

Development of a semi-automated framework for analytical metadata for comics

Valerio De Luce^(a)

a) University of Macerata, <https://orcid.org/0009-0006-6190-6406>

Contact: Valerio De Luce, valeriodeluce@libero.it

Received: 17 December 2025; **Accepted:** 03 June 2026; **First Published:** 15 September 2026

ABSTRACT

The study addresses the inadequacy of traditional cataloguing standards applied to comics, a historically marginalised medium, proposing a practical solution consistent with the ongoing debate. The framework developed here, based on the principles of Augmented Humanities, integrates GPT-4o for the automatic extraction of metadata from a sample of 100 digital comics. The results show high accuracy in extraction, filling in numerous fields of an analytical bibliographic record proposed and validated for the experiment. The main objective is to demonstrate the potential of AI in improving the efficiency of cataloguing practices. However, the study highlights persistent problems such as hallucinations, difficulties in handling incomplete historical data on the material, and the need for human supervision for ambiguous fields. Finally, the work outlines a roadmap for integrating AI into descriptive processes, promoting the bibliographic recognition of comics as a cultural resource.

KEYWORDS

Generative artificial intelligence; Bibliographic description of comics; Automated metadata generation; Library and information science.

Sviluppo di un framework semi-automatizzato per la metadattazione analitica di fumetti

ABSTRACT

Lo studio affronta l'inadeguatezza degli standard catalografici tradizionali applicati al fumetto, questo medium storicamente marginalizzato, proponendo una soluzione pratica coerente con il dibattito in corso. Il framework qui sviluppato, fondato sui principi delle Augmented Humanities, integra GPT-4o per l'estrazione automatica di metadati da un campione di 100 fumetti digitali. I risultati mostrano un'elevata accuratezza nell'estrazione, compilando numerosi campi di una scheda bibliografica analitica proposta e validata per l'esperimento. L'obiettivo principale è dimostrare il potenziale dell'IA nel migliorare l'efficienza delle pratiche catalografiche. Tuttavia, lo studio evidenzia problemi persistenti come le allucinazioni, le difficoltà nella gestione di dati storici incompleti sul materiale e la necessità di supervisione umana per campi ambigui. Infine, il lavoro delinea una roadmap per l'integrazione dell'IA nei processi descrittivi, promuovendo il riconoscimento bibliografico dei fumetti come risorsa culturale.

PAROLE CHIAVE

Intelligenza artificiale generativa; Descrizione bibliografica del fumetto; Generazione automatica di metadati; Biblioteconomia e scienze dell'informazione.

1. Introduzione

L'integrazione dell'intelligenza artificiale (IA) nelle pratiche di descrizione automatica di oggetti culturali suscita grande interesse nel dibattito contemporaneo all'interno delle Library and Information Science (LIS) (De Martin, Leporini, e Pellegrino 2021, 480-83). Tale interesse si manifesta con duplice valenza: da un lato emergono preoccupazioni legate alla perdita di controllo umano, all'opacità degli algoritmi e al rischio di riduzione della qualità descrittiva; dall'altro si riconoscono le potenzialità trasformative offerte dal deep learning, se correttamente compreso e integrato nei processi lavorativi del mondo librario e documentale (Guerrini 2025, 23-6).

Queste dinamiche si amplificano in presenza di materiali ad alta componente visuale e narrativa, che spingono gli strumenti innovativi ai limiti delle loro capacità operative. I fumetti costituiscono un caso di studio esemplare; nonostante una crescente rivalutazione culturale e commerciale negli ultimi decenni (Amato 2024, 25-62), sostenuta anche da autorevoli studiosi come Umberto Eco, rimangono storicamente esclusi dai contesti accademici e istituzionali. Ciononostante, si tratta di un medium caratterizzato da una forte ibridazione e stratificazione espressiva.

Tuttavia, la metadattazione del fumetto continua a soffrire di una strutturale inadeguatezza: le informazioni bibliografiche disponibili risultano spesso frammentarie, generiche e non standardizzate, limitandosi a eccezioni isolate (Formiga 2024, 25-45) più che a criteri sistemici (Sferruzza 2023, 297-311). Le conseguenze: il fumetto sfugge ai modelli descrittivi tradizionali, generando un vuoto normativo che non è soltanto tecnico, ma anche epistemologico. Questa carenza è insieme causa ed effetto della marginalizzazione del medium nei contesti bibliografici.

In parallelo, la ricerca internazionale sta esplorando con rinnovato slancio il ruolo dell'IA nella generazione e gestione dei metadati (Soykan, Yuret, e Sezgin 2024, 257-81), anche nel settore dei Comics Studies, con risultati promettenti in termini di analisi visuale, comprensione e produzione (Vivoli et al. 2024, 3-8), anche di schede descrittive (Gamage e Wanigasooriya 2024, 257-84). Le applicazioni più interessanti emergono proprio in quei contesti in cui l'intervento umano è limitato o inefficiente, dimostrando come l'IA possa contribuire a individuare elementi altrimenti trascurati da operatori non specialisti, migliorando così l'accessibilità e la valorizzazione dell'opera.

Il presente contributo persegue due obiettivi principali. Il primo consiste nello sviluppo e nella sperimentazione di un framework operativo per la metadattazione semi-automatizzata di materiali bibliografici visivi, tramite un sistema progettato specificatamente per integrare un Large Visual Language Model (LVLM) (Zhu et al. 2024, 1-3), in particolare il modello GPT-4o. Il corpus di riferimento è costituito da 100 fumetti digitalizzati selezionati dal *repository* Digital Comic Museum.¹ I metadati generati, in formato strutturato (JSON) e non strutturato (TXT), vengono valutati in termini di conformità, coerenza e interoperabilità rispetto alla normativa catalografica vigente.

Il secondo obiettivo nasce dalla constatazione che non esista, ad oggi, almeno in Italia, una normativa specifica per la metadattazione dei fumetti che sia condivisa e accettata uniformemente; questi sono generalmente descritti tramite adattamenti della normativa generale.² In risposta a questa

¹ Digital Comic Museum, Homepage, «We are the best site for downloading free public domain Golden Age Comics», <https://digitalcomicmuseum.com>.

² Con gli standard condivisi, si fa riferimento alle REICAT, a RDA e ad altre linee guida, che però rappresentano un "vestito stretto" per il fumetto, che il catalogatore deve necessariamente adattare forzatamente all'oggetto, parafrasando Sferruzza nel suo articolo del 2023.

lacuna, si propone un modello sperimentale di scheda descrittiva bibliografica analitica, orientato alla valorizzazione del fumetto in quanto oggetto culturale complesso.

Il metodo adottato si ispira ai principi delle Augmented Humanities (Bryant 2021), che concepiscono l'IA come strumento a supporto dell'attività umana, e non come sua sostituzione. L'ambizione si limita alla volontà di contribuire al dibattito scientifico e professionale con una proposta che non introduce nuovi standard normativi, ma un framework adattabile alle diverse esigenze istituzionali e progettuali. In quest'ottica, la metadattazione non è solo un atto tecnico, bensì una forma di mediazione cosciente, critica e culturale, capace di rispecchiare la polifunzionalità del materiale descritto (Guerrini 2020, 13, 86).

