



Conversione dei dati archivistici in pubblicazioni digitali Open Access. Un caso di studio: *Carte d'autore online*

Manuela Ferraro

Il Centro di Studi «Aldo Palazzeschi» ha promosso un progetto finalizzato alla pubblicazione in rete di inventari del materiale palazzeschiiano tramite un processo innovativo che ha origine dai dati archivistici della piattaforma di schedatura. La pubblicazione è il frutto di un flusso di lavoro che ha previsto l'esportazione dei dati del programma d'archivio, basati sullo standard ISAD (International Standard Archival Description), nel nuovo modello dati EAD (Encoded Archival Description) che definisce la codifica elettronica in formato XML dei record di descrizione archivistica. Partendo da un'analisi approfondita della struttura archivistica del fondo Palazzeschi, sono stati progettati dei tracciati XML con l'obiettivo di estrapolare, in maniera organizzata, dati specifici contenuti nel Fondo. I tracciati sono stati la base di partenza di un lavoro editoriale multifase che mira alla pubblicazione in Open Access degli inventari dei manoscritti di Palazzeschi e della biblioteca palazzeschiiana.

Study Center «Aldo Palazzeschi» has promoted a project for the purpose of the online publication of inventories of Palazzeschi's material through an innovative process that originates from the archival data of the cataloguing platform. The publication is the result of a workflow which involved exporting the archive program data, based on the ISAD (International Standard Archival Description) standard, into the new EAD (Encoded Archival Description) data model which defines the electronic coding in XML format of archival description records. Starting from an in-depth analysis of the archival structure of the Palazzeschi Fund, XML paths were designed with the aim of extrapolating, in an organized way, specific data contained in the Fund. XML paths were the starting point of a multi-phase editorial work in order to the Open Access publication of the inventories of Palazzeschi's manuscripts and the Palazzeschi's library.

Parole chiave: Archivi, Open Access, DOI, Editoria digitale

Keywords: Archives, Open Access, DOI, Digital publishing

Sommario: Caso di studio - Premessa - Descrizione del progetto - Da ISAD a EAD - Predisposizione tracciato XML - Fasi di composizione dell'inventario - Modalità di distribuzione - Conclusioni

Peer review

Submitted 20/07/2022

Accepted 21/09/2022

Published 25/10/2022

Open access

© 2023 | Attribution - Non commercial - Non derivatives (IT)

Cita come Manuela Ferraro, *Conversione dei dati archivistici in pubblicazioni digitali Open Access. Un caso di studio: Carte d'autore online* in Rivista DILEF - II, 2022/2 (gennaio-

DOI

dicembre), pp. 296-307. 10.35948/DILEF/2023.4309
10.35948/DILEF/2023.4309

Caso di studio

L'articolo si concentra sull'applicazione di uno specifico caso di studio, partendo da una banca dati e offrendo indicazioni metodologiche e operative sul processo di formazione di inventari archivistici.

Il caso di studio si riferisce al progetto di ricerca "Studio e progettazione dei tracciati XML-TEI e revisione delle carte dell'Archivio Palazzeschi per la conversione dei dati e dei testi e la loro esportazione in un flusso editoriale per la pubblicazione in Open Access", promosso dal Centro di Studi "Aldo Palazzeschi" del Dipartimento di Lettere e Filosofia dell'Università di Firenze.

Premessa

Le carte dell'archivio Palazzeschi sono disponibili in rete sul portale "Carte d'autore online. Archivi e biblioteche digitali della modernità letteraria italiana", che consente di accedere, oltre alle schede d'archivio, a riproduzioni di manoscritti, lettere, foto e trascrizioni. Tale ambiente di consultazione dialoga con il sistema di back office "WCM modulo Theke", la piattaforma web per la gestione degli archivi sviluppata dal partner tecnologico Progettinrete¹.

L'obiettivo del progetto è riutilizzare i dati archivistici inseriti nella piattaforma di schedatura e riorganizzarli in vista di un inventario. La redazione dell'inventario in questo caso non seguirà l'impianto "tradizionale" ma si servirà dei dati già presenti sull'applicativo gestionale.

