

Una banca dati per Petrarca e il suo tempo: criteri, modelli e obiettivi¹

«DigItalia» 2-2021
DOI: 10.36181/digitalia-00037

Nicola De Nisco - Università per Stranieri di Perugia
Sandra Gorla - Università degli Studi di Napoli Federico II
Alessia Valenti - Università degli Studi di Siena

L'articolo presenta le fasi di elaborazione e i modelli del database del progetto PRIN Itinera – Italian Trecento Intellectual Network and European Renaissance Advent. Sviluppata attraverso il framework Cadmus, la banca dati raccoglierà le ricerche sui corrispondenti di Petrarca e quelle sui testi, insieme ai dati codicologici e paleografici relativi alla tradizione manoscritta dei Rerum vulgarium fragmenta.

1. Il progetto Itinera

Itinera (Italian Trecento Intellectual Network and European Renaissance Advent) è un progetto finanziato nell'ambito dei PRIN 2017 e diretto da Natascia Tonelli, responsabile dell'unità di Siena, a cui partecipano altre tre unità: Università Roma Tre, Università per Stranieri di Perugia e Università Federico II di Napoli². Il progetto è dedicato allo studio e all'analisi delle reti intellettuali individuabili a partire dalle corrispondenze di Petrarca. Attraverso il censimento e lo studio delle corrispondenze, Itinera intende ricostruire l'universo culturale e storico del Trecento italiano, indagando inoltre la prima fortuna delle opere petrarchesche, che proprio presso i corrispondenti trovarono il primissimo canale di diffusione. L'attenzione si rivolge in particolare alla tradizione manoscritta delle rime del *Canzoniere*. Avviato a gennaio 2020, Itinera vede collaborare mediolatinisti, filologi italiani, italianisti, paleografi, codicologi ed esperti di Digital Humanities e ha tra i suoi obiettivi la costruzione di una banca dati digitale consultabile online in open access, in cui confluiranno le ricerche prosopografiche sui corrispondenti di Petrarca e quelle sui testi insieme ai dati codicologici e paleografici relativi alla tradizione manoscritta dei *Rerum vulgarium fragmenta*. Lo sviluppo della banca dati, che ha impegnato gran parte del primo anno di lavoro ed è tuttora in corso, è stata affidata ad un informatico esterno, Daniele Fusi, ricercatore presso l'Università Ca'

¹ Questo articolo è frutto della stretta collaborazione dei tre autori. All'interno di una concezione e di una elaborazione comune e condivisa, i paragrafi 2, 4.1 e 4.2 sono da attribuire a Nicola De Nisco; il paragrafo 4.4 a Sandra Gorla; i paragrafi 1, 3, 4.3 e 5 ad Alessia Valenti. Il contributo non avrebbe mai visto la luce senza il supporto costante di Daniele Fusi e i generosi consigli di Paul Gabriele Weston: a entrambi va il ringraziamento di chi scrive.

² I responsabili delle tre unità sono rispettivamente Luca Marcozzi, Daniele Piccini e Marco Corsi.

Foscari, che ha lavorato a stretto contatto con alcuni membri del progetto per la modellizzazione dei dati³.

2. Dal corpus al database

I corpora sui quali insiste Itinera sono molteplici. Un primo nucleo comprende le corrispondenze di Petrarca, latine e volgari, in prosa e in versi, e altre tipologie di testi⁴. Il loro studio è funzionale alla ricostruzione delle reti di rapporti che si dipanano a partire da Petrarca, di cui i corrispondenti costituiscono i nodi. Di questi, si intende condurre un primo censimento completo e fornire schede bio-bibliografiche aggiornate ed esaustive. Un altro filone di ricerca è costituito dalla tradizione manoscritta del *Canzoniere*, con particolare attenzione alla prima e più antica: dei codici si intende approntare descrizioni di prima mano complete e aggiornate.

Il gruppo di lavoro ha discusso preliminarmente la tipologia di dati da estrarre e trattare, le modalità di raccolta e di organizzazione degli stessi, i *desiderata* rispetto alle possibilità di interrogazione della banca dati, tenendo conto di esperienze e progetti analoghi e degli standard descrittivi adottati dalle varie discipline. Un censimento preliminare si è reso necessario per individuare categorie funzionali alla schedatura e per organizzare le fasi successive di lavoro. Le diverse unità hanno prodotto dei prototipi di descrizione per ogni corpus, che hanno costituito il punto di partenza per una riflessione strutturata sui requisiti e sulle caratteristiche che il software doveva possedere.

Sulla base di questi presupposti, la scelta è ricaduta sul framework di editing open source Cadmus e sulla collaborazione con il suo sviluppatore Daniele Fusi⁵.

³ La progettazione dei modelli e le successive fasi di test sono state affidate a tre assegnisti impegnati nel progetto, Nicola De Nisco (unità di Perugia), Sandra Gorla (unità di Napoli), Alessia Valenti (unità di Siena), sempre supportati dal professor Paul Gabriele Weston (unità di Siena) e in collaborazione con Daniele Fusi. I risultati del loro lavoro sono stati poi periodicamente sottoposti all'intera équipe per le necessarie verifiche e discussioni e per essere approvati. Per la parte codicologica ha collaborato in un primo momento Daphne Grieco, dottoranda presso la Scuola Superiore Meridionale, ed è stata costante la supervisione del professor Marco Corsi.

⁴ Il corpus include le lettere di Petrarca confluite negli epistolari (*Epystole; Familiarium rerum libri; Res Seniles; Liber sine nomine*), quelle cosiddette *disperse*, le lettere dei corrispondenti a Petrarca, le tenzoni e i testi poetici genericamente di scambio che lo coinvolgono, compresi quelli raccolti nei *Rvf*, i testi composti da amici e ammiratori alla morte del poeta.

⁵ Il framework Cadmus è stato sviluppato da Daniele Fusi nell'ambito di una *research fellowship* presso il *Venice Centre for Digital and Public Humanities* dell'Università Ca' Foscari di Venezia. Cadmus è un progetto open source: tutti i repository relativi a Cadmus e ai progetti basati su di esso, compreso Itinera, sono pubblici e ospitati presso VeDPH GitHub, disponibili all'indirizzo: <<https://github.com/vedph>>. Hanno adottato Cadmus anche i seguenti progetti: "Musisque deoque"; "Catullus online"; PURA (Purism in Antiquity: Theories of Language in Greek Atticist Lexica and their Legacy), PAGES (Priscian's Ars Grammatica in European Scriptoria: A Millennium of Latin and Greek Scholarship); "The Transmission of Ancient Linguistics: Texts and Contexts of the Roman Grammatical Studies"; "Thesaurus dubii sermonis: Digital Critical Collection of Ancient Latin Linguistics (1st century BC – 8th century CE)"; "Thesaurus dubii sermonis: repertorio critico digitale

L'architettura logica di Cadmus permette di costruire sistemi per la creazione web-based di contenuti, ed è in grado di rappresentare ogni tipo di oggetto perché, non ricorrendo a schemi predefiniti e statici, si basa su una modellizzazione dinamica e modulare dei dati che consente di organizzarli in set indipendenti e componibili in strutture più complesse. Rispetto alla prospettiva testo-centrica della marcatura in XML, in cui l'oggetto testuale regge tutto il peso della codifica dei metadati, Cadmus adotta infatti un paradigma incentrato sui dati⁶. La sua struttura semantica risulta dunque particolarmente adeguata alle esigenze di Itinera, il cui database non sarà costituito solo da un corpus di testi, ma in primo luogo dalle informazioni relative agli oggetti schedati, siano essi testi, manoscritti o altro.

L'unità di base dell'architettura di Cadmus è la *part*, un insieme di dati semanticamente coerente, autonomo e virtualmente indipendente, strutturato secondo un proprio modello. Ogni *part* contiene un certo numero di *properties* ed eventualmente *classes* (pattern elementari di dati composti da una serie di *properties*): la somma di più *part* costituisce un *item*, l'unità di grado superiore. Un *item* è dunque un contenitore in grado di rappresentare qualsiasi tipo di oggetto complesso a partire dall'assemblaggio di *part* funzionali a quella rappresentazione.

Il vantaggio di un framework improntato all'*open modelling* risiede nella possibilità di creare *item* altamente specifici e strutturati: è infatti possibile costruire modelli descrittivi per le singole tipologie di oggetti da trattare a seconda dei diversi obiettivi di ricerca. Risulta dunque particolarmente efficace per le finalità di Itinera, che studia tipologie di oggetti molto diversificate. Il carattere modulare di Cadmus permette il riuso di parti ed elementi già formalizzati per la creazione di nuovi modelli (*composite modelling*): *part* largamente generalizzabili possono essere sfruttate per la realizzazione di nuovi *item*, come *properties* e *classes* per nuove *part*. Dal punto di vista operativo, le caratteristiche dell'*open modelling* e del *composite modelling* svincolano dunque la progettazione dall'uso di set predefiniti di modelli, non necessariamente adeguati alle esigenze della ricerca. Al contempo permettono di ricorrere a sottomodelli o modelli preesistenti – se efficaci e funzionali ai propri obiettivi –, evitando la difficoltà di creare ex novo gli *item*. La progettazione in Cadmus risulta inoltre estremamente facilitata perché ogni parte è dotata di un'interfaccia user-friendly che, oltre a servire per il data entry, permette un'immediata visualizzazione grafica della struttura dei dati già in fase di definizione dei modelli.

della riflessione linguistica antica (I sec. a. C. – VIII sec. d. C.); "MapAeg: Buondelmonti's Liber Insularum"; "Inquisitions Graffiti".