2. La normativa attuale: limiti e inadeguatezze

Descrivere i fumetti all'interno di una collezione implica il confronto con un quadro normativo complesso e stratificato, caratterizzato da un'assenza complessiva di protocolli condivisi specifici relativi alla struttura narrativa unitaria, alla completezza dell'arco narrativo e alla distinzione tra volumi autoconclusivi e opere serializzate rispetto alla generica categoria delle monografie (REICAT 1.5.2.C). In Italia, il riferimento principale è costituito dalle REICAT (Regole italiane di catalogazione), mentre per la codifica dei metadati si utilizza il formato UNIMARC, previsto dal Servizio Bibliotecario Nazionale (SBN). A livello internazionale, gli standard di riferimento sono ISBD (*International Standard Bibliographic Description*) e RDA (*Resource Description and Access*), implementati attraverso i formati UNIMARC o MARC21, a seconda del contesto istituzionale.

Sono tuttavia evidenti diversi problemi. Nessuno di questi standard offre una trattazione sistematica del fumetto, che viene solitamente assimilato alle risorse illustrate o a contenuto misto. Le REICAT forniscono indicazioni limitate e frammentarie: la regola 4.5.2.4(B) permette di sostituire "ill." con "fumetti" nella descrizione fisica, mentre la 8.1.2 riconosce l'inscindibilità tra testo e immagine, aspetto certamente importante per il fumetto, ma non sviluppato operativamente. RDA, d'altro canto, presenta riferimenti altrettanto frammentari, collocando i fumetti genericamente tra le risorse a contenuto eterogeneo.

Tra i problemi principali vi è la distinzione tra monografia e pubblicazione seriale: mentre alcuni periodici come *Topolino* sono correttamente trattati come seriali, le graphic novel necessiterebbero di criteri più specifici rispetto alla generica categoria delle monografie (REICAT 1.5.2.C). Anche la gestione di titolo e menzione di responsabilità presenta ulteriori complessità, particolarmente evidenti nei manga, spesso caratterizzati da un doppio frontespizio. Occorre quindi privilegiare quello coerente con il verso di lettura originale dell'opera (da destra a sinistra nei manga giapponesi), generando però contrasti con la prassi operativa standard.

È inoltre essenziale documentare tutte le figure autoriali coinvolte nella realizzazione dell'opera. Questo aspetto genera ancora oggi discussioni tra esperti e professionisti, che faticano a trovare una sintesi condivisa. Un ulteriore tema è rappresentato dalle numerose varianti editoriali e ristampe: richiedono manifestazioni distinte o è sufficiente creare collegamenti tra edizioni, prestando attenzione a titoli varianti e pseudonimi?

A fronte di queste problematiche, alcuni contesti nazionali hanno sviluppato soluzioni specifiche. Tra le esperienze più rilevanti si segnala quella francese, dove le RDA-FR includono disposizioni

esplicite per i fumetti, affrontando sistematicamente sia la forma dell'espressione (6.15.1.3) sia le menzioni di responsabilità condivise (6.27.1.3). Questo approccio consente maggiore coerenza e specificità rispetto al panorama italiano.

Negli Stati Uniti, grazie al lavoro del Graphic Novels and Comics Round Table (GNCRT) dell'ALA, dal 2022 è disponibile una linea guida intitolata *Best Practices for Cataloging Comics and Graphic Novels Using RDA and MARC21* (Bailund et al. 2023, 62-8). Essa fornisce istruzioni operative più chiare, distinguendo nettamente tra comic books seriali e graphic novels. Sebbene pensata per MARC21, questa guida può essere adattata anche a contesti UNIMARC.

Rilevante è anche il Metadata Framework for Manga sviluppato dall'Università di Tsukuba in Giappone, che propone un modello ibrido basato su FRBR, Wikidata e MPEG-7. Questo approccio si distingue per l'attenzione ai metadati e all'accessibilità, includendo buone pratiche del contesto Wikidata (Sugie et al. 2025, 69-97).

Come sottolinea Gupta in un suo articolo del 2025, sarebbe auspicabile l'adozione di uno standard condiviso, ma finora si è proceduto con iniziative isolate (Gupta 2025, 1-4). Questi esempi internazionali dimostrano che la rigidità degli standard tradizionali potrebbe essere superata attraverso adattamenti specifici, che devono però essere accompagnati da strategie mirate sin dalle fasi preliminari di un progetto di catalogazione.

È in questo scenario di inadeguatezza normativa e frammentazione internazionale che emerge con forza la necessità di esplorare soluzioni innovative, tra cui l'impiego dell'IA per l'automazione dei processi descrittivi. L'adozione dell'IA non costituisce solo una scelta tecnologica, ma rappresenta una risposta strategica a criticità strutturali. Le ragioni sono molteplici: in primo luogo, la natura ibrida del fumetto richiede capacità analitiche che superino i limiti delle categorie tradizionali; in secondo luogo, la crescita esponenziale della produzione fumettistica rende insostenibile un approccio esclusivamente manuale alla catalogazione. L'automazione, se sfruttata in modo intelligente, può garantire scalabilità ed elevata coerenza descrittiva.

3. L'IA come risposta alle problematiche catalografiche

L'impiego dell'IA generativa nella metadattazione dei fumetti può rappresentare una risposta concreta ai limiti degli standard tradizionali, offrendo la possibilità di automatizzare processi complessi e garantire maggiore coerenza metodologica nella descrizione. Fin dagli anni '90, la ricerca informatica si è interessata all'elaborazione automatica delle tavole a fumetti, concentrandosi inizialmente sulla segmentazione grafica (Rigaud 2014, 14-6), sul riconoscimento del testo nei balloon (Dutta, Biswas, e Das 2021, 49-62) e sull'analisi semantica sequenziale dei contenuti nei pannelli (Zhang e Hotta 2023, 76-90).

La comprensione automatica dei fumetti costituisce un ambito di ricerca complesso, poiché richiede l'integrazione di contenuti visivi e testuali all'interno di una struttura narrativa multimodale. I modelli basati su computer vision, machine learning e deep learning devono affrontare una scomposizione analitica degli elementi del fumetto come vignette, balloon, personaggi, testi, onomatopее, per consentire analisi precise e interoperabili, finalizzate a un'affidabile comprensione dell'opera (Vivoli et al. 2024, 9-22).

Le principali sfide tecniche derivano dalla natura complessa del medium fumettistico. La segmen-

tazione delle immagini risulta problematica a causa della varietà dei layout, che rende difficile isolare automaticamente le vignette. Ad esempio, un OCR (*Optical Character Recognition*) deve confrontarsi con font irregolari e balloon atipici, che ostacolano l'estrazione accurata del testo (Ponsard, Ramdoyal, e Dziamski 2012, 471-78). Mentre l'associazione tra testo e immagine richiede algoritmi sofisticati per riconoscere chi parli e collegare balloon e personaggi. Contemporaneamente, l'analisi del contenuto narrato necessita di un'elaborazione temporale e logica altrettanto complessa, data la frequente non linearità delle trame.

Nel dominio biblioteconomico, le esperienze condotte dalla Library of Congress hanno portato a test per la generazione automatica di metadati tramite IA (Lincoln et al. 2020, 28-31), mettendo in luce sia le potenzialità della tecnologia sia la necessità di una supervisione umana per garantire conformità semantica e normativa.

