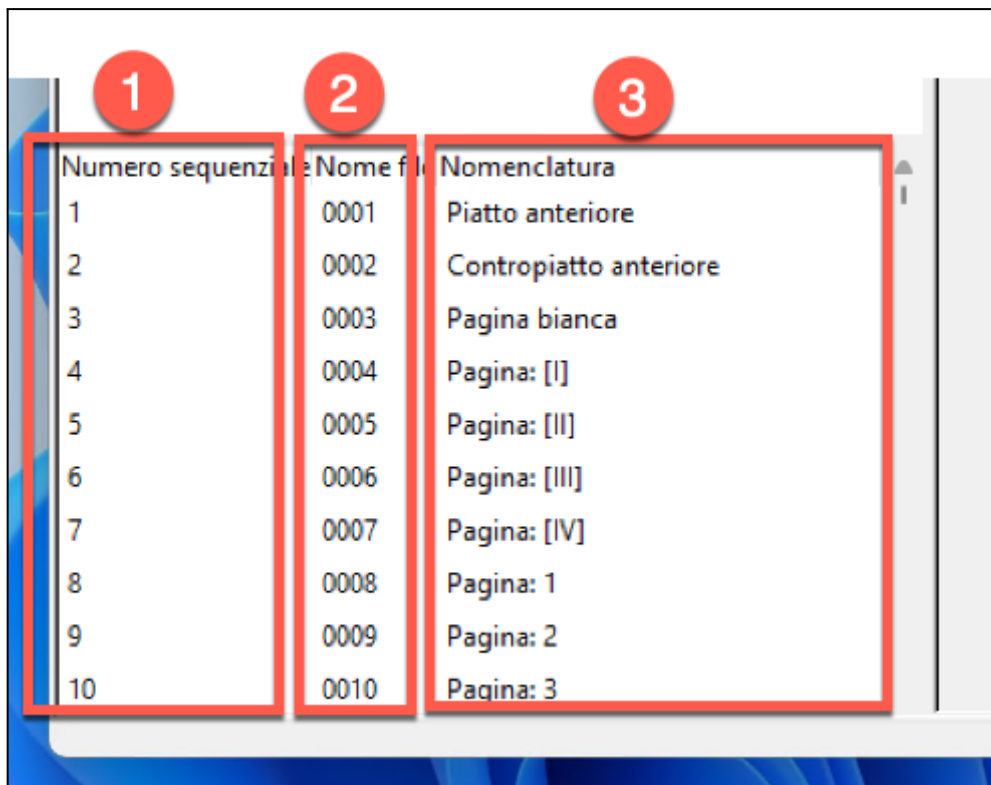
NOTA IMPORTANTE SULLA SELEZIONE DELLE PAGINE

Il box di selezione delle pagine presenta 3 colonne (vedere immagine di seguito):

1. la **colonna 1** è un numero sequenziale assegnato alle pagine (diverso dal numero di pagina e dal nome del file) che non ci interessa e che viene generato automaticamente dal Magtool per il file xml che conterrà i metadati.
2. la **colonna 2** presenta il nome del file dell'immagine assegnato in fase di digitalizzazione dal software dello scanner (o rinominato da noi con il software Advanced Renamer, se lo abbiamo fatto).
3. la **colonna 3** riporta invece il nome dell'immagine che abbiamo impostato noi nella durante la loro nomenclatura nella scheda “Nomenclatura immagini”.

E' la colonna 3 quella importante per la selezione delle pagine (ignorare le altre 2).

Se abbiamo fatto bene la nomenclatura, le pagine corrisponderanno a quelle del libro e quindi la selezione sarà più semplice ed accurata!



1	2	3
Numero sequenziale	Nome file	Nomenclatura
1	0001	Piatto anteriore
2	0002	Contropiatto anteriore
3	0003	Pagina bianca
4	0004	Pagina: [I]
5	0005	Pagina: [II]
6	0006	Pagina: [III]
7	0007	Pagina: [IV]
8	0008	Pagina: 1
9	0009	Pagina: 2
10	0010	Pagina: 3