⁶ Sui limiti della codifica TEI e sui presupposti teorici del paradigma *data oriented* alla base di Cadmus si veda Daniele Fusi, *Sailing for a Second Navigation: Paradigms in Producing Digital Content*, «Seminari romani di cultura greca», 7 (2018), p. 213-276.

3. La modellizzazione in Cadmus

I prototipi di scheda descrittiva dei vari corpora sono stati progressivamente tradotti nelle componenti modulari di Cadmus a formare una prima versione degli *item*. La progettazione è stata frutto della mediazione tra i requisiti di efficacia funzionale dei modelli rispetto all'architettura logica del sistema e quelli di efficacia rappresentativa rispetto agli obiettivi del progetto, verificata a più riprese in successive fasi di test condotte attraverso l'inserimento di dati reali. Il processo di definizione degli *item* e delle *part* ha coinciso inizialmente con l'aggiunta progressiva di componenti e sottocomponenti che permettessero di rappresentare gli oggetti con un alto grado di specificità e granularità, comportando un incremento di complessità descrittiva. Un successivo confronto tra le strutture dei singoli *item* così formalizzate ha evidenziato aree di parziale sovrapponibilità concettuale, poi ricondotte a modelli uniformi grazie ad un processo di astrazione che puntava in seconda battuta a creare le condizioni per un loro potenziale uso da parte di altri progetti.

È stato così possibile creare modelli che, pur ampiamente generalizzabili, presentano un alto livello di specificità descrittiva. Sono stati predisposti, comunque, campi note a testo libero per l'inserimento di dati non riconducibili ai modelli elaborati, il cui uso sarà tuttavia limitato a casi eccezionali. Allo stesso tempo, il grado di dettaglio nella rappresentazione degli oggetti può essere definito di volta in volta in base alle finalità specifiche della ricerca e a fattori quali la fattibilità del progetto. Vale a dire che è sempre prevista la possibilità di compilare gli *item* nella misura più adatta all'oggetto, agli obiettivi e alle risorse disponibili, grazie alla modularità delle singole componenti e alla discrezionalità del loro impiego.

Il vantaggio di disporre di una base di dati così granulari e strutturati è costituito da possibilità multiple di ricerca, aggregazione (facilitata dall'uso di *ID* human-readable e di *tag*)⁷ ed esportazione (per esempio su mappe geografiche e timeline), che non rinuncia però ad una presentazione più tradizionale dei dati di tipo discorsivo.

Nel corso della progettazione si è cercato di privilegiare un data entry che permettesse il trattamento automatico dei dati oltre che dei metadati. Quando è stato possibile delineare tassonomie o stabilire valori predefiniti, sono state adottate modalità di inserimento controllate: compilazione a scelta multipla, con possibilità di selezionare uno o più valori all'intero di thesauri; valori booleani; sistemi di ranking numerici. Accanto a questo, sono previste stringhe di testo libere numeriche e alfanumeriche.

In alcuni casi si è optato per la definizione di una sintassi che permettesse l'estrazione automatica dei dati dalla stringa inserita, come nel caso degli intervalli nu-

⁷ L'*ID* human-readable è un identificativo univoco di oggetti e singole componenti, assegnato sulla base di convenzioni stabilite all'interno del progetto, che permette l'incrocio degli *item*, anche di tipo diverso. I *tag* sono etichette con le quali è possibile raggruppare oggetti con caratteristiche simili secondo logiche definite di volta in volta a seconda delle esigenze di ricerca, dunque assegnate, come l'*ID*, sulla base di convenzioni definite dal gruppo di ricerca.

merici (*ranges*) per l'indicazione delle carte all'interno delle schede codicologiche o dell'ordine dei componimenti petrarcheschi nella *Poem ranges part*⁸, dotata inoltre di un visual editor specifico in grado di tradurre in serie i singoli elementi che compongono gli intervalli (Fig. 1).

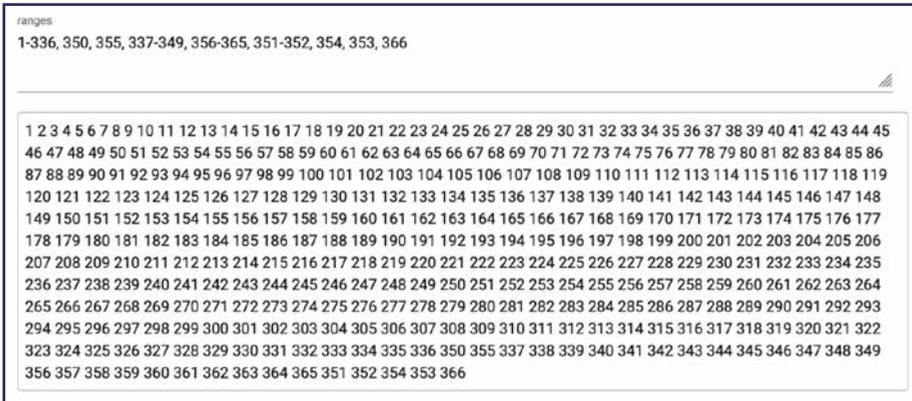


Figura 1. L'interfaccia di editing per l'elemento *ranges* e il relativo visual editor in *Poem ranges part*

In pochi, numerati casi, per facilitare la fruizione dell'utente o quando una maggiore strutturazione dei dati non fosse utile ai fini della ricerca, si è preferito ricorrere ad una formulazione discorsiva, come per il campo *subject* in una delle sezioni degli *item* testuali, dove si danno sommarie indicazioni sul contenuto del testo schedato (Fig. 2). All'interno dell'*item* o per parti specifiche è inoltre previsto il collegamento a risorse esterne (authority file, digitalizzazioni di manoscritti o documenti d'archivio ecc.) attraverso URI o altri tipi di identificativo.



Figura 2. L'interfaccia di editing per il modello *Serial text information part*, con il campo *subject* a testo libero

⁸ Per cui v. *infra*.

4.1 I modelli descrittivi: struttura, caratteristiche e compilazione

Allo stato attuale di avanzamento del progetto, gli *item* elaborati sono i seguenti⁹:

- Lettera (*item Letter*);
- Testo poetico di corrispondenza (*item Poetic text*);
- Testo (*item Text*)¹⁰;
- Corrispondente (*item Correspondent*);
- Manoscritto (*item Manuscript*).

Di seguito se ne illustreranno struttura complessiva e caratteristiche salienti¹¹. Per esigenze di chiarezza, alla descrizione dei singoli *item* è premessa quella delle loro parti comuni.



Figura 3. L'interfaccia di editing per il modello *Historical date*

Per la progettazione si è potuto fare ampio ricorso a componenti già esistenti nel framework Cadmus, come nel caso di *Historical date* per l'inserimento della data (fig. 3), impiegato come elemento singolo o come tassello di parti più complesse e autonome (*Chronotopics part*)¹².

Allo stesso modo, sono stati diffusamente riutilizzati in più parti modelli progettati ad hoc per Itinera. Uno di questi è *Document references*, pensato per inserire a corredo delle informazioni degli *item* fonti puntuali, siano esse rimandi testuali o riferimenti bibliografici (fig. 4). Nel primo caso può trattarsi di citazioni di testi letterari, interni o esterni al corpus, richiamati occasionalmente con un riferimento abbreviato; nel secondo, le voci vengono depositate in forma estesa e completa in un database autonomo e parallelo appositamente elabo-

⁹ Nei mesi successivi alla stesura dell'articolo, i modelli hanno conosciuto modifiche e rielaborazioni anche di un certo rilievo. Cambiando la fisionomia dei modelli, non cambiano però i principi e i presupposti di progettazione qui presentati.

¹⁰ Mentre le lettere saranno descritte attraverso l'*item Letter* e i testi poetici attraverso l'*item Poetic text*, in *Text* confluiranno i dati relativi a opere che attestino in vario modo, ad esempio per mezzo di dediche, i legami tra Petrarca e i suoi contemporanei. Questo *item*, tuttavia, è ancora in fase di perfezionamento e non sarà presentato nella disamina che segue.

¹¹ In considerazione della sede di pubblicazione di questo articolo, si è ritenuto opportuno dedicare una trattazione più ampia all'*item Manuscript*.

¹² Il modello concettuale alla base di *Historical date* è illustrato in D. Fusi, *Sailing for a Second Navigation*, cit., p. 264-266.