Anche il panorama italiano presenta esperienze che, seppur non direttamente focalizzate sui temi di questa ricerca, dimostrano la maturità tecnologica necessaria: un esperimento condotto dal Polo Archivistico dell'Emilia-Romagna (ParER)³ ha recentemente testato l'impiego di modelli di IA per il confronto normativo e il calcolo automatico dell'indice di interoperabilità dei formati. Si tratta di un precedente di rilievo, poiché documenta come l'IA potrebbe essere impiegata nel riconoscere formati di file e confrontarli secondo regole complesse per deliberare delle sintesi normative.

4. Linguaggi, formati, standard per una metadattazione semi-automatizzata

Per delineare un quadro generale dell'infrastruttura descrittiva e semantica entro cui si colloca il presente lavoro, è utile distinguere tre livelli: il primo occupato dai set di elementi e standard di codifica di ausilio nel dominio bibliografico-digitale; il secondo, dai linguaggi del Web semantico e i Linked Open Data; infine, i formati di serializzazione e scambio dati impiegati a livello applicativo.

Sul primo piano, il Dublin Core, set di elementi descrittivi sviluppato per la metadattazione e interoperabilità tra risorse digitali e Web applicato sin da subito in ambito bibliografico-digitale, offre una struttura semplice ma efficace (Jackson et al. 2008, 5-21); MARCXML e UNIMARCXML, rappresentazioni XML dei record MARC21 e UNIMARC, consentono la conversione e lo scambio dei record bibliografici fra sistemi informativi diversi.

Sul piano dei linguaggi semantici, il Resource Description Framework (RDF), modello generale del Web semantico, permette di rappresentare i metadati come triple soggetto-predicato-oggetto, creando una rete di conoscenze collegabili nei Linked Open Data.

Su un piano più alto, quello degli standard di contenuto e dei modelli concettuali della catalogazione, il progetto BIBFRAME, iniziativa Linked Open Data della Library of Congress, supera il formato MARC con un modello articolato in tre entità principali (*Work*, *Instance* e *Item*); lo standard di contenuto RDA, basato sul modello concettuale IFLA-LRM, mantiene l'indipendenza dal formato di codifica e garantisce coerenza semantica fra standard differenti.

³ L'esperienza è stata presentata al convegno Conference on Information and Research Science Connecting to Digital and Library Science (IRCDL 2026) e sarà pubblicata sia nel numero speciale 2026 della rivista Umanistica Digitale, sia negli atti della conferenza. <https://ircdl2026.unimore.it/>.

È fondamentale chiarire il fine e l'obiettivo di una descrizione per determinare quali standard, formati di codifica e linguaggi adottare. Non si tratta soltanto di prassi legate a specifici contesti geografici, ma anche di scelte strategiche definite a monte, in funzione di obiettivi progettuali. A questi piani si aggiunge inoltre quello, più strettamente operativo, dei formati di serializzazione e scambio dati fra applicativi, dal quale il framework qui presentato attinge per ragioni legate alla pipeline di elaborazione.

Nell'esperimento presentato nei paragrafi seguenti, si è scelto di adottare, accanto al testo normale, il formato JSON, formato di serializzazione di scambio dati fra applicativi, come principale output. Questa scelta è motivata dalle caratteristiche intrinseche dei modelli di linguaggio: gli LVLN e gli LLM sono ottimizzati e affidabili quando si tratta di generare o convertire in JSON, in quanto linguaggio strutturato standard per l'interazione tra programmi. Il formato JSON presenta il vantaggio di essere facilmente leggibile anche per gli operatori umani, inoltre è facilmente convertibile in altri metalinguaggi e non richiede strutture complesse predefinite. Ad oggi rappresenta uno standard per l'interazione semplice tra sistemi e applicazioni.

5. Fondamenti metodologici: Augmented Humanities e principi etici

La componente sperimentale del framework risiede nell'automazione della metadattazione attraverso l'impiego di IA, applicata in modo analitico e specificamente orientato ai fumetti digitalizzati. L'esperimento si è focalizzato sulla qualità, sulla struttura e sulla standardizzazione dei metadati generati, con l'obiettivo di ottenere descrizioni coerenti e potenzialmente integrabili nei sistemi informativi bibliografici (Maggi et al. 2023, 373-79).

L'intera metodologia si ispira ai principi delle Augmented Humanities, secondo cui l'IA non sostituisce l'intervento umano, ma ne potenzia le capacità operative (Nguyen e Elbanna 2025, 6-9). Questa visione collaborativa ha consentito di perseguire obiettivi fondamentali, tra cui l'affidabilità e la qualità dei risultati sperimentali. Nel framework sviluppato, ad esempio, l'IA viene impiegata per l'estrazione automatica dei metadati visivi e testuali, mentre la validazione finale e la risoluzione di ambiguità rimangono prerogative dell'operatore umano specializzato.

Tale impianto metodologico mira a garantire qualità dei metadati e rispetto etico nel lavoro catalografico (Floridi et al. 2018, 689-707), seguendo i principi di trasparenza, equità, non maleficenza, responsabilità e privacy delle linee guida globali sull'IA (Jobin, Ienca, e Vayena 2019, 389-99). Seguendo questo tracciato, il framework è stato concepito per essere replicabile, scalabile e applicabile anche ad altri oggetti bibliografico-documentali, soprattutto nei casi in cui la componente visiva sia preponderante e non esistano linee guida consolidate o uniformemente condivise.

6. Selezione del corpus e preparazione dei dati

Il progetto ha avuto inizio a partire da un corpus di 100 fumetti digitali (circa 50.000 immagini JPEG) selezionati dal Digital Comic Museum, un *repository* online di fumetti della Golden Age⁴

⁴ Golden Age comics fa riferimento ad uno "standard" o attributo che viene conferito a quei fumetti che condividono le caratteristiche di quelle opere prodotte fra il 1938 e il 1956 circa.

di pubblico dominio. I metadati inseriti manualmente nel *repository* da un utente sono stati utilizzati come *ground truth* (riferimento di confronto considerato veritiero per la valutazione delle risposte del modello) per la validazione degli output generati dall'IA, al fine di ridurre bias soggettivi nella valutazione.

Dunque, tutti i fumetti sono stati selezionati facendo riferimento al lavoro di un unico utente, le cui schede presentavano campi descrittivi completi, coerenti e conformi allo schema bibliografico analitico adottato nel prompt. Il corpus è stato inizialmente riorganizzato in un database Excel per facilitare l'analisi comparativa finale dei metadati estratti dall'IA. L'uniformità dei metadati ha garantito la coerenza nei criteri descrittivi, facilitando la valutazione dell'esperimento.

