



UNIVERSITÀ DI PISA

Corso di Laurea in Informatica Umanistica

RELAZIONE

**Archivio digitale: raccolta di cartoline della  
Grande Guerra**

**Candidato:** *Dominga Elia*

**Relatore:** *Enrica Salvatori*

**Correlatore:** *Angelo Mario Del Grosso*

Anno Accademico 2020-2021

# INDICE

Introduzione	2
1. Prima guerra mondiale	3
1.1 La nascita delle cartoline	3
1.1.1 Corrispondenze dal fronte	4
1.1.2 Propaganda e censure	5
2. Cartoline Donazione Borghi	7
2.1 Museo Civico Etnografico “G. Podenzana”	7
2.2 Public history	8
2.3 Corpus di cartoline	10
2.3.1 Retro cartolina	10
2.3.2 Fronte cartoline	14
3. Sviluppo del progetto	18
3.1 Prima fase: trascrizione XML	18
3.2 Seconda fase: XSLT e corpus XML-TEI	21
3.3 Terza fase: SaxonJS 2	26
4. Risultato finale: Archivio digitale	35
5. Conclusioni	40
Bibliografia	42
Sitografia	43

# Introduzione

Il progetto di tesi di laurea si articola su una raccolta di cartoline, scritte durante la Prima Guerra Mondiale, conservate presso il Museo Etnografico “Giovanni Podenzana” della Spezia, a seguito della donazione di Rossana Borghi. Tale progetto è frutto della collaborazione, oltre che del Museo Etnografico “G. Podenzana”, anche del Laboratorio di Cultura Digitale dell’Università di Pisa, dell’Istituto di Linguistica Computazionale del CNR e della Società Storica Spezzina.

Il lavoro illustra l’importanza che le cartoline hanno avuto durante la Grande Guerra. Esse sono state un enorme strumento di comunicazione, ma allo stesso tempo oggetto di propaganda. Le immagini, che trasmettono una legittimazione della guerra e delle sue sacre motivazioni, appaiono in contraddizione con i messaggi che i combattenti inviavano, pur sotto censura, ai loro familiari. Le cartoline sono un elemento importante anche per raccontare la storia e i vari aspetti che l’hanno caratterizzata. Grazie all’avanzamento tecnologico e al web, è possibile adottare le buone pratiche indicate dalla *digital public history*, cioè la diffusione di contenuti storici pensati per un pubblico non specializzato attraverso mappe storiche interattive, visualizzazioni di dati, timelines, e siti internet di divulgazione come nel caso dei materiali oggetto del presente contributo.

Contestualmente verrà discusso il lungo processo che ha portato alla realizzazione di un archivio digitale, che raccoglie il corpus di cartoline per presentarlo ad un ampio pubblico. Lo sviluppo dell’archivio ha fatto uso di varie tecnologie, per lo sviluppo delle varie fasi che hanno portato a questo risultato. Dalla trascrizione e codifica delle cartoline, alla realizzazione del sito web in tutte le sue funzionalità.

Obiettivo principale è stato quello di creare un sito web, accessibile e fruibile da tutti, in modo che chiunque possa apprendere, attraverso l’oggetto d’archivio “cartolina”, la storia della guerra avvenuta negli anni 1914-1918, e al tempo stesso vivere le emozioni, le paure, le speranze dei soldati e dei loro familiari.

# 1. Prima guerra mondiale

La prima guerra mondiale, chiamata anche la grande guerra, fu il più grande conflitto armato mai combattuto fino a quel tempo. Lo scontro ebbe inizio il 28 luglio 1914 quando a Sarajevo, per mano dello studente bosniaco Gavrilo Princip, fu assassinato l'arciduca Francesco Ferdinando. Questo portò ad una dichiarazione di guerra, da parte dell'impero austro-ungarico, al Regno di Serbia. Nel decennio precedente la Gran Bretagna aveva perso il primato di potenza, per cui la Germania e gli USA minavano al primato inglese. Di conseguenza l'Inghilterra, nel 1907, si alleò con la Francia e la Russia dando vita alla triplice intesa. Lo schieramento opposto invece era formato dagli imperi centrali: Germania e Austria. L'Italia inizialmente rimase neutrale ma, successivi eventi, portarono l'Italia a prendere una decisione: il 26 aprile 1915 firmò segretamente a Londra un patto che la impegnava a schierarsi con le forze dell'intesa. Determinante per l'esito finale fu l'ingresso degli Stati Uniti d'America nel 1917.

La guerra si concluse definitivamente l'11 novembre 1918 con la firma dell'armistizio da parte della Germania. Ma la grande guerra aveva seminato morte e devastazioni in tutta Europa, e gli accordi di pace, mal gestiti, prepararono il terreno a un nuovo conflitto. La pace del 1918, i trattati e le promesse furono solo una tregua nel corso di uno scontro che sarebbe durato fino alla fine della seconda guerra mondiale.

## 1.1 La nascita delle cartoline

Le cartoline sono un oggetto postale complesso, l'unico oggetto bidimensionale che presenta due lati, fronte e retro, ma tre partizioni: la prima per le informazioni riguardanti il destinatario e i segni postali, la seconda per il messaggio e la terza sezione per l'immagine.

La prima cartolina in assoluto risale al 1 ottobre del 1869, la famosa Correspondenz-Karte, emessa dalla Posta austro-ungarica. Successivamente, l'uso di tale strumento di comunicazione, venne assunto anche da altri stati in meno di tre anni. Nel 1897 il V Congresso dell'Unione Postale Universale stabilì la validità internazionale della cartolina di produzione privata, questo diede inizio ad una vera e propria mania, soprattutto nei paesi anglosassoni, cioè la cartofilia, o deltiologia, vale a dire il collezionismo di cartoline, si diffuse addirittura l'abitudine di conservare le cartoline in album, come se fossero delle fotografie.

Il fenomeno delle cartoline di inizio Novecento, fu un fenomeno internazionale legato all'ascesa della borghesia. Furono principalmente le classi medio-alte della borghesia a rivolgersi ad essa, soprattutto per uso pratico, in sostituzione della lettera, per chi voleva inviare semplicemente un messaggio stringato, o per chi non era portato a scrivere belle e lunghe lettere. Le classi più alte invece snobbarono le cartoline perché troppo economiche per loro, e troppo volgari in quanto il messaggio era scoperto e visibile agli occhi di tutti. Anche le classi più basse ignorarono questo fenomeno per i motivi opposti, ovvero troppo costose per loro, ed anche inutili per coloro che non avevano potuto permettersi un'istruzione, e per questo non sapevano né leggere né scrivere.

### 1.1.1 Corrispondenze dal fronte

La prima guerra mondiale è stato uno dei conflitti più sanguinosi dell'umanità. Nei quattro anni e sei mesi di ostilità, considerando tutte le nazioni del mondo, si stimano circa 9.722.000 soldati deceduti con oltre 21 milioni di feriti, 650.000 morti solo in Italia. I civili non furono risparmiati, morirono circa 950.000 a causa delle varie operazioni militari, e tanti altri persero la vita a causa delle conseguenze e degli effetti collaterali che questo grande conflitto provocò, come per esempio la carestia, la carenza di generi alimentari, malattie ed epidemie.