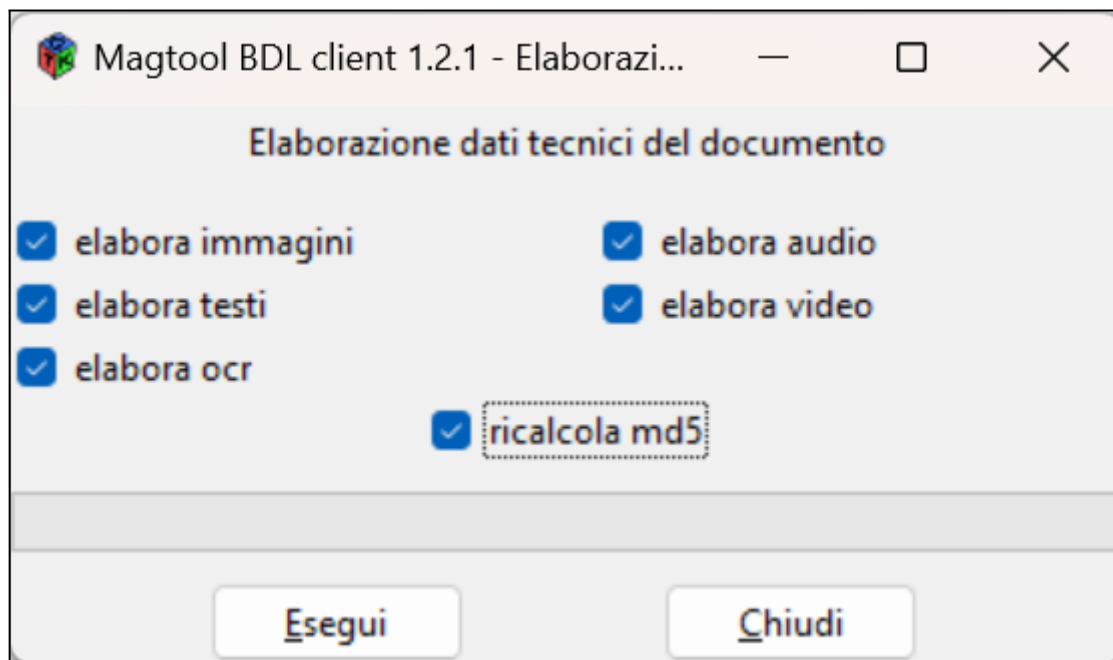
Fasi finali della metadatazione

Terminata la metadatazione effettiva è necessario effettuare alcune operazioni finali che descriviamo nel seguito.

Elaborare i DATI TECNICI

Selezionare dal menu del Magtool la voce *File* → *Elabora dati tecnici*.

Nella finestra che compare, lasciare tutte le caselle invariate selezionando solo la casella “**ricalcola md5**” e cliccare su “*Esegui*”



Si verrà avvisati che l’oggetto deve essere salvato e chiuso prima di procedere.

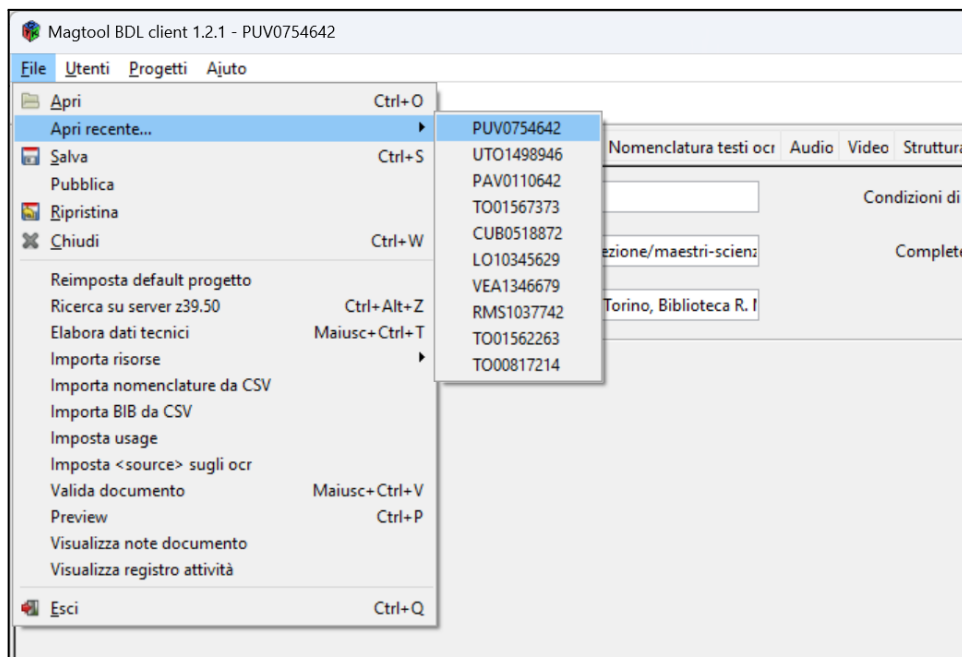
Confermare.