La selezione ha privilegiato i fumetti degli anni '40-'60 sia per la loro disponibilità in pubblico dominio, sia per assicurare la riproducibilità dell'esperimento, includendo al contempo un'ampia varietà di generi, stili e autori. Al momento, questa scelta è stata condizionata dall'assenza di altri casi studio o di istituzioni disponibili a fornire materiale simile e più recente, senza vincoli legali. Va segnalato che nei record preesistenti mancavano codici identificativi (come ISBN o EAN); tali campi sono stati quindi esclusi, insieme ad alcune informazioni presenti nella scheda e richieste dal prompt. Poiché i fumetti analizzati appartengono a periodi in cui gran parte di questi dati non veniva ancora registrata, essi risultano del tutto assenti e, di conseguenza, impossibili da estrarre direttamente dai file digitalizzati. Si è pertanto ritenuto necessario escluderli dalla valutazione finale, al fine di evitare distorsioni nei risultati. Questa esclusione, pur rappresentando una scelta pragmatica per il progetto pilota, mette in evidenza una delle principali sfide legate all'uso dell'IA nel settore del patrimonio culturale: la gestione di dati storici incompleti o disomogenei.

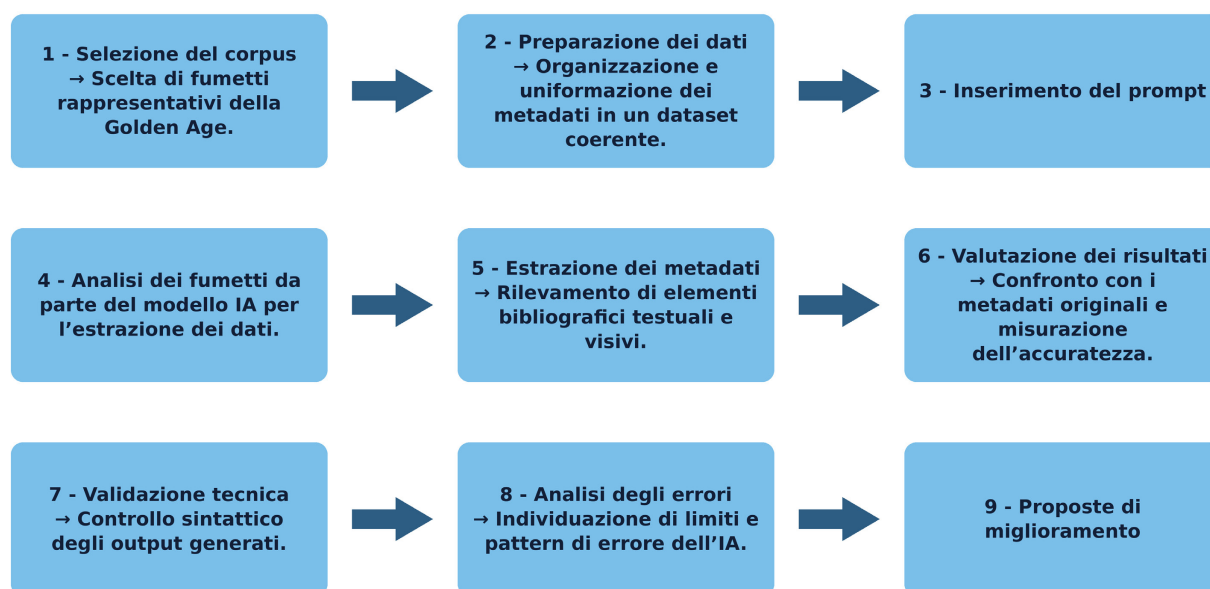


Figura 1. Schema del framework.

7. Architettura del framework e implementazione tecnica

I fumetti sono stati selezionati e caricati singolarmente in un programma progettato per integrare il modello GPT-4o di OpenAI, considerato al momento dell'esperimento il più avanzato e accessibile nel combinare elaborazione testuale e visiva. L'obiettivo era la generazione semi-automatizzata di schede descrittive di tipo bibliografico-analitico, in linea con gli standard catalografici vigenti.

Il processo si articola in tre macro-fasi operative:

1. Estrazione dei dati: ogni immagine è inviata all'LVLMM tramite un prompt ottimizzato, frutto di un processo iterativo di sintesi normativa e test sperimentali.
2. Analisi e trasformazione: il modello restituisce l'output in formato testuale (.txt) e strutturato (JSON), comprensivo di metadati bibliografici, elementi fisici e caratteristiche contestuali.
3. Valutazione qualitativa: i risultati vengono confrontati con i metadati ufficiali secondo tre livelli di accuratezza:
 - VV: completamente corretto e pertinente
 - V: parzialmente corretto/incompleto ma rilevante
 - X: errato o non pertinente

Particolare attenzione è stata riservata alla coerenza delle descrizioni, includendo anche attributi spesso trascurati, come qualità di stampa, note editoriali e presenza di contenuti pubblicitari. Il sistema è stato implementato su Google Colab con l'utilizzo della libreria Python Pillow per la gestione delle immagini, e la libreria ufficiale OpenAI per l'interazione con le API.

Campo chiave	Ground Truth (DCM)	Output IA (GPT-4o)	Valutazione	Nota qualitativa
Titolo	Jumbo Comics #1	Jumbo Comics n. 1	VV	Titolo corretto, numerazione adattata
Data di pubblicazione	Settembre 1938	1938	V	Mese assente, ma anno corretto
Editore	Fiction House	Fiction House	VV	Dato perfettamente allineato
Prezzo di copertina	10€	-	X	Dato mancante

Tabella 1. Esempi di comparazione dei risultati dell'esperimento.

8. Progettazione del prompt per l'estrazione di metadati bibliografici

Un elemento vitale del progetto è stata la realizzazione del prompt, strettamente connessa allo studio della normativa di riferimento per il trattamento catalografico dei fumetti. L'elaborazione del prompt ha seguito un approccio rigoroso, fondato su un'analisi comparativa degli standard catalografici internazionali e sulle esigenze specifiche del materiale fumettistico. Ci si è avvalsi dell'IA per attuare delibere e confrontare i vari standard, della consulenza di esperti del settore per completare e validare la scheda, tra cui docenti in ambito archivistico e IA, dirigenti dell'ICCU, dipendenti delle Biblioteche Nazionali Centrali di Roma e Firenze, membri dei tavoli tematici per la catalogazione del fumetto, nonché professionisti dell'ambito fumettistico.

La struttura del prompt riflette le categorie ritenute più rilevanti della descrizione bibliografica, adattate alle peculiarità del medium fumettistico, ed è organizzata in sezioni. Il modello così ha

potuto operare secondo una logica strutturata, volta a minimizzare omissioni e incongruenze nella descrizione bibliografica analitica.

Un aspetto innovativo del prompt riguarda l'inclusione di informazioni commerciali, raramente considerate nei sistemi catalografici tradizionali, che possono offrire spunti importanti per analisi storico-economiche o sociologiche legate alla circolazione e ricezione del fumetto.

Il prompt si conclude con una sezione dedicata alla classificazione e indicizzazione. Tuttavia, nella pratica, non è stato possibile ottenere risultati coerenti per questa sezione. Le cause principali sono da ricondurre alla specificità del materiale e all'assenza di dati espliciti nel corpus di partenza.

Per garantire l'interoperabilità con sistemi bibliografici, il prompt prevede l'esportazione dei dati in due formati di output complementari: un file di testo lineare (per la consultazione e la verifica umana) e un file JSON strutturato (per l'elaborazione automatizzata e l'integrazione con sistemi informativi). Questa doppia modalità di output facilita sia la verifica qualitativa sia l'eventuale conversione dei dati mediante mapping dei campi verso formati standard.