In questo periodo di guerra e di tragedia, le cartoline assunsero un ruolo importante, e lo sono tutt'oggi. Grazie ad esse abbiamo delle testimonianze che delineano gli aspetti peculiari della guerra, sono fonti necessarie per ripercorrere la storia, ma soprattutto i comportamenti che questi uomini ebbero nei confronti della guerra, “le cartoline sono utilizzate dai soldati per comunicare sentimenti, paure, speranze” (Nuvoli, 2004)<sup>1</sup>. Esse raccontano il dolore dei soldati protagonisti, ma anche e soprattutto dei familiari che nutrivano speranza, ma allo stesso tempo paura per i loro cari.

La prima guerra mondiale ha rappresentato un eccezionale laboratorio di pratica di scrittura per milioni di soldati scarsamente alfabetizzati. Questo era avvenuto già in precedenza, nei primi decenni del XIX secolo, a causa delle diverse situazioni regionali che avevano portato a fenomeni migratori rilevanti. Tale fenomeno aveva favorito l'ingresso nell'universo della scrittura di milioni di uomini e donne comuni, che per poter comunicare con i loro cari, a causa della lontananza, furono costretti a prendere per la prima volta in mano una matita e a scrivere. La Grande Guerra ha dato una forte accelerata a questo fenomeno, la frequenza con la quale i soldati al fronte scrivevano a

---

<sup>1</sup> Tomasoni, Nuvoli. 2004. *La Grande Guerra raccontata dalle cartoline*.

casa dimostrava la necessità che questi ultimi avevano di comunicare, di scrivere e soprattutto di ricevere posta. I soldati avevano trasformato la scrittura in uno strumento di sopravvivenza.<sup>2</sup>

Scrivere era un elemento importante per le loro vite, da una parte per documentare e testimoniare la propria esperienza di vita in quegli anni, ma soprattutto per assicurare i loro cari e allontanarsi loro stessi, anche solo per un attimo, dagli orrori della guerra, e trovare un rifugio nei loro familiari e nelle loro case. Sul fronte italiano si stima un recapito di quasi quattro miliardi di lettere e cartoline postali, in tutti nacque un forte desiderio di scrittura, dai soldati alle famiglie, qualsiasi fosse il loro livello culturale. Relativo proprio a questo, i dati italiani sono ancora più sorprendenti perché, nel 1911, si contava un tasso di analfabetismo pari a 37.6%, questo vuol dire che oltre due milioni di soldati degli oltre cinque milioni mobilitati non possedevano le competenze alfabetiche opportune al momento dell'inizio della guerra.

### 1.1.2 Propaganda e censure

Come in altre occasioni, anche durante la prima guerra mondiale, si cercò di sfruttare a pieno i mezzi di comunicazione per divulgare messaggi che andassero ad influenzare l'opinione pubblica. Solitamente, venivano utilizzate le locandine dei quotidiani o i manifesti, mentre, la novità del primo conflitto, sta proprio nell'utilizzo delle cartoline postali. Venne scelto lo strumento delle cartoline perché era un mezzo di comunicazione rapido ed economico, fruibile da chiunque, inoltre con le immagini era possibile comunicare anche a coloro che ancora facevano parte di quella parte di popolazione analfabeta.

Le cartoline postali presentavano, insieme ai saluti e agli abbracci ai familiari, messaggi propagandistici più o meno espliciti. L'elemento con maggiore rilevanza erano le illustrazioni, cioè raffigurazioni, realizzate da disegnatori esperti delle redazioni di periodici umoristici, che segnavano le tappe più importanti delle battaglie, o esortavano la donna ad un maggiore impegno all'interno della famiglia e della società civile. Molto spesso alcune rappresentazioni demonizzavano il nemico, lo ritraevano come essere ripugnante, quasi animalesco, mettendolo in ridicolo. L'idea era quella di alimentare, sempre più, il senso di avversione nei confronti dell'esercito nemico.

---

<sup>2</sup> Croci Federico. Genova 1992. *Scrivere per non morire: lettere dalla grande guerra del soldato bresciano Francesco Ferrari*.

Insieme alla propaganda si ha anche la censura da parte di tutte le autorità civili, militari e perfino religiose, per giustificare di fronte all'opinione pubblica e rendere accettabili ai combattenti scelte di ordine politico, economico, sociale e militare eticamente discutibili.<sup>3</sup> Già dal 23 maggio 1915, con Regio Decreto, venne istituita la censura postale. La normativa prevedeva diversi divieti, tra cui il divieto di inviare cartoline illustrate con paesaggi o panorami di città, era vietato usare sistemi criptati di comunicazione o la stenografia. Le buste venivano bollate con il numero del censore, ispezionate internamente, successivamente richiuse con il nastro gommato prestampato "VERIFICATO PER CENSURA". Anche i prigionieri di guerra subivano prima una censura sul suolo nemico, poi le cartoline erano prese in carico dalla Croce Rossa Internazionale, attraverso alcuni punti di frontiera con la Svizzera, e fatti pervenire all'ufficio preposto presso il Ministero delle Poste per una ulteriore censura italiana. Solo al termine di questi controlli, le cartoline erano immesse nel circuito civile per la distribuzione.

Così, nel 1914, in un clima di consenso sociale che toccava tutti i paesi coinvolti nel conflitto, gli organi di controllo e di censura, alimentati da leggi straordinarie che conferivano ai governi pieni poteri, aumentarono notevolmente il loro numero e il loro potere.<sup>4</sup> In Italia la censura fu piuttosto severa, soprattutto perché mancava il largo consenso alla guerra, sia da un punto di vista politico che sociale, quindi i governi introdussero misure repressive per contenere il malcontento del paese.

---

<sup>3</sup> <sup>4</sup> J. M. Winter, *Il mondo in guerra - Prima guerra mondiale*, Milano, Selezione dal Reader's Digest, 1996

4

## 2. Cartoline Donazione Borghi

Il corpus di cartoline, oggetto dell'archivio digitale realizzato per questo elaborato di tesi, è stato donato da Rossana Borghi al Museo Civico Etnografico "G. Podenzana" di La Spezia. È costituito esattamente da 227 cartoline, scritte durante il periodo della Prima Guerra Mondiale, tra il 1914 e il 1918.

### 2.1 Museo Civico Etnografico "G. Podenzana"

Il Museo Civico Etnografico è situato a La Spezia, in Liguria, ed è stato fondato nel 1873, per poi essere inaugurato ufficialmente nel 1906, e rappresenta una delle sezioni più antiche del Museo Civico della Spezia. Oggi è intitolato a Giovanni Podenzana, nato a La Spezia nel 1864 e morto a Brugnato nel 1943, era un musicista, etnografo, naturalista e primo Conservatore dei Musei Civici della città. In oltre trent'anni di viaggi e ricerca sul campo, Giovanni Podenzana raccolse migliaia di oggetti riferibili all'antropologia extraeuropea, alla cultura e alle tradizioni popolari della comunità della Lunigiana. Questi elementi offrono al visitatore un'interessante sintesi della vita in Lunigiana nel Settecento e nel primo Novecento.

La storia di questo museo è abbastanza tormentata, in particolare per quanto riguarda la localizzazione. Infatti il museo ha avuto diverse sedi: la prima era costituita da due sale al pianterreno del vecchio Teatro Civico, nel regolamento del 1887, scopo del museo era quello di raccogliere e conservare reperti naturalistici, archeologici e delle memorie storiche della città. Nel 1875 il museo fu spostato nei locali dell'ufficio d'Arte in corso Cavour e nel 1884 fu assunto Giovanni Podenzana, prima come preparatore tassidermista, poi come conservatore ed infine come direttore tecnico delle raccolte di storia naturale. Nel 1884 la direzione fu affidata a Davide Carazzi, fisiologo e naturalista il quale si adoperò affinché il museo ottenesse una sede migliore, infatti fu spostato al piano terra delle scuole sull'attuale via Don Minzoni. Il successore divenne Ubaldo Mazzini, sotto la sua guida il museo accrebbe notevolmente soprattutto a livello internazionale, ed è proprio in questo periodo che si formò la raccolta etnografica per opera di Podenzana. Nel 1905 ci fu ancora una volta un cambio di posizione, il museo fu collocato nella Palazzina Crozza in corso Cavour, e rimase lì fino agli anni Trenta. Con lo scoppio della seconda guerra mondiale, era stata aperta una nuova sede nell'ex convento delle Clarisse, ma quest'ultima fu distrutta dai bombardamenti del 1943 e non fu mai più ricostruita.