Si dovrà attendere qualche istante e poi lo si dovrà riaprire.

La durata di questa operazione dipende dalle dimensioni del volume digitalizzato. Può richiedere da pochi secondi a diversi minuti.

Per verificare quando l’operazione termina non c’è un avviso da parte del sistema.

Bisogna provare ad aprire il documento selezionando *File → Apri recente (o Apri) → selezionare il BID dell’oggetto appena trattato*:



Se compare un messaggio che l'oggetto è ancora in stato bloccato (locked) significa che l'operazione è ancora in corso.

Attendere e riprovare. Quando l'operazione sarà terminata l'oggetto potrà essere aperto nuovamente.

Impostare lo USAGE per immagini e documenti

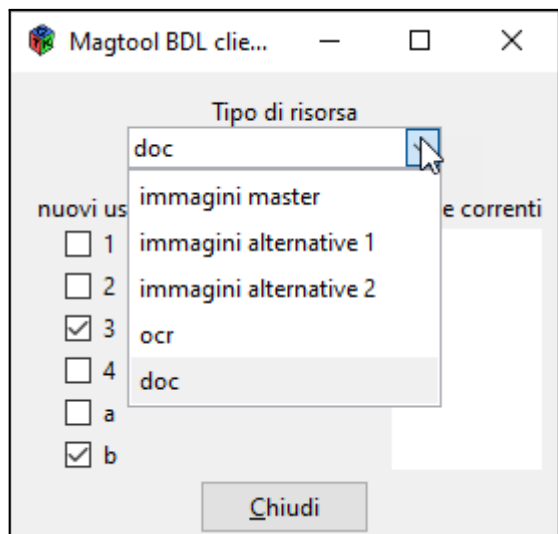
Prima di validare il documento è necessario impostare lo usage delle immagini e dei documenti.

Solitamente lo usage delle immagini viene impostato dallo strumento quando vengono importate: le immagini master (i TIFF) e le immagini alternative 1 e 2 (i JPEG) vengono impostati ai valori usage 1, usage 2 e usage 3 rispettivamente.

Vanno impostati manualmente gli usage dei testi OCR e di eventuali documenti allegati (es. il PDF).

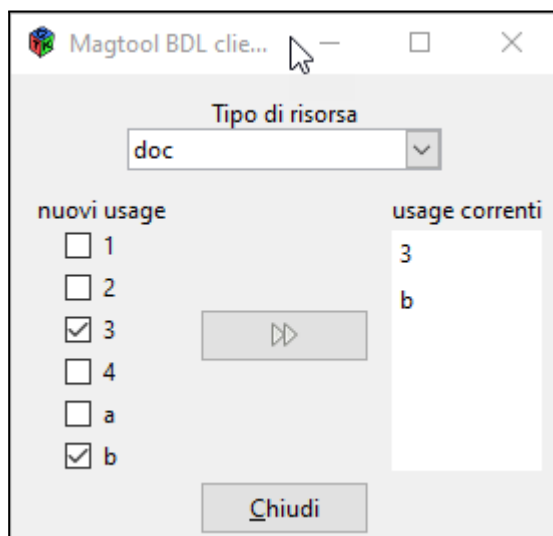
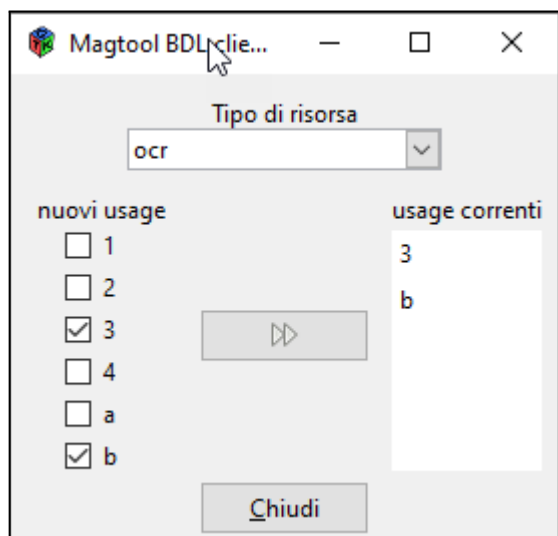
Per farlo cliccare su *File* → *Imposta usage*

