

Allegato A – Requisiti tecnico-funzionali

SERVIZIO DI DIGITALIZZAZIONE DELLA RIVISTA “GENOVA” AL FINE DI CONSENTIRNE LA CONSULTAZIONE ONLINE E LA PUBBLICAZIONE SULLA TECA DIGITALE DELL’ISTITUTO CENTRALE PER IL CATALOGO UNICO DELLE BIBLIOTECHE ITALIANE (ICCU)

Progetto finanziato nell’ambito del Programma Operativo Nazionale per le Città Metropolitane – PON Metro 2014-2020

ART. 1 – DISPOSIZIONI GENERALI

ART. 2 – FINALITA’ DEL SERVIZIO

ART. 3 – OGGETTO DEL SERVIZIO

ART. 4 – SERVIZIO RICHIESTO

ART. 5 – ATTIVITA’ E SPECIFICHE TECNICHE

5.0_Generalità

5.A_Scansione ottica, trattamento OCR

- **5.A.1 Procedure di scansione**
- **5.A.2 File immagine**
- **5.A.3 Trattamento OCR**

5.B_Identificazione delle immagini, nomenclatura file e cartelle oggetti digitali

5.C_Metadattazione MAG

5.D_Predisposizione e consegna della versione digitale della rivista “Genova”

ART. 6 – CONTROLLI DELLA QUALITÀ

ART. 7 – TERMINI DI CONSEGNA

ART. 1 – DISPOSIZIONI GENERALI

Le condizioni di esecuzione delle prestazioni contrattuali richieste sono quelle indicate nel documento “Condizioni particolari del contratto”.

Per quanto non previsto nel citato documento si rinvia alle disposizioni di Legge e di Regolamento vigenti in materia.

Il Comune di Genova è esente da ogni responsabilità per danni a persone e/o cose provocati durante la realizzazione del servizio oggetto del presente capitolato

ART. 2 – FINALITA' DEL SERVIZIO

Al fine di assicurare la conservazione di un patrimonio bibliografico e documentario particolarmente rilevante come fonte per la storia di Genova e della Liguria nel Novecento, favorendone la valorizzazione e la facilità di fruizione, nel rispetto della vigente legislazione sul diritto d'autore, il Comune di Genova intende realizzare la versione digitale della rivista “Genova”, propria testata periodica di informazione ufficiale, pubblicata regolarmente dal 1921 fino al 1978.

La realizzazione della versione digitale della rivista “Genova” costituisce parte integrante e completamento del progetto “Servizi bibliotecari on line” (GE1.1.1.o), elaborato e finanziato nell'ambito del Programma Operativo Nazionale - Città Metropolitane 2014-2020 (PON Metro 2014-2020).

Tale collezione digitale incrementerà il catalogo delle pubblicazioni digitali delle biblioteche di Genova, accessibili dal portale www.bigmet.org, e sarà affidata alla Teca Digitale dell'Istituto centrale per il catalogo unico delle biblioteche italiane (d'ora in poi: ICCU) che la renderà consultabile tramite il portale www.internetculturale.it.

ART. 3 – OGGETTO DEL SERVIZIO

Il Comune di Genova (d'ora in poi: Ente) intende affidare la digitalizzazione della rivista “Genova”, che dovrà essere realizzata in conformità a Linee guida e standard definiti dall'ICCU, disponibili sul portale dell'Istituto, alla pagina <http://www.internetculturale.it/it/1131/linee-guida-e-standard>, e come di seguito meglio specificato.

Le annate oggetto di digitalizzazione si compongono di **fascicoli mensili**, in alcuni casi rilegati in volumi per annata. Ogni fascicolo generalmente è composto da più parti distinte, prevalentemente due, spesso stampate su carta di qualità e spessore differenti: la prima parte riporta articoli e immagini di cronaca e vita cittadina, la seconda è dedicata ai bollettini statistici; la paginazione dei fascicoli varia molto nel corso degli anni e sono presenti tavole fuori testo, semplici o ripiegate, di formato vario, non comprese nella numerazione; analogamente, sono presenti pagine di pubblicità, fuori testo e non numerate. Tutte queste parti componenti saranno oggetto di digitalizzazione, indicizzazione e metadattazione.