Negli anni Novanta e i primi anni duemila, con la direzione di Marzia Ratti, ci fu una riqualifica museale della città e parte delle collezioni storiche vennero distribuite al Museo Archeologico presso il Castello di San Giorgio e nel Museo Etnografico presso l'ex Oratorio di San Bernardino.

Ad oggi il museo si trova in via del Prione 156, nel centro storico cittadino. Ha sede nell'antico oratorio di San Bernardino, edificio del XV secolo nei pressi della porta di Genova, precedentemente sede della Pubblica Assistenza cittadina.

## 2.2 Public history

Le cartoline sono uno strumento per raccontare la storia. In questo caso, attraverso la donazione delle cartoline di Borghi, siamo stati in grado di rivivere la sofferenza e la tragedia che ha colpito parte del mondo con il primo conflitto mondiale. In questo modo tutti siamo in grado di fare storia, tutti possiamo condividere esperienze di vita, testimonianze, documenti ritrovati, che possono dare un contributo al racconto storiografico. Ed è per questo che possiamo parlare di public history.

La public history nasce negli Stati Uniti negli anni Settanta, ed è costituita da una vasta gamma di attività svolte da persone, con una formazione nella disciplina della storia, che operano generalmente al di fuori di ambienti accademici specializzati. La scelta del termine in inglese è motivato da una deliberata intenzione di fare riferimento ad un crescente movimento internazionale. Alcune delle istituzioni più comuni per la pratica della public history sono i musei, le dimore e i siti storici, archivi, parchi e tutti i livelli di amministrazione pubblica. La storia tradizionale, che viene insegnata nelle università, è percepita come disciplina teorica, senza capacità di presa sull'attualità e dunque poco utile, al contrario la public history si differenzia soprattutto per le sue finalità, ragioni e motivazioni e soprattutto per un diverso approccio con la società, la storia effettuata dai public historian si fa in rete, nei media, in pubblico, in diretto contatto con le comunità coinvolte nei processi di elaborazione del passato.<sup>5</sup>

---

<sup>5</sup> LA PUBLIC HISTORY innovazioni metodologiche e prospettive divulgative nella scienza storica. Una discussione con Serge Noiret, presidente del Consiglio Direttivo dell'AIPH

Oggi, in Italia, fare storia pubblica significa non solo insegnare e diffondere conoscenza su temi storici, ma significa anche arrivare all'obiettivo di coinvolgere il più possibile le comunità locali nella formazione dell'identità, della memoria collettiva e del senso di appartenenza.

Nel 2015 a Jinan, in Cina, durante la conferenza del Comitato Internazionale di Scienze Storiche che ha ospitato la seconda conferenza annuale della Federazione Internazionale di Storia Pubblica, è nata l'idea di creare un'associazione nazionale italiana di Storia Pubblica (AIPH). Quest'ultima ha deciso di scrivere un Manifesto di Storia Pubblica Italiana, attraverso il quale l'AIPH si adopera affinché i risultati e le metodologie della ricerca storiografica siano noti ad un pubblico più ampio, che la storia venga comunicata bene, e con maggiore coinvolgimento emotivo da parte del pubblico. Per gli storici pubblici italiani diventa essenziale considerare il pubblico, specializzato o meno, come interlocutori privilegiati e come potenziali protagonisti della storia, in questo modo la storia pubblica contribuirà al restauro sia della storia che degli storici interni o esterni alle università.<sup>6</sup>

Con l'avvento della tecnologia, e di una nuova fase del web, internet ha offerto a molti l'accesso alla documentazione storica in rete e alla comunicazione di forme narrative del proprio passato. In questo modo la storia è accessibile a chiunque entri in rete, e le nuove modalità di scrittura nel web permettono un'interazione tra il lavoro di chi scrive e quello di chi legge. La partecipazione pubblica alla rete è chiamata *crowdsourcing*, cioè ognuno può effettuare degli interventi critici o suggerimenti, perfino aggiungere direttamente, senza alcuna mediazione, fonti documentarie. Si parla di *public digital history*, disciplina che mette in rapporto le tecnologie di rete e la storia attraverso piattaforme sociali e i social media. Ciò ha contribuito ad aprire le porte ad un pubblico sempre più vasto, a portare la cultura alta anche in modo partecipato e, in alcuni casi, filtrato da professionisti della storia, denominati gli storici pubblici. "La storia digitale riscrive e reinterpreta il mestiere della storia attraverso la padronanza di nuove pratiche digitali".<sup>7</sup>

Quindi i ricordi di famiglia, con i relativi materiali e fonti primarie riscoperte in casa e completati con contenuti online, possono essere facilmente condivisi sul web. La storia pubblica digitale riguarda in gran parte la creazione collettiva di archivi digitali inventati, cioè quelli che fisicamente non esistono, e che sono stati caricati su internet grazie a contributi individuali e procedure di

---

<sup>6</sup> Noiret an overview of public history in Italy no longer a field without a name. Serge Noiret, European University Institute, Florence, Italy.

<sup>7</sup> Serge Noiret. Chapter Seven Digital Public History.

*crowdsourcing*. Queste attività di public history digitale favoriscono l'incontro faccia a faccia con la storia e le sue fonti, e si crea un rapporto diretto con il passato nelle comunità locali.

## 2.3 Corpus di cartoline

Il corpus è formato da 227 cartoline appartenenti alla famiglia Coliola. In particolare Giovanni Coliola era un pronipote di Giovan Battista Culiolo, il cosiddetto Maggior Leggero, uno degli attendenti più fidati di Giuseppe Garibaldi. Oliva Turtura era la fidanzata di Giovanni, poi divenuta moglie, che raccolse tutte le cartoline in un album, le quali, grazie a Rossana Borghi, nipote di Giovanni, sono giunte al museo nel 2018. Le cartoline possono essere analizzate da vari punti di vista: attraverso il retro possiamo individuare i messaggi e le informazioni che i corrispondenti si inviavano, cosa volevano esprimere tramite di essi, ma possiamo anche analizzare come funzionava la corrispondenza di un tempo. Tramite il fronte, invece, si ha una visione più artistica, attraverso le varie raffigurazioni, talvolta anche in opposizione con la tematica di guerra.

### 2.3.1 Retro cartolina

Focalizzandoci inizialmente sul retro, che presenta la parte scritta dal mittente e la corrispondenza, possiamo notare che la maggior parte delle cartoline sono state scritte da Giovanni Coliola alla sua amata Oliva. Sono per lo più cartoline d'amore dove Giovanni esprime il suo sentimento e la mancanza nei confronti di sua moglie, la sofferenza per la lontananza dall'amata, e la speranza di poterla riabbracciare ancora. Giovanni, che si firma con lo pseudonimo di Giovannino, scrive delle vere e proprie poesie d'amore, innalza il loro amore al di sopra di tutto (si veda Fig.1).