Nella progettazione del prompt è stata posta attenzione alla chiarezza espositiva e alla precisione terminologica, adottando un registro tecnico-scientifico adatto al contesto biblioteconomico. L'istruzione introduttiva "*act as an expert librarian and cataloger*" è stata formulata per orientare il modello verso uno standard di accuratezza e coerenza tipico della prassi professionale consolidata. Il prompt utilizzato nell'esperimento è riportato per intero in Appendice A (Figura 2); un esempio rappresentativo del corrispondente output è riportato in Appendice B (Figure 3 e 4).

Il processo di affinamento del prompt ha richiesto numerose iterazioni, test e ottimizzazioni. In ciascuna iterazione, gli output generati sono stati confrontati sistematicamente con il *ground truth*.

9. Sperimentazione e valutazione dei risultati

La valutazione dei risultati è stata condotta confrontando i metadati estratti automaticamente dal modello con quelli già presenti nel *repository* online e ricreati in un database Excel colmando eventuali lacune con altre fonti, in alcuni casi con un confronto diretto con esperti. I risultati ottenuti sono complessivamente incoraggianti: il 71,88% dei campi è stato valutato come pienamente corretto (VV), il 12,80% come parzialmente corretto (V), e solo il 15,32% come errato o mancante (X). Ciò corrisponde a circa 3.200 metadati corretti su un totale stimato di 4.500, considerando 53 campi estraibili, di cui 8 esclusi per evitare bias nella valutazione.

L'aggregato positivo (VV + V), pari all'84,68%, suggerisce un buon grado di affidabilità del framework, che si dimostra già parzialmente efficace, pur con ampi margini di miglioramento. Rispetto ad approcci documentati in letteratura, spesso basati su OCR generici o flussi semi-automatizzati di metadattazione, come nei progetti Comic Book Markup Language (CBML) (Walsh 2012, 1-4) o OCR-D, il framework proposto si distingue per un più elevato grado di automazione e una maggiore adattabilità alle specificità del fumetto.

I risultati sono stati ottenuti senza ricorrere a complesse fasi di pre-elaborazione (Lenadora et al. 2020, 5-8) o pre-addestramento. Un pre-addestramento mirato avrebbe contribuito a una maggiore uniformità degli output, che attualmente variano in modo arbitrario senza compromettere la qualità complessiva.

L'analisi qualitativa ha evidenziato una variabilità delle prestazioni a seconda delle categorie di

metadati. Alcuni campi hanno raggiunto un'accuratezza superiore all'85%; in quei casi il modello si è dimostrato efficace nell'identificazione e classificazione degli elementi visivi. Al contrario, altri campi hanno mostrato prestazioni meno uniformi, con tassi di correttezza compresi tra il 40% e il 70%. In quei casi il modello tenta inferenze basate su segnali visivi e testuali, ma la qualità delle scansioni o la scarsa leggibilità del materiale ne compromette l'accuratezza, generando dati parziali o errati.

È interessante notare che, in linea generale, il modello ha identificato i personaggi con maggiore puntualità e dettaglio rispetto a quanto indicato dall'utente che ha inserito i dati originali nel *repository*. Alcune categorie sono risultate più critiche: gli identificatori standard (es. ISBN) sono del tutto assenti, poiché i fumetti in analisi sono antecedenti alla loro introduzione; le menzioni di responsabilità (come sceneggiatore, illustratore, inchiostatore⁵) presentano un tasso di errore superiore al 75%, principalmente a causa della frequente mancanza di tali informazioni nei materiali esaminati. Analogamente, campi di natura contestuale o storica, come premi ricevuti, ristampe o note sulla prima pubblicazione, tutti dati che richiederebbero l'accesso a fonti esterne non disponibili al modello, risultano non compilati o errati.

Un ulteriore elemento critico è rappresentato dal fenomeno delle allucinazioni, ovvero la generazione di dati plausibili ma non supportati da evidenze concrete. Questi errori emergono in particolare nell'attribuzione dei ruoli creativi, dove il modello tende a replicare pattern frequenti piuttosto che estrarre fedelmente le informazioni effettivamente presenti nel materiale.

Sebbene nella maggior parte dei casi le allucinazioni si limitino a singole voci o campi marginali, è assai rilevante sottolineare che in circa dieci casi il modello ha prodotto schede interamente difformi dallo schema previsto. Si tratta di errori che hanno dato origine a brevi report narrativi o generalizzazioni, anziché seguire la struttura tabellare richiesta. Questo comportamento evidenzia una vulnerabilità di fondo del sistema, che, sebbene in fase sperimentale possa essere tollerata, pone un ostacolo rilevante all'adozione in contesti produttivi.

Tuttavia, l'analisi empirica ha mostrato che reiterando la richiesta sullo stesso fumetto fino a un massimo di tre volte, si è generalmente ottenuta una scheda conforme allo schema previsto, suggerendo che il tipo di errore sia sistematico ma la sua comparsa episodica. Questo dato rafforza la necessità di meccanismi di controllo, validazione automatica e riesecuzione condizionata (Wang et al. 2023), che possano intercettare le risposte alterate e garantire l'uniformità degli output. L'integrazione di questi strumenti rappresenta un passo importante per l'evoluzione del framework da ambiente sperimentale a sistema affidabile per l'estrazione automatica di metadati.

10. Riflessioni finali e prospettive future

L'esperimento condotto ha dimostrato come l'impiego dell'IA generativa nella metadattazione bibliografica analitica⁶ dei fumetti potrebbe costituire un valido supporto al lavoro catalografico,

⁵ Si intende quell'insieme di professionisti con un ruolo esplicito o derivato che ha portato alla realizzazione di un'opera fumettistica, ad esempio, sceneggiatore, disegnatore, inchiostatore, colorista, etc.

⁶ Va specificato, anche se già più volte indicato nella trattazione come riferimento, cosa si intende con l'attributo "analitica" in riferimento alla descrizione bibliografica. Per farlo è necessario spostare la lente direttamente sul lavoro del catalogatore, si tratta di una prassi spesso ripetitiva e condizionata da tempi ristretti, a causa dell'elevata mole di materiali da trattare.

pur evidenziando alcune limitazioni. In assenza di meccanismi automatici di validazione tramite risorse esterne, la qualità delle scansioni e la presenza effettiva delle informazioni nei documenti rimangono fattori determinanti per l'accuratezza degli output.

Come accade anche per un catalogatore umano, alcune categorie di metadati non risultano realisticamente estraibili, né attribuibili dal modello in maniera autonoma qualora siano assenti nel materiale originario. È il caso, ad esempio, della classificazione bibliografica o della determinazione della fascia d'età consigliata, che richiederebbero l'accesso a repertori normativi esterni non disponibili al sistema. Nonostante queste difficoltà, il modello ha mostrato una sorprendente coerenza nell'inferenza di tali dati, con attribuzioni plausibili nel 90% dei casi.

Sulla base dei risultati ottenuti, e con un'attenta selezione dei campi da includere, si può affermare che il metodo e il framework risultano già oggi coerenti ed efficaci come supporto al lavoro umano, in particolare se applicati alla redazione di schede conformi alle REICAT e secondo lo schema ISBD, notoriamente più sintetico rispetto alla descrizione analitica adottata in questa ricerca.