Complessivamente, si tratta di **60.000 pagine**, formato prevalente cm 32 x 24 (altezza x larghezza), di cui sono parte circa 70 tavole ripiegate di formato vario, da semplice multiplo della pagina, casistica prevalente, fino alla dimensione massima di cm 90 x 150.

Lo stato di conservazione di tutte le parti componenti, oggetto di digitalizzazione, non presenta alcuna criticità.

ART. 4 – SERVIZIO RICHIESTO

Il servizio in oggetto comprende le seguenti attività:

- A. scansione ottica di tutte le parti componenti della rivista “Genova”, e produzione dei file immagine nei formati e secondo le specifiche indicate al successivo art. 5; trattamento OCR - Optical Character Recognition (d’ora in poi: OCR) di tutte le parti testuali, nel formato e secondo le specifiche indicate al successivo art. 5;
- B. identificazione delle immagini, nomenclatura e indicizzazione di file e cartelle secondo gli standard richiesti dall’ICCU e nelle modalità previste al successivo art. 5;
- C. metadattazione MAG secondo gli standard richiesti dall’ICCU e nelle modalità previste al successivo art. 5;
- D. predisposizione e consegna del prodotto finale nelle modalità e nei tempi di cui ai successivi art. 5 e 7;
- E. controllo di qualità nelle modalità e nei tempi di cui al successivo art. 6.

ART. 5 – ATTIVITA’ E SPECIFICHE TECNICHE

5.0_Generalità

Sede e trasporto

L’attività di scansione ottica, così come tutte le successive attività elencate all’art. 4, dovranno essere realizzate presso i laboratori dell’Azienda aggiudicataria del servizio.

I laboratori in cui saranno realizzate le attività e tutti i locali in cui il materiale bibliografico consegnato sarà conservato, dovranno garantire condizioni ambientali e di sicurezza adeguate e conformi a quanto previsto dalla normativa per la tutela del materiale bibliografico cartaceo di pregio.

Imballaggio e trasporto dei materiali saranno a carico dell’Azienda aggiudicataria. Le modalità di imballaggio e trasporto dovranno essere adeguate alla tipologia di materiali bibliografici di pregio.

Consegna e restituzione copie originali

La consegna dei fascicoli/volumi della rivista sarà compito del personale dell’Ente e avverrà presso il magazzino civico sito in via Chiusone (Genova – Sampierdarena).

La data di consegna verrà concordata con l’Azienda aggiudicataria, compatibilmente con le attività di predisposizione del materiale a cura dell’Ente.

Ad attività terminate, i fascicoli/volumi dovranno essere tempestivamente riconsegnati, presso la stessa sede dell’avvenuta consegna, o altra del Sistema Bibliotecario Urbano di Genova che sarà cura dell’Ente comunicare.

La ricollocazione a scaffale sarà compito del personale dell’Ente.

Assicurazione

L’Azienda aggiudicataria dovrà stipulare adeguata polizza assicurativa a copertura di eventuali danni colposi o dolosi arrecati ai materiali bibliografici durante il trasporto e/o nel corso della lavorazione.

Il valore assicurativo è stimato complessivamente in € 10.000,00. La polizza assicurativa dovrà avere validità fino al termine del servizio oggetto del presente capitolato.

Personale

Il personale impegnato nel servizio oggetto del presente capitolato dovrà essere in possesso di adeguata formazione specifica e di esperienza pregressa in attività di digitalizzazione, indicizzazione e metadattazione MAG di materiali bibliografici. Prima dell'avvio delle attività, l'Azienda aggiudicataria dovrà presentare all'Ente i curricula del personale che sarà impegnato in esse.

Strumenti per la scansione

Dovranno essere utilizzati scanner planetari professionali. Il sistema di acquisizione (sorgente di luce, ottica, sensore, software di acquisizione e di calibrazione) deve essere conforme a quanto previsto dalla normativa sulla tutela del materiale bibliografico cartaceo di pregio e dovrà garantire la qualità di immagine richiesta dal servizio oggetto del presente capitolato senza danni ai materiali originali.

In particolare

- il sistema di illuminazione deve essere a luce fredda senza emissione di UV e IR;
- il campo di ripresa deve essere adeguato a tutti i diversi formati che il materiale oggetto della scansione presenta;
- il supporto per l'appoggio dei fascicoli/volumi da scansionare deve essere idoneo a evitarne il danneggiamento, rivolgendo verso l'alto la superficie da digitalizzare.