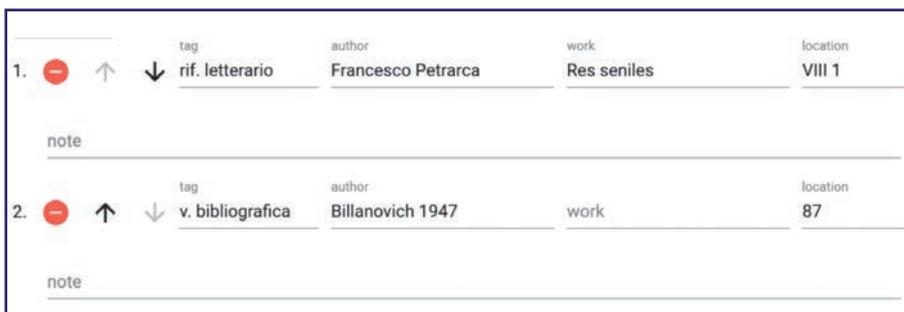


Figura 4. Un riferimento letterario e una voce bibliografica nell'interfaccia di editing per il modello *Document references*

rato, quindi richiamate negli *item* attraverso una chiave¹³. Entrambe le tipologie di richiamo possono essere qualificate con un *tag*. Un altro modello comune è *DecoratedId*, al quale si ricorre ogni volta sia necessario stabilire o proporre una o più identificazioni per persone o cose (fig. 5). Questo modulo associa ad ogni ipotesi di identificazione un *ID* ed eventualmente risorse esterne e fonti bibliografiche, prevedendo la possibilità di indicare per ogni proposta il grado di affidabilità. Un campo riservato ai *tag* permette inoltre di creare sottogruppi coerenti e utili alle logiche della ricerca. *DecoratedId*, in combinazione con i modelli *Person name* e *Document references*, è alla base della *Cited person part*, che permette di collegare un'ipotesi di identificazione alle forme attestata dei nomi¹⁴. Questi elementi, come altri di cui si tratterà in seguito, possono essere re-

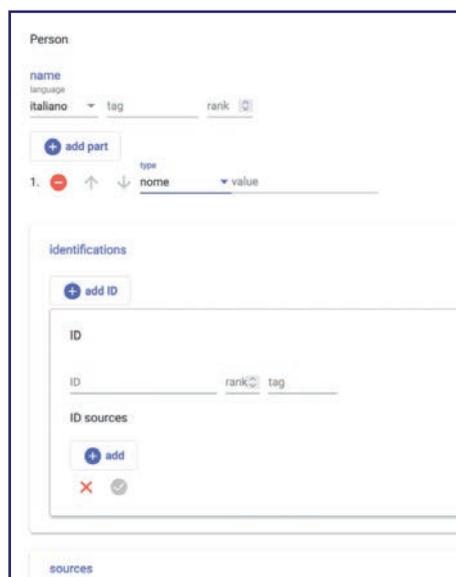


Figura 5. L'interfaccia di editing per il modello *DecoratedId*

¹³ La documentazione relativa al database bibliografico si trova all'indirizzo: <https://github.com/vedph/cadmus_biblioapi>. L'inserimento di nuove voci può avvenire accedendo direttamente a questo database oppure attraverso una *part* dedicata all'interno di ogni *item* di Itinera. In entrambi i casi la voce è registrata nel database bibliografico, dal quale viene richiamata nei singoli *item* secondo un approccio *top-down*.

¹⁴ Questo modello, pensato per sciogliere i riferimenti a personaggi storici citati all'interno dei testi, si è rivelato vantaggioso anche per dare conto di casi, frequenti nei testi poetici, in cui al nome dell'autore attestato nella tradizione manoscritta non corrisponda un'identificazione univoca. Nel primo caso il modello è utilizzato come *part* autonoma, nel secondo costituisce la struttura della sottoparte *author*.

plicati quando l'informazione da inserire attraverso di essi non coincida con uno ma con più dati. Per esempio, nel caso di un'identificazione per la quale esistano più proposte, è possibile inserire tanti *DecoratedId* quante sono le proposte, associando a ciascuna di esse un *rank* di affidabilità.

4.2. Item Letter e Poetic text

Gli *item Letter* e *Poetic text* sono progettati per raccogliere le informazioni relative a testi di vario tipo: lettere di e a Petrarca, testi poetici scambiati tra Petrarca e i suoi corrispondenti o scritti in occasione della morte del poeta (Fig. 6 e 7). Il corpus di partenza comprende testi che hanno un legame diretto con Petrarca, quindi di cui egli sia autore o destinatario; non si esclude, tuttavia, di allargare in futuro la schedatura agli scambi dei corrispondenti di Petrarca tra loro.

Itinera non punta ad estrarre dai testi informazioni di tipo linguistico o contenutistico, ma dati di realtà, che permettano in primo luogo di ricostruire la rete di scambi formatasi attorno a Petrarca (circolazione e scambi di testi, lettere e codici, incontri ecc.), e in seconda battuta di fornire dati completi sui singoli testi petrarcheschi (date e luoghi dichiarati e ricostruiti, persone citate, varianti redazionali ecc.). Ai dati desumibili direttamente dai testi sono quindi affiancati quelli ricavabili dalla bibliografia secondaria¹⁵.

La parziale identità dei dati da raccogliere in questo corpus ha reso possibile l'uso trasversale di numerosi modelli. L'*item Poetic text* è composto da sei *part*, dedicate rispettivamente alle informazioni generali sul testo (*Serial text information part*), ai dati cronologici (*Chronotopics part*), alla registrazione delle persone citate a vario titolo (*Cited persons part*) e di eventuali allegati trasmessi insieme ai testi (*Attachment part*), oltre alle due parti comuni a tutti gli *item*, utilizzate per la registrazione della bibliografia relativa all'oggetto descritto (*External bibliography part*)¹⁶ e per note di carattere generale formulate come stringhe di testo libere (*Note part*). L'*item Letter* prevede, oltre a queste, due parti ulteriori deputate a raccogliere rispettivamente le occorrenze degli pseudonimi (*Pseudonyms part*) e i riferimenti letterari presenti nel testo schedato (*References part*)¹⁷.

Serial text information costituisce la parte più articolata degli *item* testuali ed è organizzata in quattro sezioni: *general*, *recipients*, *reply-to*, *related*¹⁸. La sezione

¹⁵ È stata predisposta anche la pubblicazione dei testi, già editi o di cui si produrranno nuove edizioni nell'ambito del progetto, che costituiranno *item* autonomi da collegare ai rispettivi *item Letter*, *Poetic text* e *Text* attraverso l'identificativo univoco assegnato a ciascuno di essi.

¹⁶ È in fase di discussione la modalità di integrazione della *recensio* dei testimoni (utile soprattutto per le *Disperse* petrarchesche e per le poesie dei corrispondenti) mediante la parte *bibliography*: i manoscritti interessati costituiranno voci opportunamente registrate nel database bibliografico esterno e dotate di un *tag* specifico.

¹⁷ Per ora è previsto l'inserimento di citazioni e allusioni a testi di corrispondenti e autori contemporanei a Petrarca. Non si esclude che in futuro la parte possa essere usata anche per raccogliere sistematicamente fonti letterarie e riferimenti altri.

¹⁸ La sezione *related* è in fase di riprogettazione per essere armonizzata con la *References part* del-

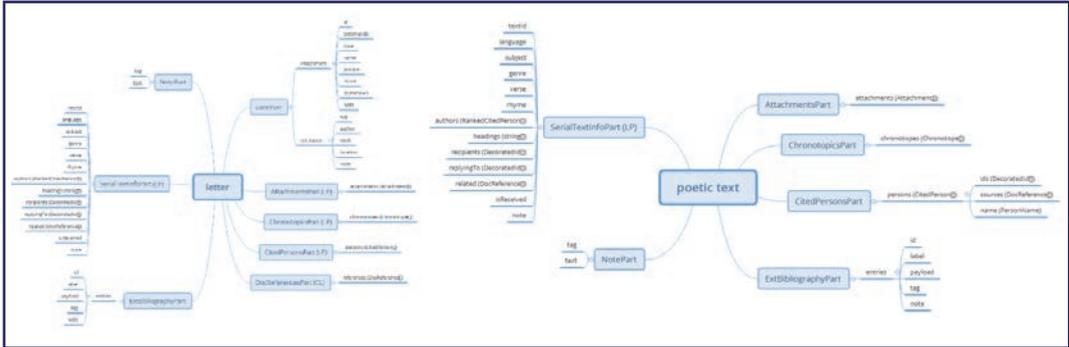


Figure 6 e 7. I modelli concettuali degli item Letter e Poetic text

general (cfr. fig. 2) ospita l'ID univoco del testo che permetterà di incrociarlo con altri *item*: per esempio con quelli relativi all'autore, al destinatario o ad eventuali testi collegati¹⁹. Campi specifici permettono di inserire altre informazioni di carattere generale: la lingua del testo (da thesaurus), un breve sommario del contenuto (stringa di testo libera), dati metrici per i testi poetici (il genere e il tipo di verso da thesauri, lo schema rimico attraverso una stringa alfabetica), l'intestazione per le lettere (stringa di testo libera). È previsto inoltre un campo note per depositare eventuali informazioni non strutturate. Ancora in *general* si raccolgono le informazioni relative all'autore attraverso un modello ripetibile che permette di dare conto anche di situazioni attributive incerte. L'attribuzione del testo ad un autore e l'identificazione dell'autore con una persona storicamente individuata sono infatti trattati come dati autonomi. In questo modo è possibile non solo immettere più ipotesi di attribuzione, ma corredarle di ipotesi di identificazione specifiche, accompagnando entrambe di opportuni riferimenti bibliografici. L'affidabilità di ogni informazione può inoltre essere classificata in modo indipendente attraverso un sistema di ranking numerico. Così sarà possibile fotografare nella sua complessità ogni situazione attributiva, ricondotta ad una struttura che garantisce la ricercabilità dei singoli dati senza ricorrere a formulazioni discorsive.