I risultati complessivi, con l'84,68% dei metadati valutati come corretti o parzialmente corretti, confermano la solidità metodologica del framework. Inoltre, l'output in formato JSON è risultato formalmente corretto nel 100% dei casi, ovvero conforme alla sintassi prevista, a prescindere dal contenuto semantico, attestandone l'affidabilità tecnica.

Ciononostante, la ricerca ha evidenziato diverse criticità che, allo stato attuale, ostacolano una catalogazione complessiva davvero efficace e affidabile. In primo luogo, per il fumetto, l'inadeguatezza degli standard catalografici tradizionali risulta evidente: questi non sono progettati per descrivere in modo esaustivo il fumetto, trattandolo piuttosto come una generica risorsa illustrata. A questo si aggiunge un'ampia frammentazione normativa e la mancanza di protocolli condivisi. Inoltre, nei fumetti "storici" come quelli del corpus è evidente la mancanza di elementi chiave come ISBN, informazioni complete sugli autori e date di pubblicazione precise. Tale lacuna compromette l'accuratezza dell'estrazione automatica.

Le limitazioni tecniche del modello emergono anche attraverso fenomeni di allucinazione; errori ricorrenti nei campi relativi alle menzioni di responsabilità e ai contenuti storici; difficoltà nell'elaborazione coerente di sequenze visive, a causa dell'assenza di memoria tra immagini consecutive. A ciò si aggiunge l'impossibilità, per il modello, di verificare dati contestuali o storici senza accesso a fonti esterne o authority file.

Nonostante questi limiti, l'esperimento ha apportato contributi rilevanti. È stato sviluppato un framework semi-automatizzato originale, basato su un modello specifico e ispirato all'approccio delle Augmented Humanities, progettato per generare schede bibliografiche analitiche da fumetti digitalizzati. Si è giunti alla definizione di un prompt specifico per questo compito, e di una scheda per il fumetto, che include elementi descrittivi bibliografici estesi, adattati alla natura del medium, con metadati sia testuali che strutturati.

L'osservazione dei fenomeni e dei risultati conferma che l'approccio proposto rappresenta una concreta opzione per implementare l'automazione di processi catalografici nel dominio del fumetto e, potenzialmente, in altri ambiti documentari con caratteristiche analoghe. Questo progetto

Per questo motivo, la descrizione tende a essere sintetica e soggetta a regole molto rigide. In questo caso, invece, si intende espandere la descrizione, andando a cogliere quelle informazioni che verrebbero escluse in una descrizione bibliografica standard.

costituisce quindi un primo passo promettente in un percorso evolutivo che, partendo dalla fase sperimentale, miri alla definizione di un protocollo operativo scalabile e replicabile, orientato all'integrazione con strumenti e risorse specialistiche, e alla validazione semantica dei metadati generati, identificata come sviluppo futuro della ricerca.

Le difficoltà riscontrate nell'applicazione della normativa catalografica vigente al fumetto, unitamente alla marginalizzazione istituzionale del medium, rendono ancor più urgente l'adozione di approcci innovativi come quello proposto, da collocare in una prospettiva dichiaratamente interdisciplinare. L'esperienza ha evidenziato la capacità del modello di interpretare contenuti ibridi, combinando efficacemente elementi testuali e visivi. Una flessibilità che apre alla possibilità di estendere l'applicazione del framework ad altri oggetti documentari, anche non illustrati.

Persistono numerose sfide: le cosiddette allucinazioni e le difficoltà nell'estrazione di dati non esplicitamente presenti dimostrano che, allo stato attuale, l'IA non sia in grado di sostituire l'expertise umana, né di compensare l'assenza di fonti esterne. Si tratta di limiti che non dipendono solo dalla tecnologia, ma anche dalla qualità e completezza del materiale di partenza, e riflettono i confini attuali delle capacità inferenziali dei modelli generativi.

Un'ulteriore direzione di sviluppo riguarda la possibilità di integrare automaticamente i metadati estratti nei sistemi informativi bibliografici, affrontando le sfide dell'interoperabilità e del mapping tra standard descrittivi eterogenei. Prima di ciò, sarà necessaria una riflessione più approfondita sull'accesso e sull'arricchimento semantico, volutamente escluso da questa fase sperimentale per non compromettere gli obiettivi principali del progetto. Si tratta di un tema che costituisce una delle sfide più importanti per una piena integrazione dell'IA nelle pratiche catalografiche avanzate.

Tutto l'impianto di questo studio è stato orientato alla ricerca di soluzioni operative e concrete a problematiche strutturali, che affliggono una delle risorse bibliografiche più diffuse, amate e al contempo trascurate. Il dibattito sulla catalogazione del fumetto appare ancora oggi frammentato e stagnante, con esperienze virtuose ma isolate.

L'esperienza documentata in questo lavoro costituisce un contributo non solo tecnico, ma anche teorico e metodologico, indicando una possibile via per l'integrazione consapevole e responsabile dell'IA nel trattamento di materiali bibliografici complessi, là dove l'intervento umano fondato su criteri standardizzati risulti difficile, se non impossibile. L'approccio proposto conferma che l'IA, se correttamente guidata, può rappresentare un alleato strategico nella valorizzazione del patrimonio culturale, in particolare per materiali non convenzionali, aprendo nuove prospettive per la descrizione, l'accesso e la diffusione della conoscenza.

Riferimenti bibliografici*

Amato, Maria Concetta. 2024. "Il mercato del fumetto in Italia: analisi e prospettive". Tesi di laurea, Università degli Studi di Genova. <https://unire.unige.it/bitstream/handle/123456789/9749/tesi30240623.pdf>

Bailund, Allison, Steven W. Holloway, Kayla Kuni, e Deborah Tomaras. 2023. "Reprints, Reboots and Retcons: Standardizing Comics Cataloging with the Best Practices Guide from the GN-CRT." *Art Libraries Journal* 48 (3): 62-8. <https://doi.org/10.1017/alj.2023.11>.

Bryant, Peter T. 2021. "Augmented Humanity: Being and Remaining Agentic in a Digitalized World." In *Augmented Humanity*, Springer Nature. <https://link.springer.com/book/10.1007/978-3-030-76445-6>.

De Martin, Chiara, Barbara Leporini, e Gregorio Pellegrino. 2021. "Verso la descrizione automatica delle immagini nell'editoria digitale accessibile: proposta di una tassonomia di immagini per gli algoritmi di IA." In *DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale. Raccolta degli abstract estesi della 10ª conferenza nazionale*, 480-83. Roma: Associazione per l'Informatica Umanistica e la Cultura Digitale. <https://arpi.unipi.it/handle/11568/1219148>.

Dutta, Arpita, Samit Biswas, e Amit Kumar Das. 2021. "CNN-Based Segmentation of Speech Balloons and Narrative Text Boxes from Comic Book Page Images." *International Journal on Document Analysis and Recognition (IJDAR)* 24 (1): 49-62. <https://doi.org/10.1007/s10032-021-00366-4>.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, e Effy Vayena. 2018. "AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28: 689-707. <https://doi.org/10.1007/s11023-018-9482-5>.