Descrizione del progetto

L'inventario riveste un ruolo fondamentale nel contesto archivistico, in quanto costituisce la chiave di accesso all'archivio, ha la funzione di descriverlo e di rappresentare la sua documentazione in forma organica, in vista di una sua fruizione; costituisce «lo strumento di ricerca concettualmente più elaborato e più rigoroso sotto l'aspetto formale»².

È dunque il frutto di un attento esame della realtà documentaria e come tale deve rispettare la struttura gerarchica del fondo.

Nello specifico, l'inventario sul quale ci focalizzeremo riguarderà i manoscritti dello scrittore Aldo Palazzeschi che, all'interno del fondo Palazzeschi, sono riuniti nella serie "Carte d'Autore".

Il fondo Palazzeschi è ripartito in serie, sottoserie, fascicolo, unità documentaria (Fig. 1).

La struttura gerarchica del fondo è conforme alla cosiddetta descrizione multilivello, rappresentata graficamente attraverso l'albero rovesciato che procede dal generale al particolare, cioè dal fondo all'unità archivistica, fino ad arrivare al documento (Fig. 2).



Fig. 1 - Struttura gerarchica del fondo Palazzeschi sulla piattaforma "WCM modulo Theke".

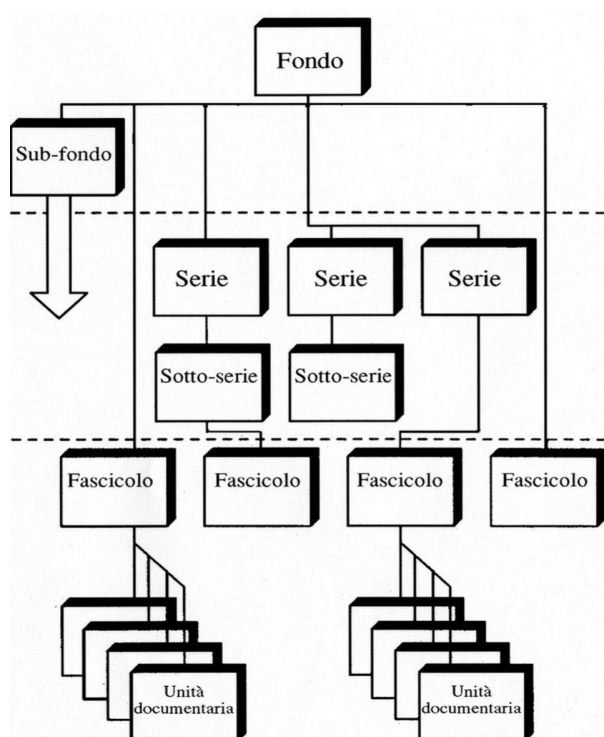



Fig. 2 - Schema dei livelli di ordinamento di un fondo secondo ISAD(G).

La piattaforma di gestione del fondo Palazzeschi, "WCM modulo Theke", aderisce agli standard internazionali di descrizione delle unità archivistiche ISAD (G)³ e di descrizione storico istituzionale del soggetto produttore ISAAR (CPF)⁴; si articola nelle sette aree informative previste dalle norme ISAD(G) (Fig. 3):

1. area dell'identificazione;
2. area delle informazioni sul contesto;
3. area delle informazioni relative al contenuto e alla struttura;
4. area delle informazioni relative alle condizioni di accesso ed utilizzazione;
5. area delle informazioni relative a documentazione collegata;

6. area delle note;
7. area di controllo della descrizione.