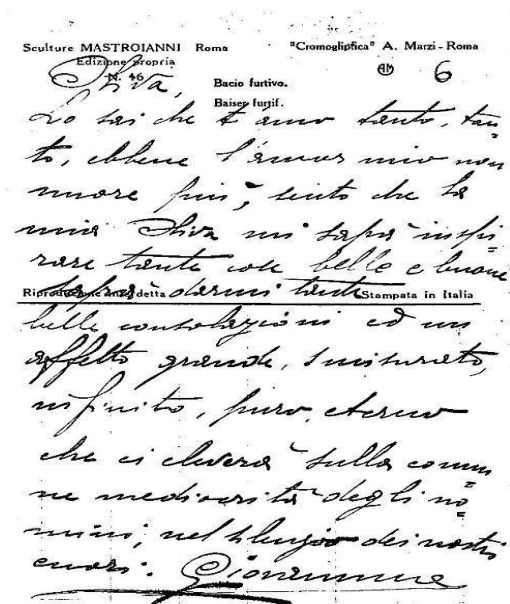


Fig.1 Cartolina 7694-006

Qui Giovanni scrive: “*Oliva, lo sai che t’amo tanto, tanto ebbene l’amor mio non muore più; tanto che la mia Oliva mi saprà ispirare tante cose belle e buone saprà darmi tante belle consolazioni ed un affetto grande, smisurato, infinito, puro, eterno che ci eleverà sulla comune mediocrità degli uomini, nel silenzio dei nostri cuori*”.

Altre cartoline invece sono state spedite, sempre da Giovanni, alla famiglia Turtura. La cartolina con codice identificativo 7694-027, qui di sotto riportata, è stata inviata ad Antonietta Turtura. Giovanni saluta tutti affettuosamente e li informa della sua attuale condizione (si veda Fig.2).



Fig.2 Cartolina 7694-027

Un altro destinatario molto frequente è Giuseppe Turtura (si veda Fig.3).



Fig.3 Cartolina 7694-047

Tutte le cartoline spedite da Giovanni Coliola presentano timbri e francobolli per cui probabilmente si può dedurre che non fosse un soldato operativo sul campo di guerra, per i quali invece erano predisposte un determinato numero di cartoline in franchigia. Questo perché, durante la prima guerra mondiale, le cartoline si dividevano in: affrancate e in franchigia.

Le cartoline affrancate erano quelle per la corrispondenza interna, oppure quelle militari indirizzate a congiunti residenti in paesi non in guerra, erano affrancate d'ufficio con francobolli dai reparti postali.

Le cartoline in franchigia, erano distribuite ai soldati, ed erano contingentate. Esse erano assegnate, in un numero limitato e stabilito, ai soldati che operavano in zone di guerra. Inizialmente i soldati avevano diritto a ricevere tre cartoline a settimana, in seguito divennero sette e poi furono ridotte ulteriormente a quattro. Si trattava di cartoline prive di raffigurazioni, esenti dal pagamento delle tasse postali. In caso fosse stato superato il numero prestabilito, esse venivano affrancate, cioè dovevano presentare timbri e francobolli postali per essere spedite. Se spedite senza il bollo di convalida del reparto o senza il timbro di posta militare, erano considerate posta civile e tassate per il doppio della tariffa.

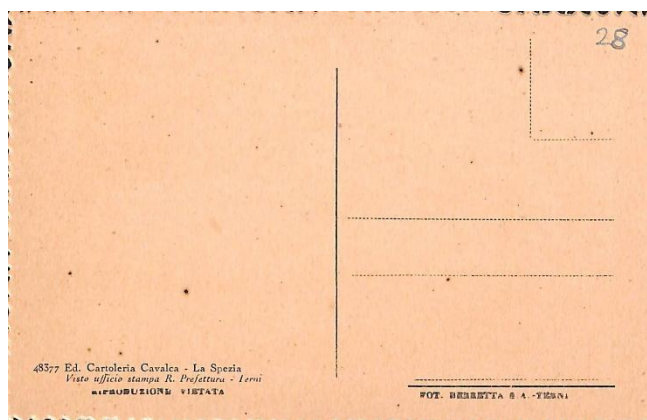
All'interno della raccolta di cartoline, sono presenti anche messaggi inviati a Giovanni stesso, da parte della famiglia o da qualche amico. Per quanto riguarda il flusso contrario, quindi dal paese alla zona di guerra, nella sezione dell'indirizzo veniva riportato il reparto e il generico "zona di guerra", questo sia per motivi di sicurezza, sia per facilitare una prima separazione dalle corrispondenze indirizzate ai civili, a quelle indirizzate ai militari in armi. Tali cartoline venivano recapitate presso i centri di smistamento militari, i quali erano in possesso di elenchi quotidianamente aggiornati dello spostamento delle truppe sui vari fronti o zone di operazioni. In seguito alla censura e alla bollatura al retro, gli uffici militari le avviavano ai vari reparti per la consegna ai destinatari.



Fig. 4 Cartolina 7694-011

La cartolina, qui di sopra raffigurata, è stata scritta a Giovanni da un suo amico, Giulio Fantoni. Come descritto precedentemente, non è presente l'indirizzo preciso, ma bensì la dicitura "zona di guerra".

Infine, sono presenti anche cartoline che non hanno viaggiato, e per questo sono prive di messaggio, non presentano alcun indirizzo, e di conseguenza non sono presenti nemmeno timbri o francobolli (si veda Fig.5).



**Fig. 5** Cartolina 7694-028

### 2.3.2 Fronte cartoline

Il fronte della cartolina presenta delle immagini molto diversificate tra loro. Solo alcune delle cartoline presentano raffigurazioni riguardanti l'ambito militare (si veda Fig.6).



**Fig. 6** Cartolina 7694-011 (fronte)

Sono rappresentati un gruppo di soldati, di grigio vestiti, armati di fucili e baionette. Il gruppo sta trasportando una bandiera italiana con al centro uno stemma del Regno d'Italia.

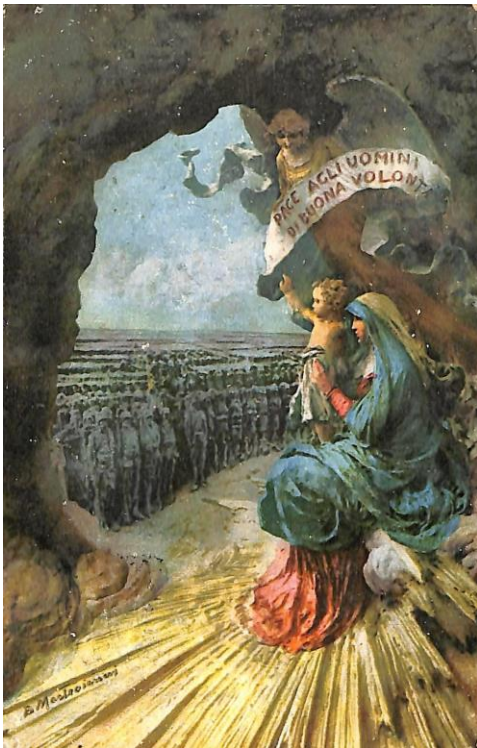
Per la maggior parte delle cartoline, le immagini sono slegate dal contesto militare, fattore dovuto al processo di allontanamento e di banalizzazione della guerra. Infatti, talvolta, sono presenti opere d'arte, questo è il caso della cartolina 7694-196 che presenta sul fronte un'opera di Raffaello Sanzio,

pittore e architetto appartenente al periodo del Rinascimento, che realizzò un autoritratto nel 1504/1506, attualmente conservato nella Galleria degli Uffizi a Firenze. Anche la cartolina 7694-225 presenta un'opera, di Antoon Van Dyck, pittore fiammingo che divenne il primo pittore di corte in Inghilterra. Principalmente realizzava ritratti, infatti in questa immagine è presente il ritratto di Amalia Solms.