Formiga, Federica. 2024. "Il fumetto nelle biblioteche: un percorso distributivo in movimento." *Sistema Editoria. Rivista internazionale di studi sulla contemporaneità* 2 (2): 25-45. <https://doi.org/10.14672/se.v2i2.2703>.

Gamage, Ruwan, e Priyanwada Wanigasooriya. 2024. "Using Generative AI for Bibliographic Description: A Study with ChatGPT-4." *Journal of the University Librarians Association of Sri Lanka* 27 (2): 257-84. <https://doi.org/10.4038/jula.v27i2.8083>.

Guerrini, Mauro. 2020. *Dalla catalogazione alla metadattazione. Tracce di un percorso*. Roma: Associazione Italiana Biblioteche.

Guerrini, Mauro. 2025. "La rappresentazione dell'universo bibliografico e delle collezioni in era digitale: l'entity modeling." *Biblioteche Oggi Trends* 11 (1): 23-6. <https://www.bibliotecheoggi-trends.it/it/articolo/4120/la-rappresentazione-dell-universo-bibliografico>.

Gupta, Asmita. 2025. "Common Issues and Uncommon Solutions to Cataloguing Comic Books in Libraries." In *Proceedings of the Annual Conference of CAIS / Actes du congrès annuel de l'ACSI*. <https://doi.org/10.29173/cais1891>.

* Ultima data di consultazione dei siti web: 25 gennaio 2026.

Jackson, Amy S., Myung-Ja K. Han, Kurt Groetsch, Megan Mustafoff, e Timothy W. Cole. 2008. "Dublin Core Metadata Harvested Through OAI-PMH." *Journal of Library Metadata* 8 (1): 5-21. https://doi.org/10.1300/J517v08n01_02.

Jobin, Anna, Marcello Ienca, e Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1: 389-99. <https://doi.org/10.1038/s42256-019-0088-2>.

Lenadora, Damitha, Rakhitha Ranathunge, Chamath Samarawickrama, Yumantha De Silva, Indika Perera, e Anuradha Welivita. 2020. "Extraction of Semantic Content and Styles in Comic Books." *International Journal on Advances in ICT for Emerging Regions (ICTer)* 13 (1). <https://doi.org/10.4038/icterv13i1.7212>.

Lincoln, Matthew, Julia Corrin, Emily Davis, e Scott B. Weingart. 2020. "CAMPI: Computer-Aided Metadata Generation for Photo Archives Initiative." Library of Congress. <https://doi.org/10.1184/R1/12791807>.

Maggi, Roberta, Tiziana Pasciuto, Martina Mazzoleni, Maria Teresa Artese, Isabella Gagliardi, e Riccardo Albertoni. 2023. "GECA 3.0-A New Tool for Cataloguing and Enjoying Cultural Heritage." In *La memoria digitale: forme del testo e organizzazione della conoscenza. Atti del XII Convegno Annuale AIUCD*, 373-79. Siena: Università degli Studi di Siena. <https://doi.org/10.6092/unibo/amsacta/7721>.

Nguyen, Trinh, e Amany Elbanna. 2025. "Understanding Human-AI Augmentation in the Workplace: A Review and a Future Research Agenda." *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-025-10591-5>.

Ponsard, Christophe, Ravi Ramdoyal, e Daniel Dziamski. 2012. "An OCR-Enabled Digital Comic Books Viewer." In *Computers Helping People with Special Needs: 13th International Conference, ICCHP 2012, Linz, Austria, July 11-13*, 471-78. Berlin: Springer. https://doi.org/10.1007/978-3-642-31522-0_71.

Rigaud, Christophe. 2014. "Segmentation and Indexation of Complex Objects in Comic Book Images". PhD dissertation, Université de La Rochelle, 2014. <https://theses.hal.science/tel-01221308v1>.

Sferruzza, Marco. 2023. "I fumetti in biblioteca: una palestra per la catalogazione." *AIB Studi* 63 (2): 297-311. <https://doi.org/10.2426/aibstudi-13878>.

Soykan, Gürkan, Deniz Yuret, e Tevfik Metin Sezgin. 2024. "ComicBERT: A Transformer Model and Pre-training Strategy for Contextual Understanding in Comics." In *Document Analysis and Recognition-ICDAR 2024 Workshops*, edited by Harold Mouchère e Anqi Zhu. Cham: Springer. https://doi.org/10.1007/978-3-031-70645-5_16.

Sugie, Noriko, Akiko Hashizume, Djoke Dam, e Mari Agata. 2025. "Are Japanese Manga Held in Dutch Libraries? An Analysis of Collection Trends Using Library Metadata." *Journal of Library Metadata* 25 (2): 69-97. <https://doi.org/10.1080/19386389.2025.2469993>.

Vivoli, Emanuele, Mohamed Ali Souibgui, Andrey Barsky, Artemis LLabrés, Marco Bertini, e Dimosthenis Karatzas. 2024. "One Missing Piece in Vision and Language: A Survey on Comics Understanding." Preprint, submitted September 14, 2024. <https://doi.org/10.48550/arXiv.2409.09502>.

Wang, Xuezhi, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, e Denny Zhou. 2023. "Self-Consistency Improves Chain of Thought Reasoning in Language Models." In International Conference on Learning Representations (ICLR 2023). <https://arxiv.org/abs/2203.11171>.

Walsh, John A. 2012. "Comic Book Markup Language: An Introduction and Rationale." *Digital Humanities Quarterly* 6 (1). <https://dhq.digitalhumanities.org/vol/6/1/000117/000117.html>.

Zhang, Yunlong, e Seiji Hotta. 2023. "Automatic Reading Order Detection of Comic Panels." In Pattern Recognition, Computer Vision, and Image Processing. ICPR 2022 International Workshops and Challenges, 76-90. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-37742-6_6.

Zhu, Deyao, Jun Chen, Xiaoqian Shen, Xiang Li, e Mohamed Elhoseiny. 2024. "MiniGPT-4: Enhancing Vision-Language Understanding with Advanced Large Language Models." In *Proceedings of the 12th International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.2304.10592>.

Appendice A. Prompt impiegato per l'estrazione dei metadati

Il prompt che segue, in lingua inglese, è quello impiegato nell'esperimento descritto nella sezione 8. È stato sottoposto a GPT-4o insieme a ciascun fumetto del corpus, accompagnato dalle immagini delle pagine.

```
Act as an expert librarian and cataloger according to reference standards. Analyze the comic book work I submit to you and extract the bibliographic metadata.
Main Fields:

Proper Title
Alternative or Parallel Title (if present)
Uniform Title
Original Title
Statement of Responsibility

Author/Writer
Artist/Illustrator
Colorist
Letterer
Cover Artist
Other Creative Contributors

Edition Statement
Place of Publication
Publisher
Publication Date
Standard Identifier

Physical and Technical Characteristics:

Physical Description

Numbered pages
Presence of illustrations
Dimensions
Reading Direction
Average Number of Panels per Page

Content Type
Media Type
Carrier Type
Content Language
Color
Paper Type
Binding
Standard Editorial Format

Serial and Collection Information:

Series or Collection
Issue Number
Number in Series or Collection
Editorial Format
Publication Status
Frequency

Content and Genre:

Narrative Genre
Target Audience
Number of Stories Contained
Synopsis or Summary
Main Characters
Reference Narrative Universe

Commercial Information:

Cover Price
Distribution Code
Target Market

Indicate any presence of advertising and if present, what type.
Classification and Indexing:

Decimal Classification
Subjects and Keywords
Age Classification

Notes and Additional Information:

Descriptive Notes
Note on First Publication
Note on Translation
Note on Censorship, Modifications, or Adaptations
Awards and Recognition Received
Reprints, New Editions, or Parallel Editions

Output Format:
Return the extracted data in two formats:

.txt: linear list, field by field, labeled
JSON: hierarchical structure suitable for interoperability with bibliographic systems
```

Figura 2. Prompt utilizzato per l'estrazione automatica dei metadati bibliografici dei fumetti.