Archivio digitale della modernità letteraria italiana > Archivio Aldo Palazzeschi > Fondo Aldo Palazzeschi > Serie Carte d'autore > Fascicolo Sorelle Materassi > U.D. carte d'autore Sorelle Materassi

 **modifica stato**

1 - Identificazione	2 - Informazioni	3 - Contenuto	4 - Accesso	5 - Documentazione	6 - Note	7 - Controllo
<p>Soggetto Produttore Palazzeschi, Aldo 102723 Modifica</p> <p>Codice Provvisorio m130a</p> <p>Codice Definitivo AP 1.4 FP90, 90.a</p> <p>IDLivello U.D. carte d'autore ▼</p> <p>Autore Palazzeschi, Aldo</p> <p>Autore non identificato</p> <p>Destinatario</p> <p>Destinatario non identificato</p> <p>Titolo (incipit) Sorelle Materassi</p> <p>Titolo Critico No ▼</p> <p>Collocazione A1, III, S4, FP90, 90.a</p> <p>Tipologia documentaria Manoscritto ▼</p>						

Fig. 3 - Focus sulla sezione identificativa della piattaforma "WCM modulo Theke".

Da ISAD a EAD

In accordo col partner informatico, si è deciso di adottare schemi XML per la descrizione degli oggetti archivistici: EAD (*Encoded Archival Description*) ed EAC (*Encoded Archival Context – Corporate bodies, Persons and Families*).

L'utilizzo di tecnologie come XML (*eXtensible Markup Language*), che permettono la conservazione e la comunicazione dei dati, mira a garantire la persistenza della struttura e del contenuto delle descrizioni, nonché la loro accessibilità nel tempo⁵.

EAD è il modello dati (DTD)⁶, composto da una struttura e da elementi descrittivi, all'interno della quale è possibile codificare la descrizione delle fonti archivistiche. Esso costituisce l'esplicitazione in formato XML della struttura descrittiva delle ISAD(G).

EAC-CPF è il modello dati (DTD) che consente la codifica delle informazioni sui soggetti produttori e/o conservatori e formalizza in XML la struttura descrittiva delle ISAAR(CPF) (Fig. 4).

La DTD detta le “regole grammaticali” di un documento XML: definisce l’insieme degli elementi e dei relativi attributi obbligatori e opzionali, l’ordine di apparizione degli elementi nel documento XML e le relazioni gerarchiche tra gli elementi.

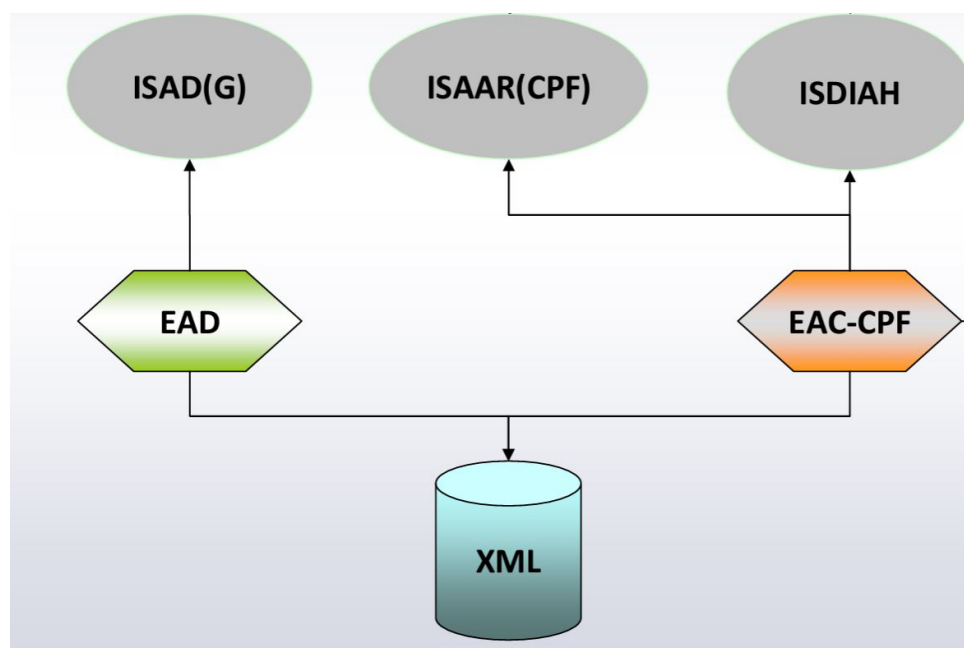


Fig. 4 - Schema esemplificativo sulle relazioni tra gli standard archivistici.