Nella storia dell'arte, uno dei soggetti più ricercati e realizzati è sempre stata la donna. In effetti, tante illustrazioni riguardano proprio le donne, queste ultime sono rappresentate in tanti modi diversi: ritratti a mezzo busto di eleganti donne aristocratiche, donne con i fiori, donne con animali, donne sdraiate, o in posa, o in procinto di leggere o scrivere una lettera.

Molto frequenti sono i paesaggi e i luoghi di città, come per esempio la cartolina 7694-017, dove viene rappresentato uno scorcio di Adige in Trentino, le cartoline 7694-028 e 7694-029 presentano delle fotografie della città di La Spezia, e ancora, la 7694-200, dipinto di una splendida Venezia.

Altre raffigurazioni, spesso presenti, sono quelle con riferimento religioso. L'arte è sempre stato uno degli strumenti per rappresentare la religione, soprattutto attraverso il simbolismo.



**Fig. 7** Cartolina 7694-005 (fronte)



**Fig. 8** Cartolina 7694-059 (fronte)



Nella prima raffigurazione (si veda Fig. 7) è presente la madonna col bambino, e nella parte superiore un angelo con un'insegna *'pace agli uomini di buona volontà'*. La seconda immagine (si veda Fig. 8), in bianco e nero, rappresenta Gesù che, in groppa ad un asino, disseta due figure. Inginocchiato a terra c'è un uomo anziano e probabilmente zoppo, dato che sulle sue ginocchia possiede una stampella. In alto a destra si nota un'immagine scolorita, quasi impercettibile, di un cuore dal quale fuoriesce la luce con sopra una croce, simbolo della cristianità.

Autori di gran parte di queste raffigurazioni sono T. Corbella e D. Mastroianni, e Aurelio Bertiglia, un illustratore, pittore e pubblicitario italiano. Egli realizza maggiormente disegni di tipo fumettistico, i soggetti sono dei bambini che mettono in scena e interpretano per esempio la musica, la poesia, l'architettura, la scultura e la letteratura. Ma anche narrazioni di diverso tipo, come per esempio l'incontro tra Dante e Beatrice, sempre con protagonisti dei bambini (si veda Fig.9).



**Fig. 9** Cartolina 7694-101 (fronte)

In ultima analisi, le immagini più frequenti sono le coppie di innamorati (si veda Fig.10). Trattandosi di una raccolta di cartoline d'amore, oltre ai messaggi che Giovanni inoltra alla sua Oliva, anche le illustrazioni rappresentano il grande amore e il sentimento che lega i due corrispondenti.

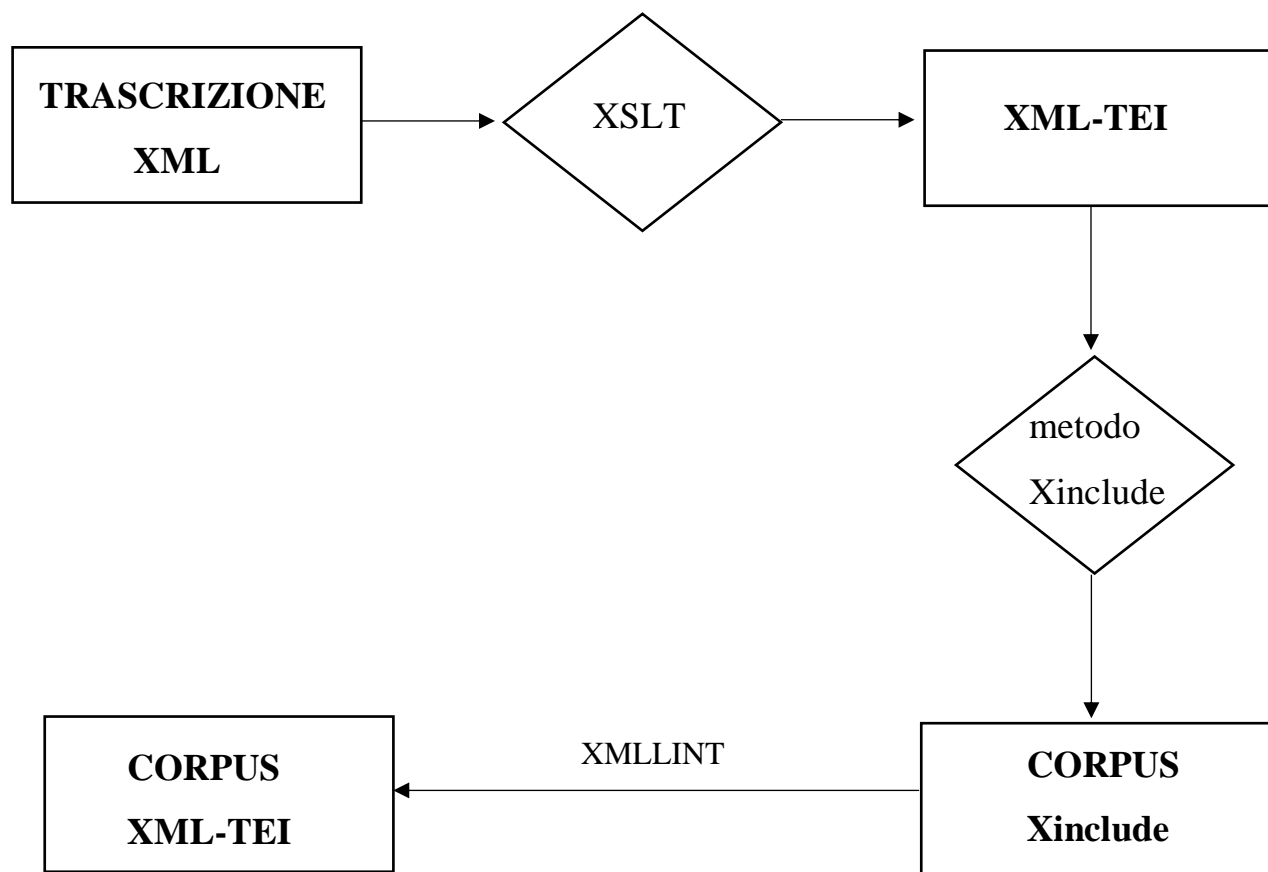


**Fig. 10** Cartolina 7694-036 (fronte)

In questo dipinto intitolato *'sempre mia'*, due giovani in un cortile, la ragazza seduta su un muretto vicino ad uno stagno, ed il ragazzo che le tiene il braccio mentre la guarda assorto. Così come nell'immagine, anche nel messaggio c'è amore, speranza, e un augurio di un per sempre insieme, infatti Giovanni scrive: *"Oliva mia, non so esprimerti la gioia che sento nell'affermarti che il 27 mattino col treno delle 8 io sarò a Ravenna, finalmente è arrivato un nostro segno; auguriamoci di cuore che tutte le nostre aspirazioni saranno anch'esse esaudite"*.