Si aprirà la seguente finestra



dove, dal menu a tendina, andremo a selezionare prima “ocr” e successivamente “doc” e per ciascuno **andremo ad aggiungere il livello b allo usage 3 già preimpostato.**

Clicchiamo su “b” e poi sulla doppia freccia centrale. La “b” comparirà nel box di sinistra. Qui sotto due screenshot che rappresentano l’operazione per gli ocr e per i doc (ovvero il pdf):



Impostare il VALORE <SOURCE> per gli OCR

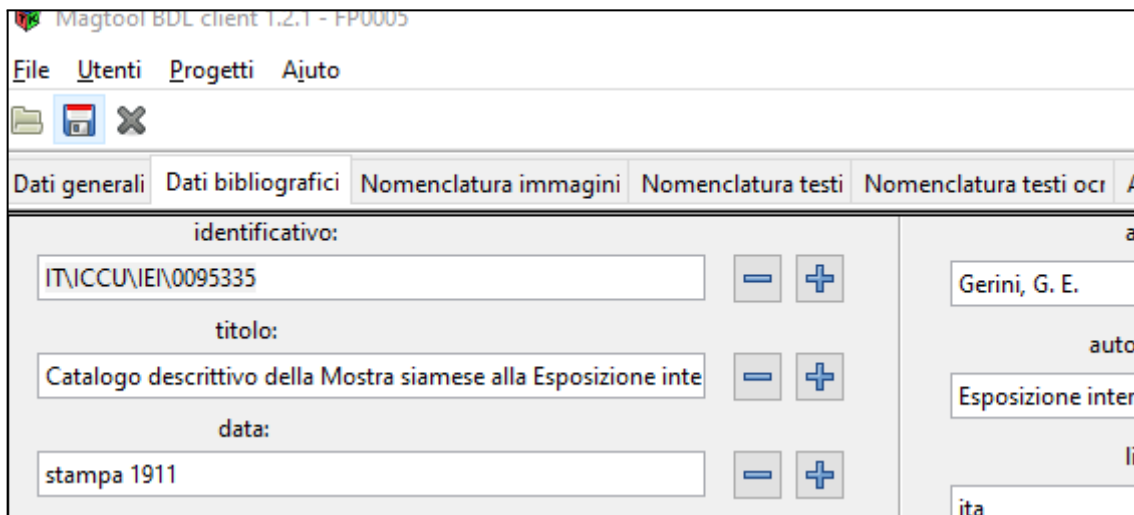
Un'altra operazione molto veloce, ma necessaria perchè la validazione avvenga con successo.

Selezionare: *File → Imposta <source> sugli ocr*
e confermare le finestre a seguire.

Salvataggio

Di tanto in tanto (e comunque alla fine della compilazione dei metadati) salvare il lavoro con il pulsante a forma di dischetto presente nella barra dei menu.