Sono state realizzate la mappatura ISAD (G) - EAD prendendo come riferimento le linee guida pubblicate dalla *Library of Congress*⁷ e la mappatura ISAAR - EAC-CPF, la quale non sarà oggetto di approfondimento in questa sede.

Ogni elemento della descrizione archivistica definito da ISAD (G) è espresso in EAD da un elemento specifico o da una combinazione di elementi.

Diversi sono i vantaggi che si possono trarre dall'utilizzo di EAD e EAC-CPF: si può elaborare qualsiasi fonte di informazione archivistica per la sua diffusione in formato elettronico, si possono conservare, gestire e aggiornare i dati intervenendo direttamente sul documento XML che è un semplice file di testo, viene facilitato lo scambio di informazioni, l'interoperabilità tra dati e formati; si garantisce una struttura di tipo gerarchico.

Predisposizione tracciato XML

Facciamo un accenno alla struttura di EAD partendo dai seguenti macro-elementi (Fig. 5):

- <eadheader> (*EAD Header*, intestazione del documento - obbligatorio) per la codifica di informazioni identificative generali sullo strumento di ricerca (titolo, compilatore, data di compilazione, edizione);
- <frontmatter> (*Front Matter*, frontespizio - opzionale) per la formattazione del documento;
- <archdesc> (*Archival Description*, descrizione archivistica - obbligatorio) contiene gli elementi per la codifica dei dati archivistici, per le informazioni e le descrizioni relative al *corpus* (fondo, serie, unità archivistica).

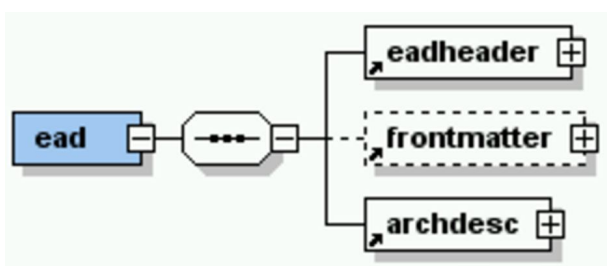


Fig. 5 - Primo livello di struttura gerarchica.

Anche lo schema EAD presenta una visualizzazione ad albero in cui gli elementi vengono annidati.

Una volta terminata la descrizione generale del complesso documentario, si giunge alla codifica dei raggruppamenti gerarchici e delle relative unità di descrizione del materiale attraverso l'elemento <dsc> (*Description of Subordinate Components*, descrizione delle componenti subordinate).

All'interno di <dsc> (*Description of Subordinate Components*) è possibile trovare diversi <c> (*Components*, Componenti) ossia unità di descrizione, ciascuna delle quali contiene sottoelementi per la codifica dei dati identificativi (Fig. 6).

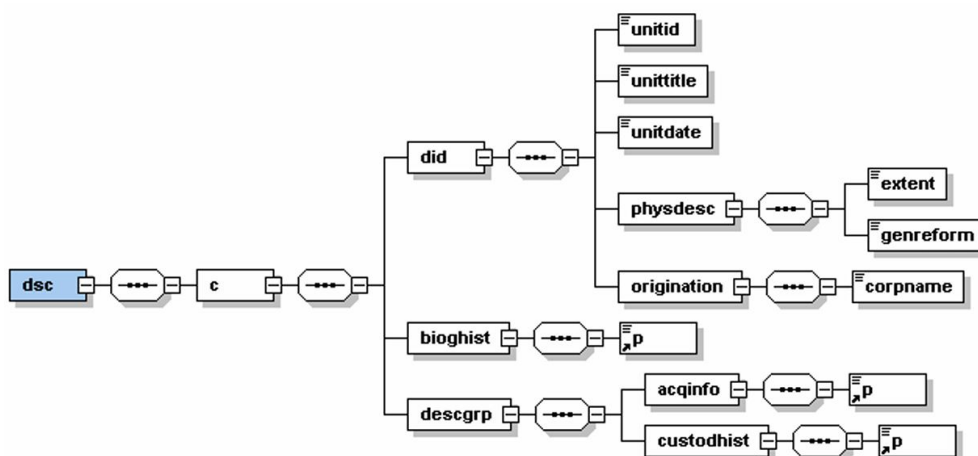


Fig. 6 - Articolazione di <c>.