5.A_Scansione ottica, trattamento OCR

5.A.1 Procedure di scansione

Si produrrà un file immagine a colori (oggetto digitale) per ogni singola parte componente della rivista, più in dettaglio:

- ogni singola pagina anche se bianca, priva di informazioni e/o di numerazione, comprese le carte di guardia, quando presenti;
- ogni singola tavola, fronte e retro;
- ogni singola tavola ripiegata, rilegata con il fascicolo o sciolta:
 - fronte e retro, se il retro riporta una qualunque informazione,
 - solo fronte, se il retro è privo di informazioni;
- ogni singola velina, quando presente a protezione o corredo delle tavole, fronte e retro, anche se priva di informazioni;
- ogni copertina e controcopertina anteriore, dorso (quando presente) e copertina e controcopertina posteriore dei fascicoli/volumi.

Le modalità di ripresa (apertura dei fascicoli/volumi, posizionamento, manipolazione) dovranno essere adeguate al pregio e alla delicatezza dei materiali oggetto di scansione, e realizzate senza causare loro alcun danno. In particolare:

- le immagini saranno acquisite appoggiando i fascicoli/volumi su un supporto idoneo e rivolgendo verso l'alto la superficie da riprodurre;
- si avrà la massima cura nello sfogliare le pagine e non sarà esercitata alcuna pressione sui fascicoli/volumi.

Non si procederà alla separazione dell'immagine in caso di illustrazioni, carte geografiche, mappe, tabelle ecc., su doppia pagina.

Le scale cromatica e millimetrica saranno posizionate una sola volta su una pagina non significativa di ciascuna unità fisica; la scala millimetrica verrà posizionata esternamente alla pagina e lungo il bordo inferiore con lo "zero" allineato al bordo verticale, il righello sarà lungo almeno quanto la pagina. La scelta delle pagine per il posizionamento delle scale cromatica e millimetrica sarà concordata con il responsabile del progetto, in fase di realizzazione del prototipo di scansione di cui all'art. 6 "Controlli della qualità".

Scala cromatica e scala millimetrica saranno posizionate nella cartella immagini al termine della sequenza di file immagini (per la disposizione delle immagini nella cartella si veda il paragrafo 5.B "Identificazione delle immagini, nomenclatura file e cartelle oggetti digitali").

Ogni oggetto di scansione sarà ripreso nella sua interezza, in modo "naturale", includendo i margini e le parti del volume/fascicolo sottostante, senza interferire con la qualità dei colori sulla profondità di campo. Le scansioni saranno effettuate su sfondo nero e in modo da comprendere un'area circostante di qualche millimetro, che consenta la visualizzazione nel contesto, comunque in modo tale che non venga mai esclusa dall'immagine nessuna parte dell'oggetto.

Nell'eventualità che si presentino lacerazioni o lacune, le pagine dovranno essere mascherate con carta di colore neutro in modo da evitare di riprendere il contenuto sottostante.

Inoltre, dovranno essere messe in atto tutte le necessarie operazioni (utilizzo di connessioni proprietarie ad alta velocità, ecc.) per minimizzare il tempo di trasferimento dei dati dalle periferiche di acquisizione verso l'hardware di controllo e le periferiche di memorizzazione.

5.A.2 File immagine

Per ogni singola parte componente della rivista "Genova", oggetto di scansione (pagina, tavola, tavola ripiegata, ecc.), saranno realizzati file immagine nei seguenti formati e rispettivi parametri:

- TIFF 6.0 non compresso, profondità di colore di 24 bit RGB (file master),
 - 400 dpi ottici, per formato inferiore o uguale ad A4;
 - 300 dpi ottici, per formato superiore ad A4.

Questi oggetti digitali saranno destinati all'archiviazione fuori linea e alla conservazione a lungo termine.

Dal file master si produrranno i seguenti file derivati:

- JPEG compresso a media qualità, con fattore di qualità 100%, 300 dpi ottici e profondità di colore di 24 bit RGB.

Questi oggetti digitali saranno destinati alla consultazione intranet.

- JPEG compresso a bassa qualità, con fattore di qualità 85%, 150/200 dpi ottici e profondità di colore di 24 bit RGB, con fattore di ridimensionamento da 2000 pixel circa lato lungo. Questi ultimi parametri saranno meglio definiti durante la realizzazione del prototipo (v. art.