La seconda sezione, *recipients*, permette di indicare attraverso un ID il destinatario o, essendo l'elemento ripetibile, i destinatari del testo (Fig. 8). Un *rank* numerico consente di qualificare l'affidabilità del dato, accompagnato da relative fonti. La presenza di un campo *tag*, previsto nel modello generale *DecoratedId* alla base della parte, permette di distinguere efficacemente tra tipologie diverse di destinatari (per esempio certi, incerti, presunti ecc.). La sezione *reply-to* sfrutta la medesima struttura per raccogliere i dati relativi ai testi legati a quello schedato nella dinamica di scambio (fig. 9): i testi di proposta e risposta, nel caso delle tenzoni, e le lettere alle quali il testo sche-

L'item *Letter*: in attesa del modello definitivo, non sarà illustrata in questa sede.

¹⁹ È importante specificare che per ogni redazione individuata di un testo sarà compilato un *item* indipendente: le redazioni α , β , γ , e di una lettera petrarchesca, per esempio, saranno schedate come tre *item* autonomi, ciascuno dotato di un proprio ID univoco, quindi collegati attraverso uno specifico metadata.

dato risponde o di cui costituisce la risposta, ciascuno individuato da un proprio *ID* e marcato da specifici *tag* («proposta», «risposta», ecc.). La descrizione dei singoli anelli della catena crea dunque i presupposti per la ricostruzione logica dell'intero scambio: un aspetto che sarà valorizzato anche graficamente in fase di pubblicazione dei dati.

The image displays two screenshots of a web interface for editing text items. The top screenshot shows the 'recipients' section with two entries:

id	rnk	tag	src
Giovanni Barrili	0	destinatario	0
Niccolò Acciaiuoli	0	destinatario	0

The bottom screenshot shows the 'reply-to' section with two entries:

id	rnk	tag	src
l' provai già quanto la soma è grave	2	proposta	1
Si come Cerer, la dea de le biave	0	risposta	0

Figure 8 e 9. L'interfaccia di editing per le sezioni *recipients* e *reply-to* di un item *Poetic text* e di un item *Letter*

I dati cronologici sono complessivamente depositati nella *Chronotopics part*, ripetibile, basata sul modello *Chronotopics*, che arricchisce *Historical date* con due campi per l'indicazione del luogo e per l'attribuzione di un *tag*. L'etichetta aumenta sensibilmente le potenzialità descrittive del modello perché permette un inserimento plurimo e connotato. È cioè possibile immettere più date e più tipologie di data a seconda degli oggetti specifici descritti: per esempio «stesura», «invio», «ricezione» e così via. Risultano particolarmente efficaci nel caso petrarchesco, perché alle date e ai luoghi dichiarati dall'autore in calce al testo è possibile aggiungere quelli eventualmente divergenti ricostruiti dalla critica.

Le parti *Cited persons*, *Attachment* e *Pseudonyms* sono state pensate specificamente per gli obiettivi di ricerca di Itinera. *Cited Persons* consente di registrare in modo puntuale per ogni testo singoli riferimenti a persone ed enti, siano essi citazioni esplicite di terze persone oppure allusioni più o meno co-

parte: il dato, così collegato ai singoli luoghi testuali che ne costituiscono la fonte, è inoltre corredato di un *tag* che permette di assegnare un ruolo alla persona citata (es. «latore», «committente», ma anche un eventuale «richiesta di recapito» a persone diverse dal destinatario ecc.). La registrazione di questi riferimenti avviene attraverso l'inserimento della forma del nome presente a testo (*Person name*) e il suo collegamento alla o alle ipotesi di identificazione con persone ed enti storicamente individuati (*DecoratedId*), al quale possono essere associati opportuni richiami bibliografici e indicatori di affidabilità. L'aggregazione dei dati attraverso l'*ID* univoco così assegnato permette di produrre, in fase di pubblicazione, registri complessivi delle occorrenze di ciascun corrispondente all'interno dei testi schedati, eventualmente incrociabili tra loro: così anche la parte *Pseudonyms*, deputata a raccogliere le occorrenze degli pseudonimi usati da Petrarca e dai corrispondenti nei loro scambi, basata però sulla registrazione della forma normalizzata dello pseudonimo al nominativo. La parte *Attachment*, infine, raccoglie le informazioni relative ad eventuali allegati che accompagnavano il testo descritto, con lo scopo di chiarire dinamiche di circolazione di opere e scritti all'interno dell'*entourage* petrarchesco, di oggetti di interesse antiquario e, più in generale, di tutto quanto veniva scambiato in accompagnamento al mezzo epistolare. Di questi oggetti è possibile precisare il nome, la natura (a partire da un apposito thesaurus: opera o testo, manoscritto ecc.) e se si tratta di oggetti perduti o sconosciuti. Un campo *ID* facoltativo permette di identificare univocamente l'oggetto qualora sia utile prevederne un collegamento con altri *item*, mentre per gli allegati testuali, che costituiscono la maggioranza dei casi censiti, il modello permette di indicare l'eventuale porzione scambiata quando non si tratti di opere o testi completi. Uno spazio note e un campo per link a risorse esterne, come digitalizzazioni di manoscritti individuati, completano la parte *Attachment*.

4.3. Item Correspondent

L'*item Correspondent* è pensato per l'inserimento dei dati relativi ai corrispondenti di Petrarca (fig. 10). L'obiettivo è creare per ogni corrispondente un *item* completo che raccolga in modo strutturato le informazioni disponibili sulla sua biografia, vagliate ed eventualmente accresciute con nuove ricerche, e sui suoi rapporti con Petrarca e altri corrispondenti. Il corpus di partenza è costituito dai soli corrispondenti diretti, ma il modello è pensato per essere esteso in futuro alla descrizione dei corrispondenti di corrispondenti. In generale, il modello risulta funzionale alla rappresentazione di qualunque entità persona di cui si intenda fornire profili strutturati, analitici o sintetici.

L'*item* comprende allo stato attuale otto parti, che raccolgono informazioni di tipo generale sulla persona e sulle sue attività ed informazioni più specifiche direttamente legate agli obiettivi del progetto. Oltre a *part* generiche comuni agli altri

item per l’inserimento di bibliografia (*External bibliography part*), riferimenti letterari (*Document references*) e note (*Note part*), le parti dell’item sulle quali si è concentrato il lavoro di modellizzazione e lo sforzo di astrazione maggiore sono quelle che raccolgono le informazioni biografiche sul corrispondente: *Person part* e *Person events part*.

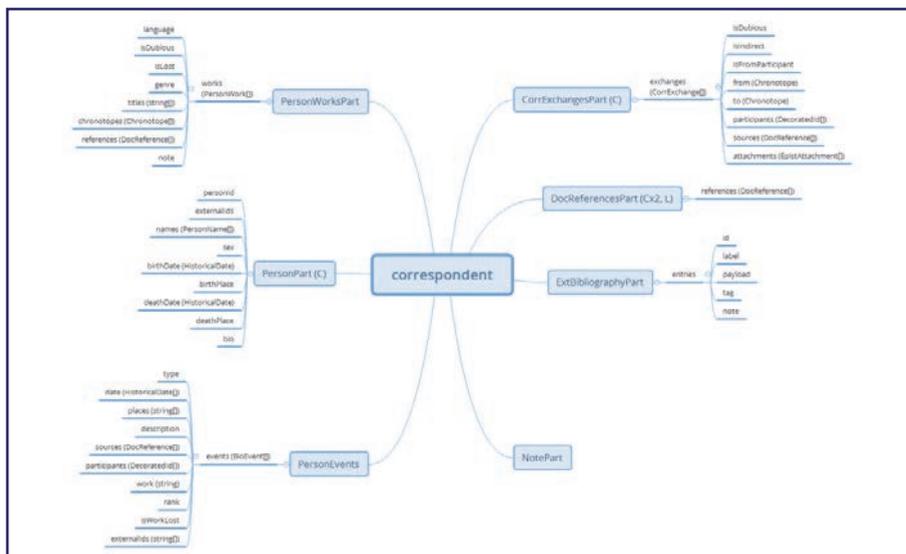


Figura 10. Modello concettuale dell’item Correspondent

La *Person part* è progettata per l’inserimento di dati biografici essenziali di qualsiasi entità-persona e comprende tre sezioni. La sezione *ID* è il luogo in cui ad ogni *item*, qui ad ogni corrispondente, viene associato un *ID human-readable* assegnato in base a convenzioni interne stabilite dall’*équipe* e opportunamente registrato in un repository consultabile dai compilatori: permetterà il collegamento e l’incrocio tra *item* interrelati, per esempio tra l’*item* di un corrispondente e quelli relativi ai testi di cui sia destinatario, autore o in cui sia citato a vario titolo, oppure con gli *item* relativi a codici di cui sia stato copista, lettore e così via. Nella stessa sezione è previsto l’inserimento del nome e di eventuali forme alternative, mentre un campo per l’inserimento di link consente il collegamento con authority file e risorse esterne – per esempio le voci del *Dizionario Biografico degli Italiani*.