Nella predisposizione del tracciato di *export* a XML-EAD sono stati adattati gli schemi di equivalenza ISAD-EAD ai campi archivistici di ciascun livello del sistema gestionale dell'archivio (Fig. 7), indicando nel tag l'etichetta del campo del programma d'archivio e il riferimento alle norme ISAD: per esempio il campo "codice definitivo" è stato riportato nel tag <unitid>, il campo "titolo" nel tag <unittitle> o ancora la data nel tag <unitdate>; in verde, invece, si possono individuare le note fornite per un corretto uso del tracciato ai fini della conversione degli standard.

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE ead PUBLIC "-//ISBN 1-931666-00-8//DTD ead.dtd (Encoded Archival Description
(EAD) Version 2002)//EN" "ftp://ftp.loc.gov/pub/ead/ead.dtd">
<ead>
  <!-- UNITÀ DOCUMENTARIA CARTE D'AUTORE-->
  <!-- in <archdesc> inserire il livello del padre-->
  <archdesc level="fascicolo" id="">
    <did>
      <unittitle>titolo scheda padre</unittitle>
    </did>
    <dsc>
      <c level="otherlevel" otherlevel="U.D. carte d'autore" audience="">
        <did>
          <unitid encodinganalog="ISAD 3.1.1 Reference code(s)" label="codice provvisorio"></unitid>
          <unitid encodinganalog="ISAD 3.1.1 Reference code(s)" label="codice definitivo"></unitid>
          <unitid encodinganalog="ISAD 3.1.1 Reference code(s)" label="segnatura originaria"></unitid>
        </did>
        <!--se il valore del campo "Titolo Critico" è sì inserire l'attributo type="TitoloCritico",
        se il valore è no non inserire alcun attributo all'interno di title-->
        <unittitle encodinganalog="ISAD 3.1.2 Title" type="titolo critico">Titolo
          <persname role="author">Autore</persname>
        </unittitle>
        <!--nel tag unitdate inserire i valori del campo Data Iniziale e del campo Data Finale separati
        da spazio e trattino (-) -->
        <!--se data iniziale e data finale coincidono inserirne solo una-->
        <unitdate encodinganalog="ISAD 3.1.3 Date(s)" normal="19520101-19641231">1885 - 1974
          <emph altrender="notes date">Note Datazione</emph>
        </unitdate>
      </c>
    </dsc>
  </archdesc>
</ead>
```

Fig. 7 - Tracciato XML di export da WCM modulo Theke a EAD.

Fasi di composizione dell'inventario

Nella fase di redazione di un inventario è sempre importante definire preliminarmente il grado di analiticità della descrizione relativa alla documentazione⁸. Pertanto, anche nel nostro caso, è stata effettuata la scelta riguardante il set di informazioni per la composizione dell'inventario.

Si è scelto dunque di produrre un inventario analitico che descrivesse le singole unità documentarie. In seguito, è stato impostato il nostro strumento di ricerca esaminando la mole di materiale e l'organizzazione della documentazione e prendendo come riferimento gli elementi essenziali per la redazione di un inventario⁹:

- strutturali (articolazioni del fondo);
- identificativi (signature originali, date estreme);
- descrittivi (note introduttive al fondo e alle singole serie, titolo ed eventuale integrazione del contenuto);
- complementari (bibliografia, indici, tavole di raffronto, appendici).

La documentazione è suddivisa in fascicoli, ciascuno contenente una o più unità documentarie. La descrizione delle carte è stata effettuata a livello del singolo

documento, tenendo conto delle varie redazioni.

In relazione alla tipologia di materiale (manoscritti), si è deciso di estrarre campi specifici del livello analizzato, estrapolando le informazioni che potessero risultare utili all'utente durante la consultazione dell'inventario.

Dopo aver stabilito la struttura dell'inventario è stato curato l'aspetto formale, ossia la disposizione grafica e la formattazione del testo: l'impaginazione, la tipologia di carattere, lo stile, la dimensione e l'interlinea, come riportato nello schema delle indicazioni tipografiche sottostante.