6 “Controlli della qualità”), rispettivamente, in base ai formati, se superiori o inferiori ad A4, e in funzione di una agevole consultabilità internet.

Questi oggetti digitali saranno destinati alla consultazione internet.

Sui file derivati sarà 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., da non eseguirsi comunque sui file master).

5.A.3 Trattamento OCR

I file TIFF (file master) saranno sottoposti a OCR e si realizzeranno tanti file PDF con OCR embedded, quanti sono i file immagine, a cui saranno assegnati la stessa nomenclatura, sequence_number e nome file del corrispondente file IMG. La sez. OCR avrà il solo usage 3.

In sostituzione, sarà possibile 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).

I file così realizzati saranno utilizzati anche per generare file PDF multipagina ricercabili, con OCR embedded.

5.B Identificazione delle immagini, nomenclatura file e cartelle oggetti digitali

Nelle memorie di massa le diverse sez. immagini di ciascun fascicolo/volume saranno contenute in una cartella così nominata: BID SBN della rivista “Genova”, da indicarsi in base all’anno di pubblicazione del fascicolo/volume, seguito da underscore e anno.

Esempio:

- Directory superiore BID_anno
 - Directory Master: tiff
 - file immagine: bid_annomese_fasc_(eventuali indicazioni supplementari)_n° progressivo
 - ...

La cartella delle scansioni di ciascun fascicolo/volume sarà organizzata collocando nell’ordine: copertina e contro copertina anteriore, dorso (quando presente), in sequenza le pagine interne, copertina e contro copertina posteriore, infine la scala cromatica e millimetrica.

In sede di realizzazione del prototipo di scansione, l’Ente fornirà all’Azienda aggiudicataria l’elenco dei fascicoli/volumi da digitalizzare con il nome da attribuire ai file immagine di ogni fascicolo (“File system” della rivista “Genova”), compilati secondo le indicazioni dell’ICCU, con l’indicazione del BID SBN corrispondente alle diverse annate della rivista.

I filename saranno completati a cura dell’Azienda aggiudicataria con la nomenclatura delle singole pagine e l’opportuna estensione fissa: “tiff”, “jpg”, “pdf”.

5.C Metadattazione MAG

Contestualmente all’acquisizione, identificazione e nomina delle immagini di cui alle attività precedenti (5.A e 5.B), l’Azienda aggiudicataria curerà la compilazione di un file XML per ciascun fascicolo.

Le informazioni bibliografiche saranno fornite dall'Ente in formato Unimarc, il mapping tra Unimarc e metadati MAG è quello previsto dall'ICCU.

I file XML dovranno avere una struttura strettamente corrispondente al "MAG Schema" versione 2.0.1, con le restrizioni per la Teca Digitale dell'ICCU, il cui schema è disponibile alla pagina

<http://www.internetculturale.it/it/1131/linee-guida-e-standard>

In sede di realizzazione del prototipo di scansione, unitamente all'elenco dei fascicoli/volumi da digitalizzare e dei nomi da attribuire ai file immagine, di cui al precedente punto 5.B, l'Ente fornirà all'Azienda aggiudicataria le relative informazioni utili per la creazione della sez. pezzi del metadato per la pubblicazione, da ottimizzare secondo le norme dello schema MAG. A tal fine, si veda il documento "Specifiche tag pezzi periodici e unità componenti 1.6" alla pagina <http://www.internetculturale.it/getFile.php?id=44403>

La nomenclatura delle parti componenti delle sez. digitali saranno a cura della Azienda aggiudicataria, con la supervisione e il controllo del responsabile del progetto.

5.D_Predisposizione e consegna della versione digitale della rivista "Genova"

Al termine dell'attività, nei tempi indicati al successivo art. 7 "Termini di consegna", tutte le serie di cartelle immagini (tiff, jpg, pdf...) della versione digitale della rivista "Genova", insieme ai rispettivi file XML-MAG, dovranno essere consegnate all'Ente nelle seguenti modalità:

- 3 copie complete, comprensive dei file master;
- 1 copia senza i file master, contenente i soli file per la consultazione internet (JPEG compresso a bassa qualità e PDF con OCR embedded). Tale copia sarà trasmessa all'ICCU per il caricamento, la validazione finale e la pubblicazione nella Teca Digitale dell'Istituto.