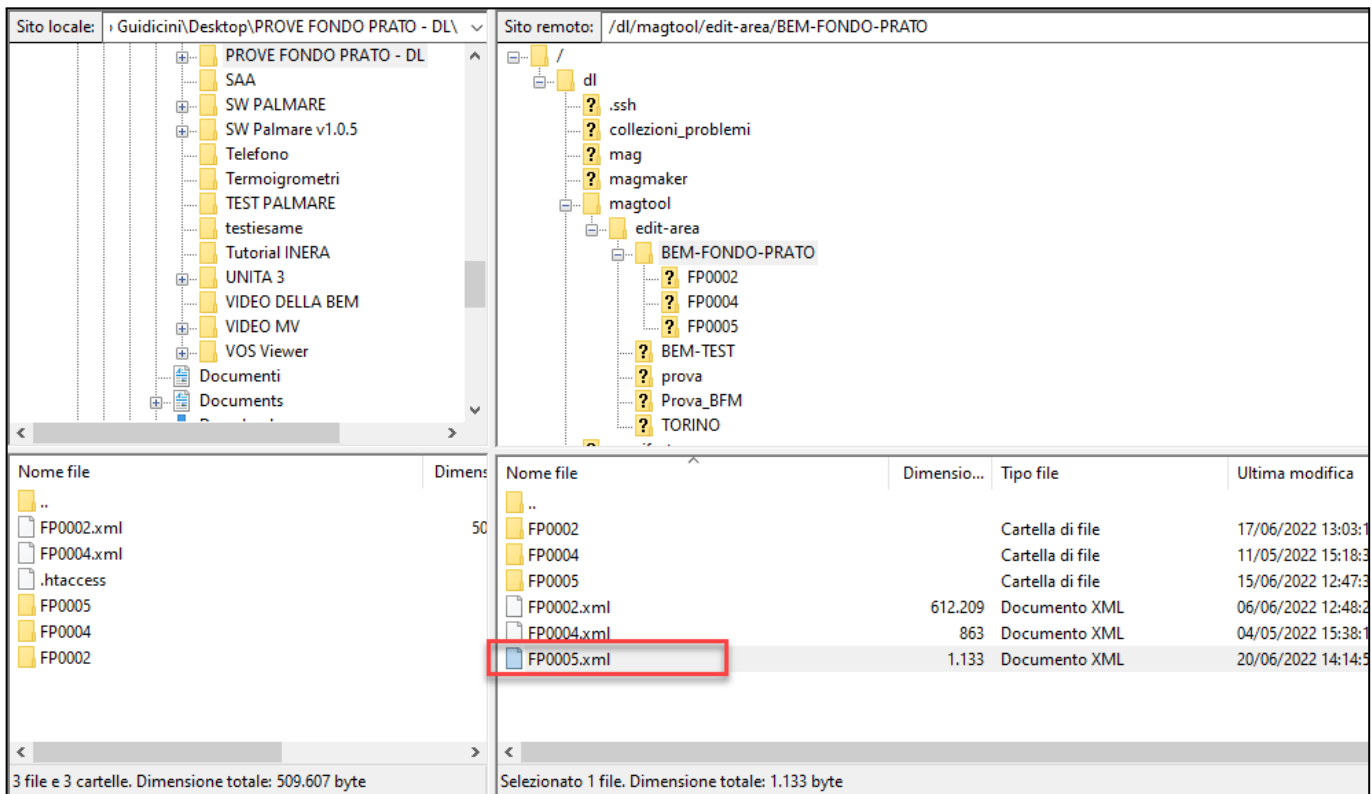
Questo consente di riprenderli in un secondo momento senza ricominciare da zero.



Questo salvataggio creerà sul server di Inera nella cartella

/dl/magtool/edit-area/BEM-FONDO-PRATO

il file FP0005.xml che conterrà i metadati salvati (vedi immagine seguente).



Nome file	Dimens	Nome file	Dimensio...	Tipo file	Ultima modifica
..		..			
FP0002.xml	50	FP0002		Cartella di file	17/06/2022 13:03:1
FP0004.xml		FP0004		Cartella di file	11/05/2022 15:18:3
.htaccess		FP0005		Cartella di file	15/06/2022 12:47:3
FP0005		FP0002.xml	612.209	Documento XML	06/06/2022 12:48:2
FP0004		FP0004.xml	863	Documento XML	04/05/2022 15:38:1
FP0002		FP0005.xml	1.133	Documento XML	20/06/2022 14:14:5

3 file e 3 cartelle. Dimensione totale: 509.607 byte

Selezionato 1 file. Dimensione totale: 1.133 byte

Validare il documento

Prima di procedere all'invio del pacchetto immagini + file dei metadati appena creato all'applicativo TECA per la pubblicazione è necessario controllare se vi siano errori formali e tecnici nel file dei metadati che è stato creato verificando che esso sia conforme allo standard MAG.

Il MAGTOOL offre una funzione che effettua automaticamente questa verifica e in caso di errori li segnala.

Gli errori vanno corretti tutti fino a quando la validazione darà esito positivo.

Solo a questo punto si potrà inviare alla TECA tutto il pacchetto.

Per procedere alla verifica cliccare su *File → Valida documento*.

Questa funzione confronta il documento con gli standard MAG e verifica che tutti i metadati necessari ci siano.