Le date e i luoghi di nascita e di morte della persona sono raccolte in una sezione a parte (*Chronotopes*) attraverso il modello *Chronotopics*. Le informazioni relative a nascita e morte del corrispondente erano inizialmente comprese nella *Person events part*: si è però deciso di scorporarle e di dare loro rilievo autonomo nella *Person part* per garantire la presenza delle informazioni fondamentali relative ad una persona in un’unica parte, considerando la possibilità di un riuso del modello da parte di altri progetti eventualmente non interessati alla *Person events part*.

L'elemento *Chronotopics*, come accennato in più punti, è replicabile per ogni tipologia di data e per ogni data da inserire: a più ipotesi di datazione corrisponderanno più elementi-data opportunamente descritti da *tag*, mentre nel caso di date incerte per le quali si disponga di più elementi utili alla loro definizione (per esempio termini *ante* e *post quem*) sarà possibile inserire ogni dato singolarmente. Il sistema di ranking numerico previsto dal modello consente anche in questo caso di indicare per ogni dato il grado di attendibilità.

La terza sezione, *biography*, è costituita da un unico campo a testo libero per l'inserimento di una biografia discorsiva del corrispondente, che rispetti criteri di sintesi, chiarezza e completezza. Le singole informazioni biografiche sono raccolte in forma strutturata attraverso la *Person events part*, come si vedrà a breve, ma si è comunque ritenuto opportuno prevedere una presentazione discorsiva della biografia del corrispondente per garantire un accesso ai dati più immediato e vicino alle consuetudini della schedatura biografica tradizionale. S'intende che, mentre l'inserimento strutturato delle informazioni nella *Person events part* copre tutte quelle disponibili per la biografia di una persona, la presentazione di *biography* si concentra su aspetti ed eventi essenziali per fornirne un primo quadro di massima: *biography* costituisce insomma una chiave d'accesso più immediata per facilitare la fruizione della banca dati, da integrare necessariamente con il complesso di dati riversato nella *Person events parts*.

La *Person events part* è dunque l'indispensabile complemento della biografia sintetica inserita nella parte *Person*, che raccoglie e descrive in modo puntuale i singoli eventi biografici relativi al corrispondente (Fig. 11). L'elemento-evento, ripetibile, è descritto attraverso tre sezioni: *general*, *sources* e *participants*. *General* fornisce le

The screenshot shows a web interface titled 'events'. At the top left, there is a blue button with a plus sign and the text 'add event'. Below this is a table with four columns: 'type', 'date', and 'places'. Each row represents an event and includes a set of control icons: a pencil for editing, an up arrow for moving up, a down arrow for moving down, and a red circle with a minus sign for deleting. The events listed are:

	type	date	places
[edit] [up] [down] [delete]	Documento	22 oct 1340 AD	Sulmona
[edit] [up] [down] [delete]	Incontro	c. 15 feb 1341 – c. 15 apr 1341 AD	Napoli
[edit] [up] [down] [delete]	Nomina	16 nov 1342 AD	Napoli
[edit] [up] [down] [delete]	Viaggio	c. 1350 AD	Roma

At the bottom of the interface, there are two buttons: a red button with a white 'x' and the text 'close', and a blue button with a white checkmark and the text 'save'.

Figura 11. Prospetto riassuntivo di elementi-eventi nell'interfaccia di editing per il modello *Person events part*

informazioni generali di un evento: la data (attraverso il modello *Historical date*), il luogo, la probabilità del suo essersi verificato, utilmente indicata in caso di eventi riportati da fonti poco o non attendibili, comunque da registrare per garantire completezza al quadro biografico. La selezione da un thesaurus elaborato a partire dall'analisi del corpus e implementabile nel corso del data entry consente di indicare la tipologia di evento descritta (*type*). Se l'efficacia operativa di questo modello risiede nella possibilità di aggiornamenti progressivi attraverso l'aggiunta o la correzione di singoli elementi che non richiedono una revisione completa dell'intero *item*, quella descrittiva dipende dalla sua capacità di raccogliere e rendere ricercabili non solo eventi propriamente detti (viaggi, incontri e così via) ma anche fonti storiche relative alla persona schedata. Nella *Person events part* è infatti possibile inserire anche documenti d'archivio nei quali la persona compaia in qualità di testimone o ad altro titolo, che, se non illuminano la biografia del corrispondente in aspetti di primario interesse, forniscono importanti indicazioni sulla sua presenza in contesti individuati, permettendo di localizzarlo entro coordinate cronologiche e geografiche precise. Un campo di testo libero permette di aggiungere descrizioni puntuali, mentre uno spazio per i link consente di rimandare ad eventuali risorse esterne come documenti o manoscritti digitalizzati.

La sezione *sources* raccoglie le fonti (documentarie, bibliografiche o letterarie) relative all'evento descritto ricorrendo al modello *Document references*, mentre grazie a *DecoratedId* è possibile indicare in *Participants* persone e istituzioni coinvolte a vario titolo nell'evento descritto – l'istituzione presso cui è stata svolta un'attività, quale notaio ha vergato il documento che testimonia un acquisto da parte del corrispondente, le persone che partecipano con la persona schedata ad un viaggio e così via (fig. 12). Se l'uso di *tag* permette di indicare il ruolo dei partecipanti secondo convenzioni stabilite dall'équipe, ancora una volta l'indicazione



Figura 12. Prospetto riassuntivo di elementi-partecipanti nell'interfaccia di editing per il modello *Person events part*

dell'ID consente di incrociare ed esportare i dati in varie forme, permettendone in fase di pubblicazione una mappatura dinamica in grado di restituire incontri, viaggi, spostamenti e così via.

Sono attualmente in fase finale di progettazione le parti più direttamente legate agli obiettivi di ricerca di Itinera: *Person work part* e *Correspondent exchanges part*. Nella prima, pensata per la descrizione dell'attività letteraria dei corrispondenti, saranno schedati i testi e le opere di cui il corrispondente è autore: ciascuno dotato di un proprio ID, di essi si indicherà titolo, cronologia di composizione (*Historical Date*), genere letterario (thesaurus) ed eventuale bibliografia relativa. La parte *exchanges*, che raccoglie dati strutturati su scambi di testi, opere, manoscritti e altro tra Petrarca e i corrispondenti, punta invece ad integrare le informazioni raccolte a partire dai testi e depositate nella parte *Attachement* negli *item Letter* e *Poetic text*. Diviso in quattro sezioni (*general*, *chronotopics*, *sources*, *attachment*), il modello permette di indicare quale sia l'oggetto dello scambio; se lo scambio sia stato diretto o indiretto (avvenuto cioè durante un incontro tra le persone coinvolte, tramite invio o consegnato da altri) e se sia ipotetico o accertato; quale ne sia la direzione (da Petrarca al corrispondente o viceversa); la cronologia e i luoghi relativi; quali persone vi siano state coinvolte e con quale ruolo (*DecoratedID*); eventuali fonti e bibliografia (Fig. 13).

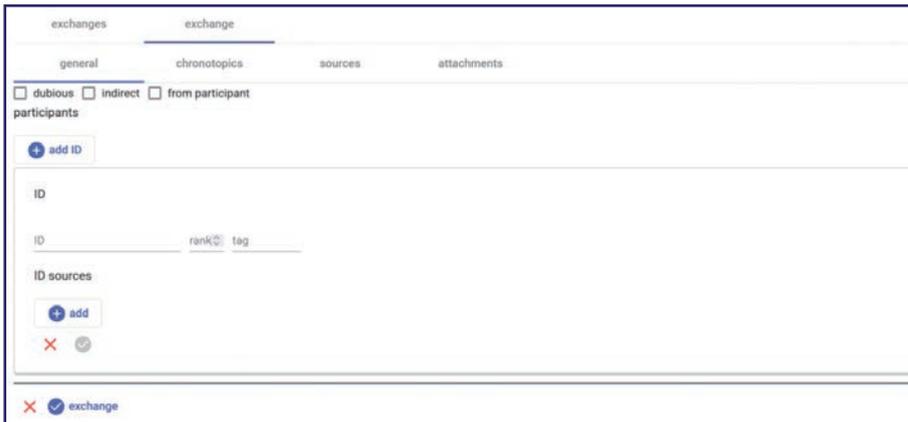


Figura 13. L'interfaccia di editing per la parte *general* di *Correspondent exchanges part*

4.4. Item Manuscript

La progettazione dell'*item Manuscript* si è dovuta confrontare con una prassi descrittiva consolidata nella tradizione catalografica²⁰, che ha conosciuto nel tempo

²⁰ All'interno della vasta bibliografia esistente a riguardo, sono stati presi a riferimento vari contributi, in particolare: Armando Petrucci, *La descrizione del manoscritto: storia, problemi, modelli*, Roma: Carocci, 2009; *Norme per la descrizione dei manoscritti*, a cura di T. De Robertis, N. Giovè Marchioli, Firenze: SISMEL - Edizioni del Galluzzo, 2021 (*Manoscritti datati d'Italia - Strumenti*; 1).