Livello + codice¹
Titolo (serie)
data²

Contenuto

Ordinamento

Consistenza

Livello + codice

Titolo:

Consistenza:

Data:

Livello + codice

Titolo:

Tipologia documentaria: (+ attributo type³, se presente)

Consistenza:

Autore:

Data:

Allegati:

L'ultima fase, prima di procedere all'estrazione dei dati per l'impaginazione, è stata la revisione delle carte d'archivio: per avere una visione d'insieme si è scelto di adoperare il programma excel che consente di ordinare le righe e verificare facilmente il formato e il valore dei campi numerici. Il partner tecnologico ha estratto tutto il materiale della serie interessata su un documento excel in cui ogni riga corrispondeva a una scheda archivistica (record), permettendoci così di effettuare il controllo sistematico di tutti i fascicoli e le unità documentarie, verificando tutti i dati, anche quelli non funzionali alla pubblicazione dell'inventario.

Laddove sono emerse delle criticità, è stata fatta una verifica sui pezzi archivistici coinvolti per una valutazione ponderata su ogni singolo caso.

Una volta puliti i dati, il documento excel corretto è stato restituito al partner tecnologico il quale ha aggiornato i dati nell'ambiente di schedatura e esportato i dati

revisionati -tramite un documento XML- al curatore dell'impaginazione. I dati interessati alla pubblicazione sono stati processati da un template realizzato *ad hoc* su Adobe Indesign, che ha permesso la visualizzazione del contenuto in forma sequenziale, seguendo precise disposizioni¹⁰.

L'impaginatore ha dato una nuova veste al documento XML e, avvalendosi delle istruzioni fornite ha esportato il contenuto in un file PDF, inviato, successivamente, al curatore (archivista) per la correzione delle bozze.

L'*iter* ha previsto un controllo massiccio del file, individuando eventuali errori sfuggiti nei passaggi precedenti e verificando che ogni dato del PDF prodotto automaticamente fosse stato esportato correttamente.

Nel caso specifico dell'inventario sui manoscritti, sono stati riscontrati sia errori sostanziali (estratte informazioni non previste dal tracciato), sia errori formali (righe vuote tra un campo e l'altro).

Dopo aver concluso il giro di bozze, corretto manualmente i dati nell'ambiente di schedatura e appurato che il file fosse pronto per la pubblicazione, abbiamo curato una nota al testo per fornire le indicazioni di consultazione dello strumento di ricerca, il resoconto degli interventi apportati, i criteri di schedatura, la tavola delle sigle e delle abbreviazioni.

Modalità di distribuzione

Vogliamo evidenziare la scelta consapevole e strategica di adottare una politica di distribuzione *Open Access* del volume, attraverso la licenza *Creative Commons*¹¹, CC BY-NC-ND 4.0, che consente di scaricare e condividere i lavori originali a condizione che non vengano modificati né utilizzati a scopi commerciali.

È possibile, tuttavia, anche acquistare la versione cartacea tramite il print on demand, dando al fruitore la possibilità di scegliere la modalità a lui più congeniale.

Al volume è stato assegnato un e-ISBN¹² e un DOI¹³ per identificare in modo univoco e permanente l'oggetto digitale.

Conclusioni

Il lavoro qui presentato è il risultato di un approccio metodologico che si è rivelato proficuo grazie alla cooperazione tra ricercatore/archivista, partner informatico e responsabile editoriale.

Tramite la messa a punto di questo flusso innovativo sono stati pubblicati al momento due strumenti di ricerca: *l'Inventario dei manoscritti dell'Archivio Aldo Palazzeschi* e *l'Inventario della biblioteca di Aldo Palazzeschi*.