Appendice B. Esempio di scheda generata

Si riporta di seguito l'output ottenuto dall'analisi automatica del primo fumetto del corpus (All True Romance, Farrell Publications, marzo 1955), selezionato dal Digital Comic Museum. L'output è prodotto nei due formati richiesti dal prompt, lineare testuale (Figura 3) e JSON (Figura 4).

```
**Main Fields:**  
- Proper Title: All True Romance  
- Alternative or Parallel Title: Intimate Love Stories  
- Uniform Title: All True Romance  
- Original Title: All True Romance  
- Statement of Responsibility: A Farrell Publication  
  
- Author/Writer: Unknown  
- Artist/Illustrator: Ruth Roche  
- Colorist: Unknown  
- Letterer: Unknown  
- Cover Artist: Unknown  
- Other Creative Contributors: Editor - Ruth Roche; Managing Editor - Robert W. Farrell  
  
- Edition Statement: No. 17  
- Place of Publication: New York, N.Y.  
- Publisher: Farrell Publications  
- Publication Date: March 1955  
- Standard Identifier: Not present  
  
**Physical and Technical Characteristics:**  
- Physical Description:  
  - Numbered pages: Not specified  
  - Presence of illustrations: Yes  
  - Dimensions: Standard comic book size  
  - Reading Direction: Left to right  
  - Average Number of Panels per Page: Varies (approx. 6)  
  
- Content Type: Text, illustrations  
- Media Type: Print  
- Carrier Type: Volume  
- Content Language: English  
- Color: Full color  
- Paper Type: Newsprint  
- Binding: Saddle-stitched  
- Standard Editorial Format: Comic book  
  
**Serial and Collection Information:**  
- Series or Collection: All True Romance  
- Issue Number: 17  
- Number in Series or Collection: 17  
- Editorial Format: Comic  
- Publication Status: Ceased  
- Frequency: Bi-monthly  
  
**Content and Genre:**  
- Narrative Genre: Romance  
- Target Audience: General audiences (teenagers and adults)  
- Number of Stories Contained: Multiple (at least 4 identified from image)  
- Synopsis or Summary: Focuses on romantic and dramatic stories  
- Main Characters: Various based on stories  
- Reference Narrative Universe: All True Romance  
  
**Commercial Information:**  
- Cover Price: 10 cents  
- Distribution Code: None identified  
- Target Market: Romance comic readers  
  
- Presence of Advertising: Yes, typical product and novelty advertisements  
  
**Classification and Indexing:**  
- Decimal Classification: Not specified  
- Subjects and Keywords: Romance, drama, relationships  
- Age Classification: Teen+  
  
**Notes and Additional Information:**  
- Descriptive Notes: Includes advice column by Roberta Bennett  
- Note on First Publication: Not the first issue  
- Note on Translation: Not applicable  
- Note on Censorship, Modifications, or Adaptations: None noted  
- Awards and Recognition Received: None noted  
- Reprints, New Editions, or Parallel Editions: None noted
```

Figura 3. Esempio di scheda generata in formato testuale lineare. Fonte: All True Romance, Farrell Publications, marzo 1955.

```
{
  "MainFields": {
    "ProperTitle": "All True Romance",
    "AlternativeOrParallelTitle": "Intimate Love Stories",
    "UniformTitle": "All True Romance",
    "OriginalTitle": "All True Romance",
    "StatementOfResponsibility": "A Farrell Publication",
    "Contributors": {
      "AuthorWriter": "Unknown",
      "ArtistIllustrator": "Ruth Roche",
      "Colorist": "Unknown",
      "Letterer": "Unknown",
      "CoverArtist": "Unknown",
      "OtherCreativeContributors": {
        "Editor": "Ruth Roche",
        "ManagingEditor": "Robert W. Farrell"
      }
    },
    "EditionStatement": "No. 17",
    "PlaceOfPublication": "New York, N.Y.",
    "Publisher": "Farrell Publications",
    "PublicationDate": "March 1955",
    "StandardIdentifier": "Not present"
  },
  "PhysicalAndTechnicalCharacteristics": {
    "PhysicalDescription": {
      "NumberedPages": "Not specified",
      "PresenceOfIllustrations": "Yes",
      "Dimensions": "Standard comic book size",
      "ReadingDirection": "Left to right",
      "AverageNumberOfPanelsPerPage": "Varies (approx. 6)"
    },
    "ContentType": "Text, illustrations",
    "MediaType": "Print",
    "CarrierType": "Volume",
    "ContentLanguage": "English",
    "Color": "Full color",
    "PaperType": "Newsprint",
    "Binding": "Saddle-stitched",
    "StandardEditorialFormat": "Comic book"
  },
  "SerialAndCollectionInformation": {
    "SeriesOrCollection": "All True Romance",
    "IssueNumber": "17",
    "NumberInSeriesOrCollection": "17",
    "EditorialFormat": "Comic",
    "PublicationStatus": "Ceased",
    "Frequency": "Bi-monthly"
  },
  "ContentAndGenre": {
    "NarrativeGenre": "Romance",
    "TargetAudience": "General audiences (teenagers and adults)",
    "NumberOfStoriesContained": "Multiple (at least 4 identified from image)",
    "SynopsisOrSummary": "Focuses on romantic and dramatic stories",
    "MainCharacters": "Various based on stories",
    "ReferenceNarrativeUniverse": "All True Romance"
  },
  "CommercialInformation": {
    "CoverPrice": "10 cents",
    "DistributionCode": "None identified",
    "TargetMarket": "Romance comic readers"
  },
  "Advertising": "Yes, typical product and novelty advertisements",
  "ClassificationAndIndexing": {
    "DecimalClassification": "Not specified",
    "SubjectsAndKeywords": ["Romance", "drama", "relationships"],
    "AgeClassification": "Teen+"
  },
  "NotesAndAdditionalInformation": {
    "DescriptiveNotes": "Includes advice column by Roberta Bennett",
    "NoteOnFirstPublication": "Not the first issue",
    "NoteOnTranslation": "Not applicable",
    "NoteOnCensorshipModificationsOrAdaptations": "None noted",
    "AwardsAndRecognitionReceived": "None noted",
    "ReprintsNewEditionsOrParallelEditions": "None noted"
  }
}
```

Figura 4. Esempio di scheda generata in formato JSON. Fonte: All True Romance, Farrell Publications, marzo 1955.