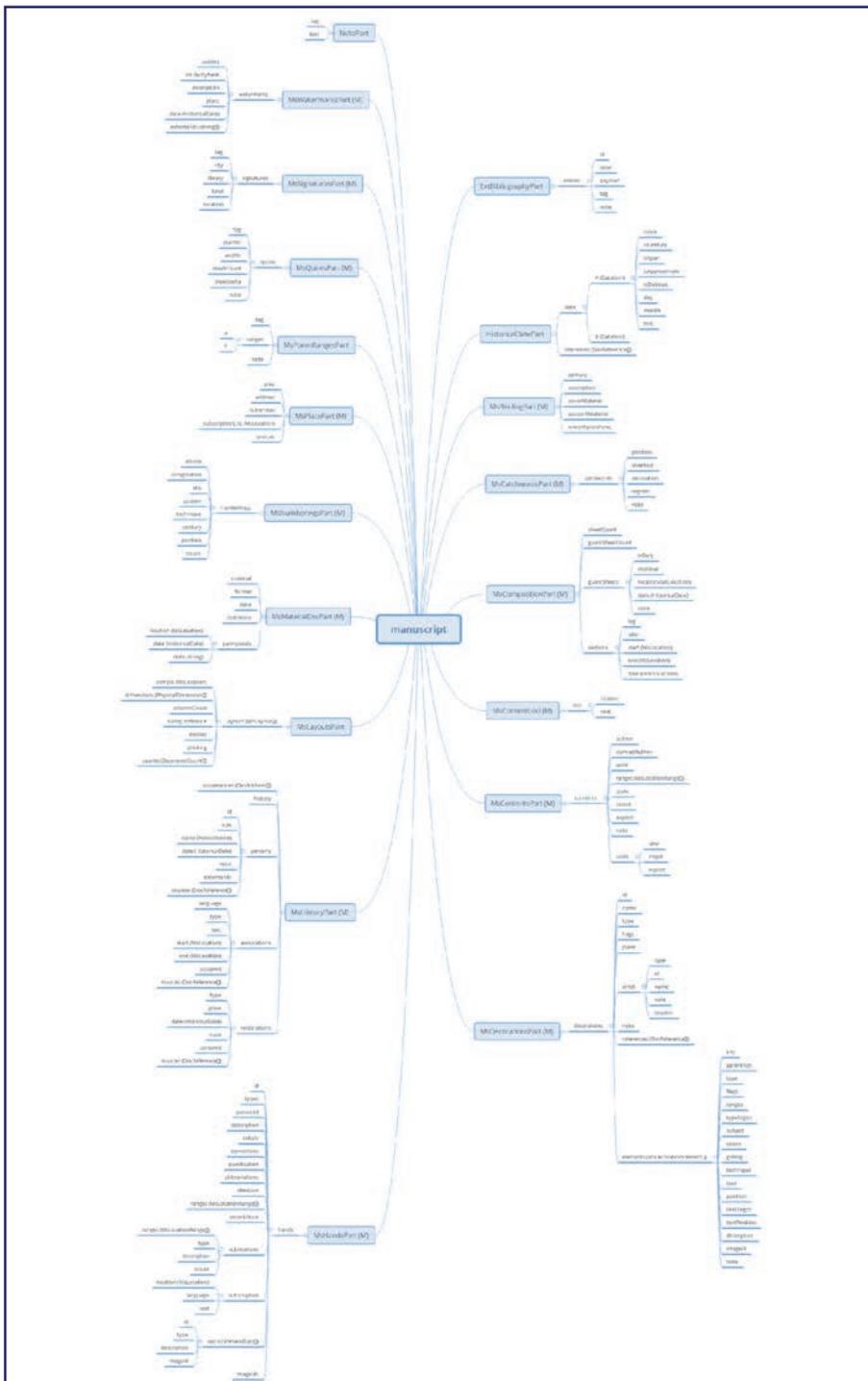


Figura 14. Il modello concettuale dell'item Manuscript

diverse realizzazioni su supporto digitale. Il passaggio dall'una all'altra modalità si è tradotto spesso nella riproduzione delle parti tradizionali di una scheda codicologica attraverso la predisposizione, per lo più, di campi a testo libero, che, prevedendo una trattazione discorsiva, permettono di indicizzare solo una minima parte delle informazioni immesse. Per la costruzione dell'*item* in Itinera, si è tentato di conciliare il rispetto delle convenzioni della descrizione codicologica tradizionale con l'obiettivo di una modellizzazione il più possibile granulare dei dati (Fig. 14).

L'*item* si compone di diciannove *part*²¹, le prime tre delle quali relative agli elementi identificativi essenziali di un codice: la segnatura attuale (*Signatures part*), il luogo di origine (*Place part*) e la datazione (*Historical date part*). Sebbene si tratti di dati di minima complessità semantica, idealmente presenti in qualsiasi tipo di scheda codicologica (indipendentemente dal grado di approfondimento o dal supporto utilizzato), esse possono risultare utili per mettere in luce con immediata evidenza quale sia la logica sottesa alla costruzione del modello di data entry per i manoscritti in Itinera.

La segnatura, per esempio, non viene immessa come un unico elemento, ma è suddivisa nelle sue componenti primarie, vale a dire: città, biblioteca, fondo (se presente), segnatura. L'identificazione del manoscritto è veicolata dalla somma di questi elementi, ma ognuno di essi risulta al contempo autonomo e può essere aggregato a tutti i dati della stessa tipologia relativi ad ogni codice descritto nel database. In questo modo sarà possibile estrarre tutti i codici di un corpus conservati in una determinata città, oltre che in una precisa biblioteca, o fornire informazioni più complesse sulla distribuzione geografica dei manoscritti. Se il procedimento di scomposizione in questo caso risulta intuitivo e trasparente perché basato su una partizione intrinseca alla struttura della segnatura, ciò non rende meno valido ed efficace il principio che ordina tale prassi, che allo stesso modo può essere applicato ad agglomerati di dati molto più complessi.

Alle prime tre parti seguono quelle deputate alla raccolta dei dati strettamente codicologici e materiali. Sono state create, a tale scopo, sette parti dedicate rispettivamente a: descrizione materiale *stricto sensu* (*Material description part*); filigrane (*Watermarks part*); composizione e consistenza (*Composition part*); cartulazione (*Numberings part*); fascicolazione (*Quires part*); sistemi di richiamo (*Catchwords part*); dimensioni della pagina e schema di scrittura (*Layouts part*). Un approfondimento sulla struttura interna di alcune di queste parti può risultare utile a esemplificare come siano state scomposte e ri-organizzate le informazioni tradizionalmente presenti in una scheda descrittiva.

Nella *part* riservata alla descrizione delle filigrane, l'immissione dei dati avviene attraverso la preliminare creazione di un nuovo elemento *watermark*, corrispondente

²¹ Sebbene le *part* siano indipendenti l'una dall'altra e non abbiano rapporti reciproci di tipo gerarchico, per il data entry sono visualizzate secondo un ordine che aspira a ricreare le logiche descrittive standardizzate in campo codicologico, così da facilitare il lavoro di compilazione.

a una singola tipologia di filigrana rilevata nel codice: a differenti filigrane del manoscritto corrisponderà dunque il medesimo numero di elementi-filigrana creati nella *part*. Ognuno di essi prevede al suo interno un determinato numero di campi che permettono di inserire, con un certo grado di parcellizzazione, le informazioni più rilevanti incluse abitualmente in una scheda codicologica. Il modello consente di indicare: il soggetto della filigrana (*subject*); un coefficiente numerico di somiglianza con eventuali esempi rintracciati nei repertori più noti (*sim. rank*) e un rinvio ad essi nelle loro versioni consultabili online (*external ID*); le coordinate topografiche e cronologiche ricavate da tali repertori (*place* e *date*); infine, un campo libero dedicato alla descrizione analitica (*description*) (Fig. 15).

The screenshot shows a web-based editing interface for a 'Watermarks' model. At the top, there are two tabs: 'watermarks' and 'watermark'. Below the tabs, there are several input fields: 'subject', 'sim.rank', 'description', 'place', 'date', and 'datation'. The 'datation' field has a red circle with a '1' and a blue download icon. Below these fields is a section for 'external IDs' with a '+ add ID' button and a list of external IDs, with the first entry being '1.' followed by a red minus icon, an up arrow, and a down arrow.

Figura 15. L'interfaccia di editing per il modello Watermarks

Un altro caso esemplificativo è quello della *Composition Part*. La consistenza di un manoscritto è convenzionalmente indicata attraverso una formula composta di vari elementi che si susseguono in un ordine fisso. Si tratta del numero totale delle carte in cifre arabe, preceduto e seguito dalla consistenza delle carte di guardia rispettivamente anteriori e posteriori (queste ultime contrassegnate da un apice) in numeri romani. Eventuali informazioni aggiuntive circa la materia delle carte di guardia, la loro provenienza e datazione, vengono fornite in modo discorsivo, più o meno sintetico e formalizzato, come nel seguente esempio:

cc. II (cart., moderne), 175, III' (I' membr., antica; II'-III' cart., moderne).

