

Il presente documento ha come finalità la realizzazione di una Collezione digitale secondo gli standard nazionali e la sua fruizione dal portale italiano <http://www.internetculturale.it>

Le collezioni digitali possono essere affidate in gestione alla MagTeca dell'ICCU, teca digitale di Internet Culturale. La MagTeca garantisce inoltre la diffusione delle collezioni digitali anche ai portali aggregatori Cultura Italia ed Europeaana.

I progetti destinati al sistema Internet Culturale devono essere realizzati secondo gli standard definiti dall'ICCU, che riguardano tutte le fasi del progetto, dalla scansione ottica dei documenti alla definizione e realizzazione dei metadati, obbligatori per la gestione e conservazione e necessari ai fini della ricerca sul web.

Per informazioni e documentazione rivolgersi a

[Istituto Centrale per il Catalogo Unico delle biblioteche italiane \(ICCU\)](#)

[Viale del Castro Pretorio, 105 - 00185 ROMA](#)

[Tel +39 0649210417 - Fax +39 064959302](#)

[E-mail l.ciancio@iccu.sbn.it](mailto:l.ciancio@iccu.sbn.it)

Vademecum per la digitalizzazione

Scansione ottica

Le attività di scansione devono essere eseguite in sede con attrezzatura idonea: scanner a luce fredda e in luogo appropriati, nel pieno rispetto della normativa vigente in tema di sicurezza del lavoro, per la sicurezza e conservazione dei documenti da trattare, l'ambiente sarà oscurato.

Tutta la strumentazione necessaria allo svolgimento del servizio viene installata e configurata dall'impresa, calibrata sempre al montaggio, prima di procedere alle scansioni.

Tenuto conto del pregio e delicatezza degli originali, le modalità di ripresa che riguardano l'apertura dei volumi, l'illuminazione, la manipolazione, dovranno essere tali da non arrecare loro danno, le immagini saranno acquisite appoggiando i piatti della legatura sopra il piano basculante dello scanner, le pagine dovranno essere sfogliate con cura e senza esercitare pressione sul documento, in particolare sui dorsi delle legature. Durante tutte le operazioni di digitalizzazione il personale addetto dovrà indossare appositi guanti di cotone.

La riproduzione deve avvenire con l'utilizzo di lampade a luce fredda.

Verrà prodotto un file immagine (oggetto digitale) per ciascuna scansione relativa a:

- ciascun *verso* e *recto* di ciascuna carta, comprese le carte di guardia, anche se prive di informazioni, e le carte bianche sia interpolate che consecutive;
- tutte le parti componenti la legatura: piatti, dorso, tagli qualora significativi, (nel caso di codici, in modo da mostrare capitelli, fermagli, borchie, cantonali);
- la scala cromatica sarà posizionata una sola volta su una carta originale non significativa di ciascun volume (o per quante unità componenti); la carta da scansionare sarà concordata con il responsabile del progetto;
- la scala millimetrica sarà posizionata una sola volta su una carta originale non significativa di ciascun volume (o per quante unità componenti); la carta da scansionare sarà concordata con il responsabile del progetto; posizionata esternamente alla carta e lungo il bordo inferiore con lo "zero" allineato al bordo verticale della carta, il righello sarà lungo almeno quanto la carta da riprendere;

- le carte saranno riprese in modo “naturale”, includendo i margini delle carte e delle parti del volume sottostante, per mostrare la materialità del documento nella sua interezza, senza interferire con la qualità dei colori sulla profondità di campo. Le riprese saranno effettuate su sfondo nero e le immagini saranno poi rifilate in modo da comprendere un’area minima scura, circostante la carta digitalizzata che ne consenta la visualizzazione nel contesto, comunque in modo tale che non venga mai esclusa dalla ripresa nessuna parte del documento. Le carte/pagine nella zona della legatura dovranno essere tagliate con un margine per mostrare la legatura e in piccola parte la pagina di fronte.
- in caso di presenza di lacerazioni, di fori di tarlo e ossidazione degli inchiostri, le carte devono essere mascherate con carta bianca in modo da evitare di riprendere il contenuto sottostante.
- In presenza di carte da riprendere, più piccole dei fogli sottostanti, porre al disotto del foglio in ripresa, una carta giapponese, di spessore tale che consenta la visibilità delle pagine sottostanti in modo tale che non interferiscano con la lettura del foglio scansionato, di dimensione pari alle misure del documento.