### 3. Sviluppo del progetto

La realizzazione dell'archivio digitale si articola in diverse fasi, con l'utilizzo di diverse tecnologie e strumenti software (si veda Fig.11). In una prima fase è stata eseguita la trascrizione dei testi delle cartoline. La trascrizione è stata realizzata mediante uno schema di codifica XML personalizzato con struttura molto semplice, utile a registrare il testo di ciascuna cartolina. Nella seconda fase sono stati realizzati fogli di stile XSLT per il trattamento automatico dei testi codificati, in cui le trascrizioni sono state standardizzate secondo le linee guida TEI, arrivando alla definizione di un corpus XML-TEI di tutte le cartoline. Quest'ultimo passaggio grazie al tool xmllint da riga di comando<sup>8</sup>. Nella terza e ultima fase è stato costruito il sito web di fruizione del corpus, la struttura della presentazione Web è stata sviluppata mediante le tecnologie HTML, CSS, jQuery e JavaScript. Le funzionalità sono state elaborate impiegando una libreria JavaScript innovativa: Saxon JS2.



**Fig.11** Diagramma di flusso della trasformazione da XML al corpus XML-TEI del testo

---

<sup>8</sup> xmllint é uno strumento che ci permette di elaborare, navigare, validare, etc dei file di XML tramite la linea di comando.

## 3.1 Prima fase: trascrizione XML

Il nostro corpus iniziale è formato da un insieme di cartoline, quindi testi trasmessi e conservati mediante supporti fisici. Per rendere disponibile questo patrimonio è necessario effettuare una transcodifica dei testi dal loro supporto originario verso il nuovo supporto elettronico.

Si ha la necessità di adottare una rappresentazione formale del testo, si parla di linguaggi di codifica basati su un insieme di istruzioni rigorose definite da grammatiche formali. La soluzione considerata ottimale per una corretta codifica del testo è l'adozione dei *markup language descrittivi* basati su XML. La codifica con linguaggi di marcatura è in sostanza un insieme di convenzioni, rese attraverso specifiche etichette intercalati nel testo principale per permettere agli elaboratori elettronici di distinguere le varie parti di un documento (documenti in formato machine readable). XML (*extensible markup language*)<sup>9</sup> è un insieme di regole per definire linguaggi di marcatura dichiarativi personalizzati e personalizzabili, esso è nato per strutturare, conservare e trasportare informazioni. XML ha un modello dei dati gerarchico e può essere rappresentato come un albero etichettato e ordinato. Per questo motivo le informazioni sono strutturate in modo ottimale se sono gerarchiche e sequenziali. Un documento XML consiste di tre parti: la prima parte è il prologo, che a sua volta può presentare l'*XML declaration*, unica componente obbligatoria, e opzionalmente *processing instructions*, commenti e *document type declaration*. La seconda parte è il *body*, esso contiene il contenuto in una struttura ad albero ordinata. La terza parte è l'epilogo, esso è opzionale e può contenere commenti finali e *processing instructions*. Un documento XML è composto da elementi e attributi. Gli elementi possono avere diversi tipi di contenuto: contenuto strutturale cioè formato solo da altri elementi, non testo, contenuto misto costituito da testo e altri elementi, contenuto testuale costituito da solo testo e non altri elementi. Gli elementi possono avere uno o più attributi. Un attributo descrive una caratteristica dell'elemento in cui appare. Un documento XML può essere ben formato - se rispetta le regole della specifica - oppure valido - se rispetta le regole definite mediante una grammatica formale o uno schema di codifica. Gli strumenti per descrivere le regole relative ad una corretta compilazione sono principalmente: *document type definition* (DTD) oppure *XML schema definition* (XSD). La DTD definisce regole relative alla struttura di un documento XML. Quest'ultima definisce gli elementi, gli attributi, le entità e le notazioni ammesse in un documento XML. Se un documento XML manca di riferirsi ad una DTD oppure non rispetta le regole di una DTD, esso può essere tutt'al più ben formato, ma sicuramente non può essere valido.

---

<sup>9</sup> <https://www.w3.org/XML/>

Per il progetto di tesi è stato realizzato un primo documento XML per la trascrizione di tutte le cartoline, che è stato associato ad una specifica grammatica DTD. La struttura di tale trascrizione è la seguente:

```
<cartoline>
  <cartolina id = "">
    <text>
      <retro id = "" />
      <fronte id = "" />
    <text>
    <note />
  </cartolina>
</cartoline>
```

**Listato 1:** frammento di codice XML per la trascrizione delle cartoline

Il listato 1 mostra gli elementi che sono stati utilizzati, gli elementi `<cartolina>` con attributo `@id` che registra l'identificativo univoco di ciascuna cartolina, costituito da una stringa quale ad esempio: "cartolina\_7694-001". Successivamente l'elemento `<text>` descrive il contenuto testuale della cartolina, esso contiene a sua volta gli elementi `<retro>` e `<fronte>`. All'interno dell'elemento `<retro>`, non solo si ha il testo della cartolina, ma anche l'indirizzo del destinatario, separato da uno spazio bianco rispetto al testo. In aggiunta, opzionalmente, si ha l'elemento `<salute>`, che contiene i saluti del mittente della cartolina, che può presentarsi sia nel retro che nel fronte di essa.

```
<cartolina id = "cartolina_7694-098">
  <text>
    <retro id = "R7694-098">
      Modena lì 8/6 - 1917
      Costanti pensieri.
      <salute> Giovannino </salute>

      Gent Sig.na
      Oliva Turtura
      Via Girolamo Rosfi 58
      Ravenna
    </retro>
```

```

</text>
<note>
Sono presenti un timbro e un francobollo;
98 scritto da seconda mano;
Diversità nella resa grafica di alcuni accostamenti di caratteri
es ss con sf;
Sul retro troviamo la stampa "Proprietà artistica riservata - N
1041", e "La Poesia";
Sul fronte è presenta il nome dell'artista: A.BERTIGLIA.
</note>
</cartolina>

```

**Listato 2:** trascrizione del testo della cartolina 7694-098

## 3.2 Seconda fase: XSLT e corpus XML-TEI

Nella seconda fase, dopo aver codificato un documento XML composto dalla trascrizione di tutte le cartoline, mediante la composizione di vari documenti XSL, si eseguono le opportune trasformazioni con l'obiettivo di strutturare il corpus secondo le direttive dello schema XML-TEI.

I fogli di stile descrivono il modo in cui un documento elettronico deve essere visualizzato. I metodi e le tecnologie utilizzate sono standard internazionali, tra essi, il più noto, è XSL (*extensible stylesheet language*) che trasforma un documento XML in documenti in altri formati o strutture: Html, Pdf, ecc. XSL incorpora tre linguaggi: *XSL formatting object* (XSL-FO) applicazione degli stili e della resa grafica di un documento XML, *XML path* (XPath) usato nei fogli di stile XSLT per selezionare le parti di un documento XML, e infine *XSL transformations* (XSLT) trasformazione di un documento XML in un altro tipo di documento. Esso è un vero e proprio linguaggio di programmazione che usa la sintassi XML per leggere e scrivere alberi XML, è basato su regole di trasformazione, le regole sono dichiarative, specificano che cosa deve essere generato quando si incontra una specifica struttura XML nel documento XML d'origine.

Le componenti di base di un foglio XSLT sono:

- l'intestazione XML: `<?xml version="1.0" encoding "UTF-8"?>`;
- l'elemento radice stylesheet e namespace: `<xsl:stylesheet version="2.0" xmlns:xsl="" http://www.w3.org/1999/XSL/Transform">`;
- eventuali istruzioni di elaborazione: `<xsl:output method="xml" version ="1-0" indent="yes" />`;

- serie di template rules: `<xsl:template match="" />...</xsl:template>`. Definisce una regola di trasformazione per i nodi di un particolare tipo/contesto. I template XSLT possono avere due forme: template rules che specificano una regola con pattern-matching (`<xsl:apply-templates>`), e i named templates che specificano regole che possono essere chiamate esplicitamente con `<xsl:call-template>`.