Nel database di Itinera, i dati veicolati da questa sequenza vengono scomposti e immessi attraverso due campi che accolgono rispettivamente il numero delle carte del corpo del manoscritto e il numero totale delle carte di guardia. È possibile poi aggiungere informazioni su queste ultime creando elementi-carte di guardia ripetibili, che consentono di precisare l'intervallo di carte descritte (attraverso un campo *start-end*), se esse siano anteriori o posteriori e di quale materiale siano costi-

tuite (scegliendo le relative opzioni da appositi thesauri), la loro datazione (tramite *Historical date*) (vd. fig. 3). È previsto inoltre un campo libero per annotazioni aggiuntive che non possano trovare posto altrove. Consentendo di creare tanti elementi-carte di guardia quanti siano necessari per ogni caso specifico, il database permette quindi di descrivere sottogruppi di carte di guardia omogenee (per tipologia, materiale, epoca) o potenzialmente ogni singola carta di guardia, qualora ciò si rendesse utile o opportuno. Riprendendo l'esempio di consistenza illustrato sopra, nel database verranno registrati separatamente tre gruppi di carte di guardia, e precisamente: le due carte anteriori; la prima carta posteriore; le ultime due posteriori (Fig. 16).

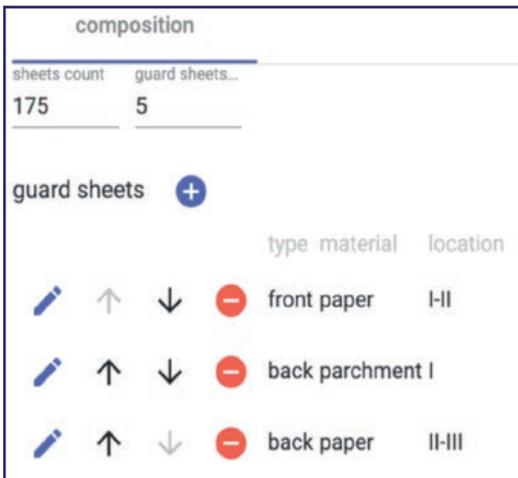


Figura 16. Prospetto riassuntivo di elementi-carte di guardia nell'interfaccia di editing per il modello *Composition*

Un'altra *part* utile ad approfondire il rapporto tra gli standard di descrizione codicologica e le soluzioni adottate in Itinera, è quella dedicata alla fascicolazione (*Quires part*). Oltre a prevedere una sezione in cui segnalare, attraverso un thesaurus, la tipologia di fascicolo prevalente nel manoscritto (o più di una, in caso di parità numerica), la parte è dotata di un campo apposito in cui registrare la fascicolazione attraverso una formula. A questo scopo si è scelto di non stabilire una sintassi descrittiva ex novo, ma di

adottarne una in uso nella prassi codicologica²². La formula viene inserita come stringa di testo secondo tale sintassi predefinita, e affinché i valori immessi siano accettati dall'editor è necessario che ne siano rispettate le regole di composizione. La sequenza produce poi un elenco con la successione dei singoli fascicoli – o degli intervalli omogenei di fascicoli consecutivi – ognuno corredato del numero delle carte che lo compongono, di eventuali carte aggiunte o mancanti, di uno specifico campo note in cui dare conto di anomalie o di altri aspetti rilevanti (Fig. 17).

²² Si tratta, in particolare, della formula codificata da Montague R. James che prevede di segnalare con cifre progressive i fascicoli, dotando ognuno di essi (o gruppi di sequenze omogenee) di un esponente che indichi il numero delle carte che lo compongono (cfr. Maria Luisa Agati, *Il libro manoscritto da Oriente a Occidente: per una codicologia comparata*, Roma: L'Erma di Bretschneider, 2009, p. 167). Tale formula, detta anche *per collazione*, è adottata dai *Manoscritti datati d'Italia* già citati. Rispetto a questi, Itinera (come spesso altri cataloghi) ricorre alla soluzione di segnalare all'interno della formula anche le carte singole aggiunte o mancanti, marcate con + e -.

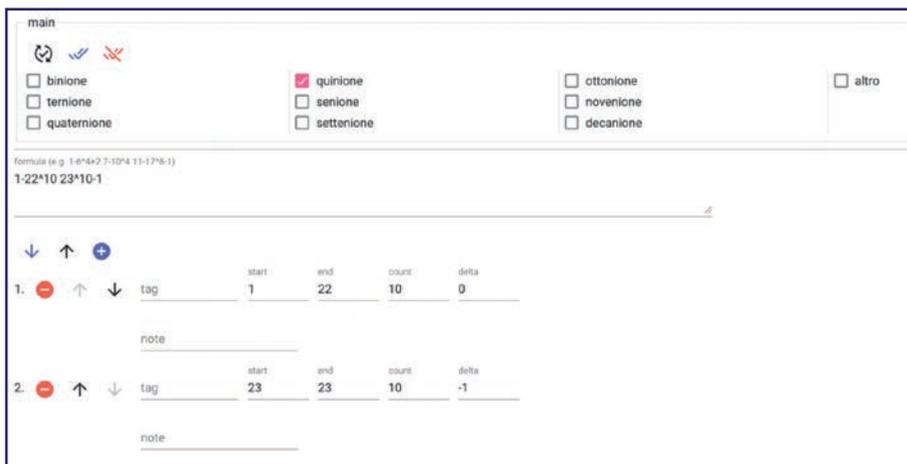


Figura 17. Il modello di *Quires part*

Anche la parte *Layouts*, riservata alla *mise en page*, trova in una formula la sua espressione più complessa. È possibile registrare singolarmente tipologie di impaginazione differenti creando più elementi *layout*. Il modello consente di indicare, in appositi campi distinti, la carta di riferimento sulla quale sono state effettuate le rilevazioni, il numero delle colonne, la tecnica di rigatura (thesaurus), il rimando al modello di *mise en page* riscontrato sul repertorio di Derolez²³, la tecnica di foratura. Le dimensioni esatte della pagina e dello schema di scrittura vengono fornite attraverso una formula che si basa su una sintassi piuttosto articolata, elaborata anche in questo caso a partire dai metodi più diffusi in campo codicologico e perfezionata per restituire un'immagine il più possibile fedele delle molte varianti che influenzano la *mise en page*. Tale formula accoglie le misure dei margini, delle colonne di scrittura, dell'intercolumnio, dei colonnini laterali che dovessero eventualmente accompagnare le colonne (a destra e a sinistra di ognuna di esse), delle righe di testa e di piede (per questi ultimi elementi, la formula permette inoltre di distinguere se gli spazi contengano o meno scrittura). Anche in questo caso, come in *Quires part*, la formula è sottoposta istantaneamente a verifica e non può essere salvata se contiene errori di composizione e non rispetta la sintassi prestabilita. Non si tratta, inoltre, di una stringa inerte: viene infatti tradotta nell'elenco di tutti i singoli valori in essa contenuti, identificati attraverso una precisa etichetta generata automaticamente (fig. 18 e 19). Questo permette di verificarne con più efficacia l'esattezza, di estrarre più rapidamente i singoli dati, di intervenire puntualmente nel caso fossero necessarie correzioni: è possibile infatti modificare un valore dell'elenco e assumere tale variazione direttamente nella formula. La *part* è dotata anche di un sistema di visualizzazione che traduce la formula in uno sche-

²³ Albert Derolez, *Codicologie des manuscrits en écriture humanistique sur parchemin*, v. 2, Turnhout: Brepols, 1984.

ma grafico costruito secondo proporzioni definite dai valori in essa contenuti. La parte permette inoltre l'inserimento di ulteriori dati numerici (ognuno da dotare di apposita etichetta e munito di un campo libero per eventuali annotazioni), sfruttabili, ad esempio, per registrare il numero di righe e linee presenti sulla pagina. È previsto, infine, un campo libero che possa accogliere annotazioni ed eventuali incongruenze di varia natura.



+ add dimension				
	ID	value	unit	
1.	height	205	mm	
2.	width	140	mm	
3.	margin-top	22	mm	
4.	area-height	134	mm	
5.	margin-bottor	49	mm	
6.	margin-left	22	mm	
7.	col-1-width	86	mm	
8.	margin-right	32	mm	

Figure 18 e 19. Formula della mise en page e sua traduzione nei valori corrispondenti nell'interfaccia di editing per il modello Layout

Alle sette *part* relative agli aspetti materiali, ne seguono due che corrispondono ad altrettante sezioni tradizionalmente presenti in una scheda di manoscritto, una dedicata alla descrizione di mani e tipologie grafiche (*Hands part*), l'altra alla descrizione di eventuali apparati decorativi (*Decorations part*), entrambe molto articolate.

La descrizione delle scritture di un manoscritto avviene di norma seguendo una prassi consolidata, che prevede una certa organizzazione interna delle informazioni. Le catalogazioni digitali sono riuscite di rado a superare un'impostazione esclusivamente discorsiva

e a strutturare i dati in modo diverso: è stato questo l'obiettivo del modello di Itinera, che affianca una parcellizzazione maggiore delle informazioni alla modalità tradizionale, comunque imprescindibile per una presentazione esaustiva della scrittura. In *Hands part* viene creato un nuovo elemento per ogni mano che si voglia descrivere, e per ciascuna è possibile indicare una o più tipologie grafiche adottate dal copista (thesaurus) e precisare l'intervallo di carte interessate (campo *start-end*). Benché sia previsto un campo libero per la descrizione discorsiva della scrittura, le osservazioni circa alcuni aspetti specifici (iniziali al tratto, correzioni, interpunzione, abbreviazioni) possono essere scorporate e trovare posto in sezioni apposite. Tale suddivisione permette di estrapolare con maggiore facilità le informazioni riguardanti questi singoli aspetti, facilitando il confronto tra abitudini scrittorie di mani diverse, tra soluzioni adottate dal medesimo copista in codici dif-

ferenti, e così via. Un'ulteriore sezione (*signs*) incrementa sensibilmente queste potenzialità, consentendo di isolare e descrivere separatamente singoli segni (lettere o segni interpuntivi), associando ad ognuno un *ID* univoco e corredandoli eventualmente della relativa immagine. La *part* include infine due sezioni specifiche che raccolgono in modo strutturato i dati relativi a rubriche (*rubrications*) e a eventuali sottoscrizioni del copista (*subscription*).