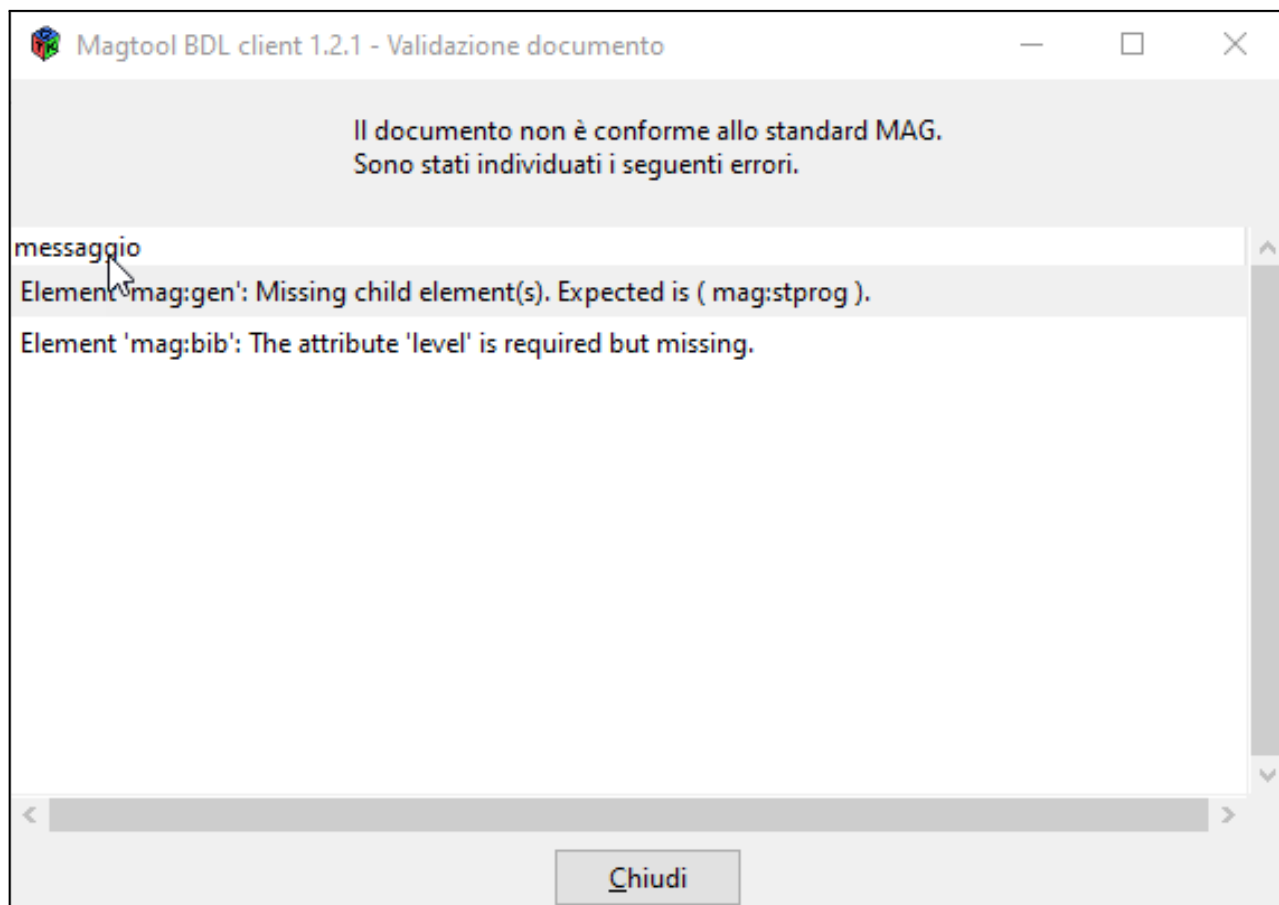
ATTENZIONE: questa funzione non controlla che i metadati siano tutti corretti. Questo controllo è lasciato all'operatore. Se alcuni metadati sono presenti (e per lo standard MAG questo è sufficiente), ma sono sbagliati, questo si rifletterà poi lato pubblico nei metadati visualizzati al pubblico.

Alcuni metadati sbagliati possono anche influire su alcune funzionalità della Digital Library.

L'esito di questa operazione di validazione potrà essere di due tipi:

1. Un messaggio di conferma che tutto è valido: si può quindi procedere alla pubblicazione.
2. Una finestra che elenca gli errori trovati che andranno risolti prima di procedere (ATTENZIONE che i messaggi di errore sono tutt'altro che chiari e vanno in qualche modo interpretati e indagati).