La trasformazione avviene tramite un processore XSLT come SaxonHE10 che permette di applicare uno o più scenari di trasformazione.

```
java -jar SaxonHE10-3J/saxon-he-10.3.jar -s:filexmlsorgente -xsl:filexslt -o:nuovofileoutput
```

Da linea di comando viene inoltrata tale istruzione, il processore XSLT legge il documento XML in input, che inizialmente contiene la trascrizione delle cartoline, e crea l'albero corrispondente, inizia a percorrere l'albero leggendo i singoli nodi, confronta ogni nodo con le regole presenti nel foglio di stile, cioè le *template rules*. Produce l'output secondo le istruzioni della regola e restituisce un albero di output. Le template rules che vengono applicate alle trascrizioni, permettono di inserire tag XML-TEI nel contenuto testuale.

Per esempio, attraverso un template, è stato possibile identificare la riga di testo che contiene il luogo e la data di spedizione della cartolina, ed è stato inserito il tag TEI `<dateline>`. Inoltre, sempre con regole dettate dai template, sono stati inseriti i tag `<opener>`, che indica la sezione preliminare all'inizio di una divisione testuale, i tag `<closer>` e `<signed>` per la sezione dei saluti finali e la firma, il tag `<address>` per l'indirizzo del destinatario.

Infine, è stata applicata un'ulteriore trasformata XSL per inserire l'elemento `<back>`, all'interno del quale è presente un `<div>` che contiene le parole del testo della cartolina e il relativo lemma.

L'associazione parola-lemma è stata ricavata grazie all'utilizzo di LinguA<sup>10</sup>, una pipeline di annotazioni linguistiche all'avanguardia che combina algoritmi basati su regole e machine learning.

Lo strumento di analisi linguistica include diversi passaggi di annotazione: sentence splitting, tokenization, part of speech tagging and lemmatization, dependency parsing. LinguA permette di

---

<sup>10</sup> LinguA è un pipeline per l'annotazione linguistica automatica realizzato dal laboratorio ItaliaNLP dell'istituto di linguistica computazionale "Antonio Zampolli" del CNR di Pisa

analizzare i testi e di scaricare l'analisi in formato CoNLL<sup>11</sup>. Le analisi dei testi delle singole cartoline, sono state inserite in un documento XML. Tale documento è descritto da una DTD, molto semplice. (si veda listato 3).

```
<!ELEMENT linguistic (analisi)+ >
<!ELEMENT analisi (#PCDATA)>

<!ATTLIST analisi
    id ID #REQUIRED>
```

**Listato 3:** DTD che descrive regole per il documento XML che contiene le analisi dei testi

Al termine di questi processi di trasformazione otteniamo file XML-TEI per ciascuna cartolina. La TEI (*text encoding initiative*) è un consorzio di istituzioni internazionali, di ambito linguistico e letterario, che ha sviluppato uno standard per la rappresentazione dei testi in formato digitale. L'obiettivo della TEI è quello di fornire linee guida per la creazione e la gestione in forma digitale di qualsiasi tipo di dato creato e usato in ambito umanistico. Esse privilegiano la descrizione del significato del testo, ossia la definizione del modello del testo con le proprie caratteristiche specifiche, piuttosto che gli aspetti di presentazione. Privilegia il modello del testo piuttosto che il formato in quanto, la formalizzazione e la struttura del documento è indispensabile per la creazione e la gestione in forma digitale di qualsiasi tipo di dato testuale. A questo fine la TEI propone una struttura modulare, per cui è possibile scegliere i moduli ed i relativi elementi (tagset) d'interesse, in modo da realizzare uno schema di codifica appropriato alle specifiche esigenze del progetto. La versione P5 della TEI, basata su XML, presenta molte novità, come per esempio maggior modularità. Si scelgono soltanto i moduli che corrispondono alle proprie esigenze, in modo da realizzare rapidamente uno schema di codifica appropriato. Ogni modulo dichiara un certo numero di elementi, tagset, gli elementi sono organizzati in classi strutturali o semantiche, gli attributi sono organizzati in classi globali o specifici. Ecco i moduli essenziali:

- **tei:** definisce le classi di elementi, le macro e i datatype che verranno usati per tutti i moduli;
- **header:** l'intestazione contenente i metadati relativi al documento TEI-XML;
- **textstructure:** elementi strutturali per qualsiasi tipo di testo;

---

<sup>11</sup> CoNLL (Computational Natural Language Learning), standard il cui scopo è fornire un metodo di annotazione universale che consenta la descrizione di un testo in maniera rigorosa e soprattutto strutturata, così da poterla adoperare nei processi di machine learning. The CoNLL 2007 Shared Task on Dependency Parsing



- core: elementi utili in qualsiasi tipo di documento.

La struttura minima di un documento TEI deve contenere una radice <TEI>, l'intestazione indicata con l'elemento <teiHeader> e la trascrizione codificata del testo, definito dall'elemento <text>.

L'intestazione <teiHeader> si compone di cinque elementi:

- <fileDesc> contiene una descrizione bibliografica completa del file digitale. Esso è l'unico elemento obbligatorio, si articola in una serie di altri elementi figli, tra cui <titleStmt>, <editionStmt>, <publicationStmt>, <noteStmt>, <sourceDesc>, non tutti sono obbligatori;
- <encodingDesc> documenta le relazioni tra il testo elettronico e le fonti da cui è stato tratto;
- <profileDesc> contiene una descrizione dettagliata degli aspetti non bibliografici di un testo, per esempio quando è stato creato, da chi, che lingue sono state utilizzate. Esso si articola in una serie di elementi figli, tra cui <correspDesc> che contiene una descrizione delle azioni relative a un atto di corrispondenza;
- <revisionDesc> informazioni sulle versioni del file;
- <xenData> fornisce un elemento contenitore in cui possono essere inseriti metadati in formati non TEI.

L'elemento <text> si divide in tre elementi:

- <front> è un elemento opzionale che contiene qualsiasi premessa (intestazioni, frontespizio, prefazioni, dediche, ecc) che si trova all'inizio di un documento, prima del corpo principale;
- <body> è un elemento obbligatorio (al netto dell'uso di group) che contiene l'intero corpo di un unico testo unitario;
- <back> è un elemento opzionale che contiene eventuali appendici che seguono la parte principale di un testo.