La cartella delle scansioni sarà organizzata collocando nell’ordine: piatto anteriore, dorso, piatto posteriore, e in sequenza le carte interne, in fondo all’intero pacchetto la scala cromatica e millimetrica.

Formati di acquisizione immagini

L’acquisizione dovrà essere eseguita rispettando i seguenti parametri:

- TIFF 6.0 non compresso ad alta risoluzione, 600 dpi ottici, profondità di colore di 24 bit RGB per formato inferiore o uguale ad A4; e con una risoluzione di 400 dpi ottici, profondità di colore di 24 bit RGB per formato decisamente superiore ad A4. Tale digitalizzazione è destinata alla conservazione fuori linea e come copia di sicurezza, chiamata (master).

Dal master una volta verificata la qualità e completezza verranno successivamente prodotti:

- JPEG in formato compresso a media risoluzione, 300 dpi ottici e profondità di colore di 24 bit RGB, destinata alla consultazione intranet (se ritenuto necessario alla biblioteca);
- JPEG in formato compresso a bassa risoluzione, 150 dpi ottici, profondità di colore di 24 bit RGB, (il fattore di ridimensionamento può essere definito in fase prototipale, in funzione di un’agevole consultabilità, a seconda della qualità del materiale (inchiostri, qualità della scrittura, della stampa, corpo dei caratteri, anche in relazione alle dimensioni dell’originale. Particolare attenzione deve essere posta nella scelta dei dpi per i periodici, determinata dal corpo dei caratteri quando molto piccolo e dalle dimensioni del foglio; è infatti indispensabile una buona leggibilità a video con lo zoom.
- un file in formato PDF con OCR embedded, tanti file quanti sono i file immagine, con la stessa nomenclature, sequence_number e nome del file del corrispettivo file IMG. La sez. OCR, avrà il solo usage 3. In sostituzione, è gradita la produzione di un unico file xml che elaborato in base allo schema ALTO di codifica, sia in grado di restituire le coordinate di ogni singola parola direttamente sul file di ciascuna immagine di usage 3 (sez. IMG per WEB).

E’ sempre possibile il ricorso a programmi di miglioramento e fotoritocco (riduzione del bordo nero esterno, correzione delle micro-rotazioni, rafforzamento del contrasto con filtri di smoothing e di riduzione del rumore, ecc.).

Prototipo scansioni e controlli di qualità

Il sistema di controllo della qualità delle immagini è finalizzato ad assicurare la buona leggibilità a video di tutto il contenuto informativo presente negli originali.

- creazione di un prototipo del prodotto sulla base di un campione significativo dei materiali fornito dalla ditta;
- verifiche periodiche a cura del Responsabile del progetto, che avrà facoltà di chiedere la ripetizione delle scansioni difettose.

Il prototipo che individua i parametri/formati più opportuni, sulla base di quanto indicato nel precedente paragrafo, viene approvato con verbale dal responsabile del progetto per la Biblioteca e dal responsabile per l'Impresa, e costituisce il riferimento di qualità per le attività di scansione.

Identificazione delle immagini nelle memorie di massa

Nomi immagini

Durante la realizzazione del prototipo di cui sopra, insieme all'elenco dei materiali da digitalizzare, il responsabile del progetto per la Biblioteca fornisce all'Impresa la lista dei nomi (BID) per la creazione delle cartelle immagini che porteranno il nome del BID, e la procedura più idonea per attribuire a ciascuna immagine un nome identificativo univoco di lunghezza prefissata (filename di venti caratteri al massimo). Sono ammessi caratteri non accentati A-Z, a-z, 0-9), e alcuni segni di interpunzione: il punto (.), il trattino (-), l'underscore (_).