Decoration part è progettata per descrivere separatamente differenti campagne decorative presenti nel manoscritto, attraverso la creazione di più elementi-decorazione. Oltre a inserire informazioni di carattere generale sulla decorazione nel suo complesso (data, luogo, paternità ecc.), è possibile descrivere singolarmente le diverse tipologie di elementi decorativi di cui è costituita, ciascuno replicabile secondo necessità: pagina incipitaria, pagine decorate, illustrazioni, ornamentazioni, iniziali semplici, iniziali filigranate, iniziali ornate, iniziali con figura, elementi paragrafematici (di mano del decoratore), spazi bianchi riservati, immagini estemporanee. Una volta selezionato l'elemento compaiono, a seconda della tipologia, campi specifici che permettono di registrarne singole caratteristiche (tecnica, strumento, posizione sulla pagina, colori ecc.), oltre ad un campo di testo libero dedicato alla descrizione discorsiva e uno per il rinvio alle relative immagini. Una struttura così articolata consente di scegliere, in base alle esigenze di ricerca e alla natura propria del singolo codice, il gradiente di specificità delle informazioni da immettere, senza rinunciare, anche nei casi in cui si voglia restituire un'immagine complessiva della decorazione, all'indicizzazione dei dati.

Seguono tre parti progettate per l'inserimento di informazioni riguardanti i testi tramandati in un manoscritto: mentre la prima (*Contents part*) serve a descrivere il contenuto secondo modalità vicine alle schede tradizionali (autore, titolo dell'opera, intervallo di carte, incipit, explicit, stato del testo ecc.), le altre due sono pensate specificamente per il corpus di codici di Itinera. *Poem ranges part* (vd. fig. 1) è stata ideata, infatti, per registrare l'ordine dei componimenti poetici dei *Rvf* trascritti in ogni testimone, mentre *Content loci part* è strutturata in modo da consentire l'inserimento di un certo numero di loci critici preventivamente individuati. Entrambe le parti hanno l'obiettivo di offrire strumenti utili a successive ricerche sulla tradizione del *Canzoniere* e sulle forme di seriazione che hanno caratterizzato la diffusione e la circolazione dei suoi componimenti, in ottica sincronica e diacronica. Nonostante risponda a questo preciso scopo, la *part* relativa ai loci critici può, evidentemente, trovare proficua applicazione in qualunque progetto di schedatura codicologica che abbia anche finalità ecdotiche. L'*item Manuscript* è provvisto, infine, di due parti comuni a tutti gli *item* di Itinera, vale a dire *External bibliography part* e *Note part*.

Il modello e i sottomodelli dell'*item Manuscript* sono stati ideati per rispondere alle finalità specifiche del progetto Itinera, nel tentativo costante, al contempo, di

produrre strutture descrittive fruibili e funzionali per qualsiasi progetto che abbia tra i suoi obiettivi l'analisi e lo studio di corpora di manoscritti²⁴, anche e soprattutto in ottica comparativa²⁵.

5. Obiettivi e prospettive future

La banca dati, in fase avanzata di elaborazione, risponde agli obiettivi fondamentali e prioritari di Itinera, di cui costituisce l'imprescindibile strumento di lavoro. Entro il termine del progetto, è previsto il completamento della schedatura dell'intero corpus di testi secondo i modelli predisposti per gli *item* relativi. Anche per corrispondenti e manoscritti si intende portare a termine la schedatura digitale, ma per garantire coerenza e completezza al lavoro si è deciso di adottare una doppia modalità di descrizione, sintetica e analitica. Se i modelli descrittivi puntano infatti ad un massimo di esaustività, la mole di informazioni e la fattibilità del progetto hanno suggerito una compilazione differenziata. La schedatura sintetica permette infatti di produrre una base di dati estesa ai corpora nella loro interezza, tenuto conto anche della possibilità di implementarla nel tempo, preferibile ad una schedatura solo analitica ma limitata ad una parte soltanto degli oggetti interessati dalla ricerca.

Per quanto riguarda il corpus dei manoscritti, si è riservata una schedatura analitica ai testimoni più antichi, compresi entro il primo decennio del XV secolo, mentre si è deciso per il momento di produrre schede sintetiche per gli altri. Per la schedatura dei corrispondenti è stato necessario considerare, insieme alle difficoltà poste dall'estensione del corpus, la distribuzione molto disomogenea dei dati, che vanno da un massimo di copertura documentaria e bibliografica ad un totale silenzio delle fonti conosciute. Per questo si è definito come obiettivo prioritario quello di raccogliere per la prima volta in un unico punto di accesso tutto il materiale disponibile, riservando ad un selezionato numero di corrispondenti analisi più approfondite e supplementi di indagine finalizzati alla produzione di schede analitiche. La progettazione ha comunque costantemente tenuto conto delle future possibilità di implementazione della banca dati (per esempio attraverso l'inserimento di edizioni digitali) e di sviluppo di output ulteriori (come mappe interattive e timeline), cercando di garantire le premesse necessarie ad entrambi.

Per il lavoro di schedatura, ripartito tra le varie unità di ricerca secondo le rispettive competenze specifiche, sono previste due fasi operative, una prima di compila-

²⁴ Alcuni modelli per la descrizione codicologica elaborati per gli *item* di Itinera, in particolare *Signatures part*, *History part*, *Content part*, sono per esempio stati adottati dai progetti PURA e The Transmission of Ancient Linguistics (cfr. n. 5).

²⁵ Le potenzialità di questo modello in termini di aggregazione ed estrazione di dati quantitativi utili a ricerche successive potranno inoltre essere valorizzate anche graficamente attraverso lo sviluppo di output specifici.

zione e una seconda di verifica e validazione, necessaria ad assicurare la correttezza e la qualità dei dati inseriti. Ruoli e responsabilità di compilatori e revisori, che corrispondono a diversi livelli di autorizzazione nel sistema, sono stati preventivamente stabiliti all'interno del gruppo di lavoro anche sulla base dei profili dei membri dell'équipe.

Per raggiungere gli obiettivi di Itinera, è fondamentale procedere ad una riflessione strutturata sulle modalità di interrogazione dei dati raccolti, che permetta di sfruttare al meglio le possibilità di ricerca e di aggregazione offerte da Cadmus²⁶: sarà questa la prossima fase del lavoro, funzionale allo sviluppo del front-end e di un sito web nel quale confluiranno anche le linee guida e i manuali approntati in fase di progettazione ad uso dei compilatori.

In corso d'opera, il progetto si è arricchito di obiettivi ulteriori rispetto a quelli inizialmente prefissati. In particolare, lo sviluppo del database bibliografico ha fornito lo spunto per uno spin-off progettuale che consiste nella creazione della prima bibliografia digitale petrarchesca. Le voci bibliografiche inserite per la compilazione degli *item* coprono infatti una porzione significativa della bibliografia petrarchesca: con un dispiego limitato di risorse, questo bacino bibliografico può facilmente essere ampliato e garantire, se periodicamente aggiornato, un quadro completo della produzione scientifica di interesse petrarchesco da mettere a disposizione degli studiosi online e in open access.

Sono attualmente in fase di discussione accordi di collaborazione e integrazione con altri progetti e terze parti per assicurare durabilità alla banca dati al di là dei tre anni di finanziamento del progetto. L'ambizione di Itinera è infatti di porre le basi per l'inclusione progressiva nel database di altri corpora, anche attraverso collegamenti strategici con progetti affini.

*The article presents the elaboration phases and the models of the PRIN Itinera project database (Italian Trecento Intellectual Network and European Renaissance Advent). Developed through the Cadmus framework, the database will collect research on Petrarch's correspondents and their texts, as well as codicological and paleographic data related to the manuscript tradition of the *Rerum vulgarium fragmenta*.*

²⁶ Il database e la modellazione utilizzati per la creazione dei contenuti sono distinti dalle innumerevoli modalità con cui essi potranno essere pubblicati. Si tratta infatti di temperare due diverse esigenze: da un lato la facilità e la versatilità di modelli ed editor durante l'inserimento dei dati, per garantire un ambiente comprensibile e familiare rispetto alla tradizione degli studi; dall'altro, la necessità di riversare questi dati in formati e modelli standard per la loro pubblicazione, producendo ad esempio documenti TEI, esportazioni di triple RDF secondo ontologie standard, endpoint API o SPARQL ecc. La pubblicazione dei dati si avvale quindi di una serie di strumenti di esportazione capaci di produrre questi ed altri tipi di output in modo automatico (inclusa un'applicazione web per l'utente finale nutrita da un tradizionale RDBMS), a partire dai dati liberamente e dinamicamente modellati e inseriti nella fase di creazione.

L'ultima consultazione dei siti web è avvenuta a dicembre 2021.