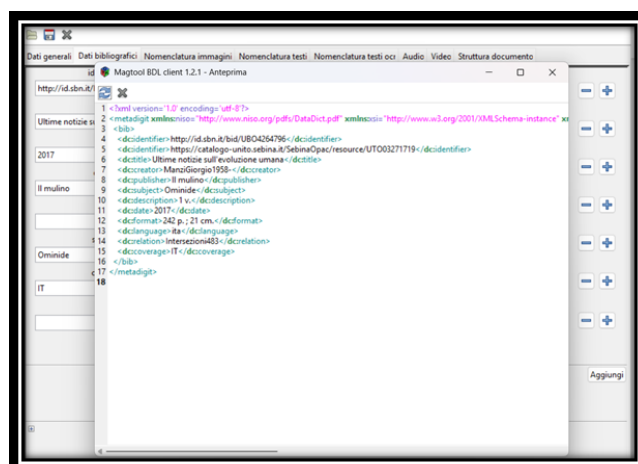
Di seguito un **esempio di messaggi di errore** (nell'esempio i messaggi si riferiscono a metadati ancora vuoti e da riempire presenti nella scheda Dati generali (mag:gen) e in quella Dati bibliografici (mag:bib):



Un altro strumento utile per indagare possibili errori nel file dei metadati è quello che consente di **visualizzare un'anteprima del file xml creato**.

Cliccare su *File* → *Preview*

Verrà aperta una finestra che mostrerà il contenuto del file xml che contiene i metadati (come nell'esempio seguente):

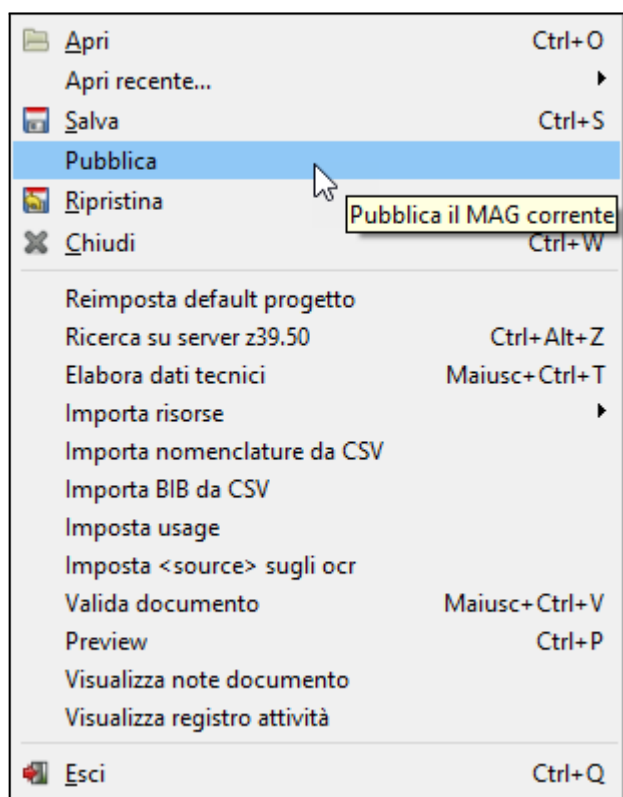


NOTA: Il file visualizzato è molto breve perchè al momento della realizzazione dello screenshot sono stati inseriti solo pochi metadati. Tuttavia dopo avere inserito anche altri metadati esso potrebbe diventare anche molto lungo.

Pre-pubblicazione dell'oggetto digitale.

Quando tutti i metadati sono stati salvati e verificati è possibile inviare tutto il pacchetto (immagini, file dei metadati, etc...) all'applicativo web TECA per poi pubblicarlo.

Questa operazione è molto semplice basta cliccare su *File* → *Pubblica*



Verrà chiesta una conferma. Se si è sicuri confermare. Il file verrà chiuso e questa operazione farà in modo che tutti i file necessari per la pubblicazione vengano spostati (sul server di INERA):

dalla cartella	alla cartella
/dl/magtool/edit-area/BEM-FONDO-PRATO	/dl/teca/import

e da qui verranno letti dalla TECA per pubblicarli nella collezione online.