Il sinergico lavoro di squadra ha già posto le basi per un'ulteriore pubblicazione riguardante l'inventario della corrispondenza, dilatando in una prospettiva futura quei criteri e quelle modalità di lavoro risultate fruttuose, per una fruizione sempre più agevole dei dati del fondo Palazzeschi.

Note

1. Per lo sviluppo informatico cfr. il sito <https://www.progettinrete.it/>.
2. Carucci - Guercio 2008, p. 104.
3. *General International Standard Archival Description*: elaborato tra il 1988 e il 1993 dalla Commissione *ad hoc* per gli standard di descrizione del Consiglio internazionale degli archivi (ICA). La prima edizione è stata pubblicata nel 1994. La seconda versione, riveduta, è stata emessa nel 1999.
4. *International Standard Archival Authority Records for Corporate Bodies, Persons and Families*.
5. Per approfondimenti si veda il sito dell'Istituto Centrale per gli Archivi - ICAR: <https://icar.cultura.gov.it/standard/standard-internazionali/ead>.
6. La *DTD - Document Type Definition* - descrive la struttura di un documento XML attraverso un insieme di regole che consentono di convalidare il file XML.
7. Cfr. il sito della *Library of Congress*: https://www.loc.gov/ead/tglib/appendix_a.html.
8. Vedi Carucci - Guercio 2008, p. 104.
9. Per maggiori dettagli sul passaggio in questione si veda Marini - Salucci 2022, pp. 147-168.
10. Ivi, p. 105.
11. Per maggiori dettagli sul passaggio in questione si veda Marini-Salucci 2022.
12. Le licenze *Creative Commons* sono contratti attraverso i quali il titolare dei diritti d'autore concede l'autorizzazione all'uso dell'opera ad una generalità di soggetti indefiniti. Per un'analisi approfondita si rimanda al sito <https://creativecommons.it/chapterIT/index.php/license-your-work/>.
13. L'ISBN - *International Standard Book Number* - è un numero che identifica a livello internazionale in modo univoco e duraturo un titolo o una edizione di un titolo di un determinato editore.
14. Il DOI - *Digital Object Identifier* - è uno standard che consente di identificare persistentemente, all'interno di una rete digitale, qualsiasi oggetto di proprietà intellettuale e di associarvi i relativi dati di riferimento, i metadati, mediante l'uso di schemi strutturati ed estensibili.

Bibliografia

- Aliprandi 2013 = Simone Aliprandi, *Creative Commons: manuale operativo. Guida all'uso delle licenze e degli altri strumenti CC*, Ledizioni, Milano.
- Carucci - Guercio 2008 = Paola Carucci - Maria Guercio, *Manuale di archivistica*, Carocci, Roma.
- Magherini 2009 = *Tradizione e modernità: archivi digitali e strumenti di ricerca: convegno di studi, Firenze, 27-28 ottobre 2006*, a cura di Simone Magherini, Società Editrice Fiorentina, Firenze.
- Marini - Salucci 2022 = Arianna Marini - Giovanni Salucci, *Flusso XML e InDesign per la realizzazione di edizioni digitali nelle Humanities in DILEF. Rivista digitale del Dipartimento di Lettere e Filosofia*, I, 1, pp. 147-168.
- Salucci - Cinotti 2021 = Giovanni Salucci - Roberto Cinotti, *Open Access e studi umanistici* in «Studi italiani», 65, pp. 159-184.
- Tommasi 2018 = Brizio Leonardo Tommasi, *Project management and digital transformation: performance measuring model of digital projects and archives* in *JLIS*, 9, 3, pp. 92-108.

Sitografia

- DOI Foundation: <https://www.doi.org/>.
- Encoded Archival Description Tag Library, Version 2002: https://www.loc.gov/ead/tglib/appendix_a.html.
- ISAD - Traduzione italiana a cura di Stefano Vitali: http://media.regesta.com/dm_0/ANAI/anaiCms/ANAI/000/0111/ANAI.000.0111.0002.pdf.
- Istituto Centrale per gli Archivi - ICAR: <https://icar.cultura.gov.it/standard/standard-internazionali>.
- Le licenze Creative Commons: <https://creativecommons.it/chapterIT/index.php/license-your-work/>.
- SAN - Strumenti di ricerca online: <https://inventari.san.beniculturali.it/inventari>.