Le copie saranno predisposte su hard disk esterni con tecnologia di connessione USB (almeno versione 3.0); uno dei tre hard disk esterni contenenti le copie complete, dovrà essere di marca o serie diversa dagli altri due; in ciascuna copia dovrà essere presente:

- un breve rapporto descrittivo dell'attività svolta e delle scelte tecniche operate,
- un documento (Excel) con l'elenco del materiale consegnato, indicando per ogni anno della rivista il numero di file per ciascun formato (tiff, jpg, ecc..) risultanti dalle attività di digitalizzazione.

ART. 6 – CONTROLLI DELLA QUALITÀ

Tutte le attività A, B, C e D, di cui al precedente art. 5, saranno sottoposte a controlli della qualità, in accordo e coerenza con le indicazioni fornite dall'ICCU.

In particolare, il sistema di controllo qualità delle immagini sarà finalizzato ad assicurare la completezza e la buona leggibilità a video di tutto il contenuto informativo presente negli originali cartacei della rivista "Genova", oltreché la qualità della ripresa nel rispetto degli stessi.

Il sistema di controllo della qualità sarà così articolato:

- 1) realizzazione, verifica e approvazione di un prototipo di oggetti digitali esito della scansione ottica. Tale prototipo sarà prodotto su un campione significativo delle diverse parti componenti della rivista "Genova", selezionato dal responsabile del progetto in accordo con il referente dell'ICCU; l'approvazione formale del prototipo costituirà riferimento di qualità per il processo di digitalizzazione;

- 2) realizzazione, verifica e approvazione di un prototipo di oggetti digitali esito del trattamento OCR. Tale prototipo sarà prodotto su un campione significativo delle diverse parti componenti della rivista “Genova” contenenti testo, selezionato dal responsabile del progetto in accordo con il referente dell’ICCU; l’approvazione formale del prototipo costituirà riferimento di qualità per il trattamento OCR;
- 3) realizzazione, verifica e approvazione di un prototipo di metadattazione MAG degli oggetti digitali. Tale prototipo sarà prodotto su un campione significativo delle diverse parti componenti della rivista “Genova”, selezionato dal responsabile del progetto in accordo con il referente dell’ICCU; l’approvazione formale del prototipo costituirà riferimento di qualità per il processo di metadattazione.

I prototipi previsti ai precedenti punti 1), 2) e 3) saranno approvati con verbale a cura del responsabile del progetto per l’Ente, del referente per l’ICCU e del capo progetto per l’Azienda affidataria del servizio.

Sia il prototipo di cui al punto 1) per le attività di scansione, sia il prototipo di cui al punto 3) per le attività di metadattazione, dovranno essere utili anche a individuare i formati e i parametri più opportuni per la realizzazione della versione digitale della rivista “Genova”, sulla base di quanto indicato nell’art. 5.

La consegna dei prototipi (ed eventuali successive correzioni) avverrà mediante la piattaforma dell’Ente per la ricezione di grandi file <https://filesender.comune.genova.it/filesender/> secondo le istruzioni e nelle modalità che saranno successivamente comunicate all’Azienda aggiudicataria.

- 4) Verifiche periodiche saranno eseguite a cura del responsabile del progetto, che avrà facoltà di chiedere la ripetizione delle scansioni difettose;
- 5) è prevista infine, una verifica finale degli oggetti digitali per la validazione dei dati e dei supporti consegnati.

ART. 7 – TERMINI DI CONSEGNA

Il servizio richiesto, articolato come previsto all’art. 4 e realizzato secondo modalità e indicazioni di cui all’art. 5, sarà completato entro e non oltre 6 mesi a far data dall’avvio delle attività, da intendersi ad avvenuta approvazione del prototipo di cui all’art. 6, punto 1).

Tale prototipo andrà sottoposto all’approvazione del responsabile del progetto per l’Ente e al referente del progetto per l’ICCU non oltre i 30 giorni successivi all’avvenuta consegna dei fascicoli/volumi della rivista “Genova”.

La conclusione del servizio sarà da intendersi ad avvenuta verifica e approvazione della qualità di immagini e metadati da parte dell’Ente e dell’ICCU.