I filename saranno completati a cura dell'Impresa con l'estensione fissa più opportuna, “.tif” o “.jpeg”, prevedendo in quest'ultimo caso un metodo per distinguere fra loro le immagini dei formati jpeg), di cui al precedente paragrafo “Formati di acquisizione immagini”.

Nelle memorie di massa, i file-immagine saranno distribuiti in più cartelle, almeno una per ciascuna unità catalografica (BID), anch'esse denominate secondo il procedimento appena esposto, in modo da preservare l'organizzazione complessiva dei materiali.

Nel caso dei Periodici, essendo complessa la casistica della periodicità che può essere usata per la preparazione delle cartelle, si invita a prendere contatto con gli uffici dell'ICCU, per entrare nello specifico dei singoli casi. In ogni modo i metadati dei periodici saranno realizzati per singolo fascicolo.

Nel caso di materiale documentario d'archivio, come corrispondenza, contratti, disegni, ecc., per la complessità si invita a prendere contatto con gli uffici dell'ICCU.

Predisposizione e Consegna degli oggetti digitali

I file-immagine saranno distribuiti in modo logico e univoco, in più cartelle, una per ogni unità catalografica, assegnando come nome alle cartelle immagini il BID del catalogo SBN, per MANUS sarà il CNMD. Quando non catalogati in un sistema elettronico, assegnare la segnatura del manoscritto, o un numero progressivo, preceduto dalla sigla di anagrafe delle Biblioteche italiane, per es. LU0022_ms.1942. Usare questi segni di interpunzione: il punto (.), il trattino (-), l'underscore (_), per separare codice da segnatura.

Supporti di consegna

Al termine dell'attività, il prodotto finale verrà consegnato su HD Desk con interfaccia USB completo delle cartelle degli oggetti digitali dei TIFF, JPG (300), JPG (150). In tre copie, insieme un breve rapporto descrittivo del lavoro svolto, di tutte le scelte tecniche operate, nonché l'elenco dei contenuti degli HD Desk. Una delle tre copie dell'intero progetto sarà consegnata su HD Desk

di marca diversa dalle altre, una delle altre due copie insieme ai metadati viene consegnata all'ICCU per l'acquisizione in MagTeca.

Produzione metadati

I metadati di Internet culturale sono scritti in formato xml e in base allo standard SCHEMA MAG, versione 2.01

<http://www.internetculturale.it/opencms/opencms/it/main/partner/servizi/standard/>

L'ICCU ha realizzato mapping da Unimarc verso MAG, per Libro antico, Libro moderno, Musica manoscritta, Cartografia, in preparazione il mapping da Manus. I mapping sono pubblicati sul portale Internet culturale nella stessa pagina dello schema MAG..

Prototipo metadati e controlli di qualità

Il sistema di controllo della qualità dei metadati è finalizzato ad assicurare l'acquisizione in MagTeca e la ricercabilità dell'oggetto digitale attraverso i motori di ricerca.

- creazione di prototipi dei metadati sulla base di un campione significativo dei tipi di materiali digitalizzati nel progetto, predisposto dalla ditta;
- verifiche in fase di acquisizione in MagTeca, che avrà facoltà di chiedere la riscrittura del mag difettoso.

Il prototipo che individua i parametri/formati più opportuni, sulla base di quanto indicato nei mapping di riferimento, viene approvato con verbale dal responsabile del progetto per la Biblioteca e dal responsabile per l'Impresa, e costituisce il riferimento di qualità per le attività di scansione.

I metadati vengono consegnati insieme alle digitalizzazioni per l'acquisizione nella MagTeca dell'ICCU.

Le indicazioni di nomenclatura della sez. IMG, indicazioni per l'OCR, le sez. GEN, e BIB vengono concordate con l'ICCU, per il rispetto degli standard e delle norme redazionali di Internet culturale.