Allo stesso livello di <text> o in sostituzione ad esso, è possibile trovare l'elemento <facsimile>, che contiene una rappresentazione di una fonte scritta sotto forma di un insieme di immagini piuttosto che come testo trascritto o codificato. Per collegare il corrispettivo contenuto di un elemento a una parte dell'immagine, è possibile includere il modulo TEI *transcr*, che rende disponibile l'attributo globale *fac*s. Il <facsimile> contiene a sua volta l'elemento <surface> che definisce una superficie di scrittura in termini di uno spazio espresso attraverso un sistema di coordinate. Opzionalmente può essere inserita una collezione di elementi <graphic> che

rappresentano graficamente lo spazio indicato. È possibile caratterizzare la superficie con successivi elementi <zone> che rappresentano specifiche regioni di interesse su di essa.

```
<TEI> <!-- radice -->
  <teiHeader> <!-- intestazione -->
    <fileDesc> <!-- descrizione bibliografica -->
      <titleStmt> <!-- titolo risorsa --> </titleStmt>
      <publicationStmt/> <!-- pubblicazione e distribuzione -->
      <sourceDesc> <!-- fonte da cui il testo deriva -->
        <bibl>
          <!-- contiene una citazione bibliografica poco strutturata -->
        </bibl>
        <msDesc>
          <!-- contiene una descrizione di un singolo manoscritto
            identificabile o altro oggetto contenente testo come i primi
            libri a stampa -->
        </msDesc>
      </sourceDesc>
    </fileDesc>
    <profileDesc> <!-- aspetti non bibliografici del testo -->
      <correspDesc/> <!-- descrizione delle azioni relative a un atto
        di corrispondenza -->
    </profileDesc>
  </teiHeader>
  <facsimile>
    <surface>
      <graphic/>
    </surface>
  </facsimile>
  <text>
    <front> <!-- avantesto --> </front>
    <body> <!-- testo --> </body>
    <back> <!-- appendici --> </back>
  </text>
</TEI>
```

**Listato 4:** file XML-TEI risultante per ciascuna cartolina

Successivamente, è stato creato un corpus XML-TEI che contenesse tutte le cartoline. Trattandosi di una grande quantità di documenti, tale corpus è stato creato grazie al metodo *Xinclude*<sup>12</sup>, esso è un modello di elaborazione che specifica un metodo generale di inclusione per la creazione di grandi documenti XML composti da più documenti XML. È stato creato un file con radice `<teiCorpus>`, che comprende un `<teiHeader>` generale dell'edizione digitale, e gli elementi *xi:include*, tanti quante sono le cartoline.

```
<teiCorpus>
  <teiHeader/>
  <xi:include href="" xpointer="" />
</teiCorpus>
```

**Listato 5:** frammento di codice XML per la codifica del corpus di cartoline facendo uso dell'istruzione *xi:include*

L'elemento *xi:include* presenta l'attributo *href* che contiene un URI che indica il percorso del file da includere, e l'attributo *xpointer* che identifica la parte di risorsa da includere (si veda listato 5).

Infine, è stato utilizzato il programma `xmllint` che permette di elaborare, navigare, validare, i documenti XML tramite riga di comando. Questo strumento è disponibile sui sistemi operativi Linux/Unix installando la libreria `libxml2`. Il comando utilizzato è il seguente:

```
xmllint --output corpusTot.xml --xinclude corpusXinclude.xml
```

Il risultato finale è un file XML, denominato *corpusTot.xml*, che contiene la codifica XML-TEI di tutte le cartoline.

### 3.3 Terza fase: SaxonJS 2

La creazione del sito web, che mette insieme la raccolta di cartoline della Donazione Borghi in un archivio digitale, è stato sviluppato impiegando vari linguaggi di programmazione e strumenti:

---

<sup>12</sup> <https://www.w3.org/TR/xinclude/>

- HTML (*hyper text markup language*), un linguaggio di markup nato per la formattazione e impaginazione di documenti ipertestuali specifici per il Web;
- CSS (*cascading style sheets*), linguaggio usato per definire la formattazione, lo stile di un documento html;
- JavaScript, linguaggio di programmazione orientato agli eventi, comunemente adottato per la programmazione web lato client, esteso poi anche al lato server, per la creazione di siti web e applicazioni web;
- JQuery è una libreria JavaScript, semplifica la selezione, la manipolazione, la gestione degli eventi e l'animazione di elementi DOM in pagine html.

In aggiunta, l'implementazione dell'archivio digitale è stato messo a punto grazie ad una nuova tecnologia: SaxonJS 2.

Saxon-JS versione 2 è un processore XSLT 3.0 scritto principalmente in Javascript e in parte in XSLT. Esso è disponibile per due ambienti Javascript: sia per lo sviluppo di applicazioni client-side, sia per lo sviluppo di applicazioni server-side. Per l'uso lato client (browser) è progettato per eseguire fogli di stile, precedentemente compilati in un formato intermedio denominato Stylesheet Export File (SEF). Saxon-JS 2 per Node.js fornisce un processore XSLT 3.0 lato server completo, che può eseguire direttamente fogli di stile XSLT ed esportarli in SEF, nonché eseguire SEF. Il pacchetto software contiene effettivamente un tempo di compilazione e un componente di runtime. Il compilatore prende il codice sorgente XSLT come input e genera SEF, che è una rappresentazione intermedia del foglio di stile compilato e ottimizzato in un formato JSON personalizzato. Infatti, Saxonica<sup>13</sup> offre due compilatori XSLT:

- Compilatore XJ è un compilatore da XSLT a SEF scritto in Java e fornito come parte del prodotto Saxon-EE;
- Compilatore XX è un compilatore da XSLT a SEF scritto in XSLT, che effettua chiamate su un parser XPath scritto in JavaScript.

Saxon-JS è in grado di elaborare i file SEF generati da entrambi i compilatori. Per la maggior parte degli scopi il compilatore XJ è una scelta migliore perché è più veloce, produce una migliore diagnostica, genera codice di runtime più piccolo e veloce, e ha meno restrizioni. Mentre, il

---

<sup>13</sup> Saxonica è stata fondata dal Dr. Michael Kay nel 2004. Negli anni hanno consegnato: oltre 2500 licenze a pagamento a più di 1300 clienti, più di 80 fornitori di software hanno integrato Saxon nelle proprie applicazioni. Tra i loro prodotti c'è SaxonJS.

vantaggio del compilatore XX è che funziona su qualsiasi piattaforma in cui è supportato Saxon-JS, e inoltre è disponibile gratuitamente.

Saxon-JS viene caricato utilizzando l'elemento di script HTML standard all'interno dell'intestazione o del corpo dell'HTML:

```
<script src = "saxon-js/SaxonJs2.rt.js"> </script>
```

L'approccio progettuale consigliato con Saxon-JS, non consiste nel rendere l'intera pagina HTML dal foglio di stile, ma piuttosto nel generare frammenti di HTML che possono essere inseriti nello scheletro statico della pagina nei punti appropriati. Quindi la pagina HTML iniziale conterrà generalmente una struttura scheletrica di elementi come *div*, ciascuno con un attributo *id*, in attesa di essere popolata dal foglio di stile. Infatti, l'istruzione XSLT *xsl:result-document* viene utilizzata per generare la struttura di output dalla trasformazione XSLT e per determinare dove questa struttura deve essere aggiunta all'HTML. L'istruzione è la seguente:

```
<xsl:result-document href = "#abcd">
```

Essa fa sì che l'albero dei risultati venga scritto come l'ultimo, o unico, figlio dell'elemento HTML con un determinato valore *id*, in questo caso 'abcd'.

La pagina HTML dell'archivio di cartoline è formata da diversi *div*, vuoti, che presentano solo l'attributo *id*. Per esempio:

```
<div id = "home"> </div>  
<div id = "lecartoline"> </div>  
<div id ="info"> </div>
```

Ogni *div* riguarda una porzione di sito web, che viene generata da un file XSL, che successivamente verrà trasformato in file SEF. Ci sono tanti file xsl/sef, quante sono le porzioni html da generare.

Ciascuna pagina xsl ha inizio con l'intestazione XML e l'elemento radice per la trasformazione di un documento (si veda listato 6).

```
<?xml version="1.0" encoding="UTF-8"?>  
<xsl:stylesheet version="2.0"  
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
```

```
xmlns:tei="http://www.tei-c.org/ns/1.0"
xmlns="http://www.w3.org/1999/xhtml">
```

**Listato 6:** intestazione obbligatoria presente in un file xsl

Segue l'istruzione `<xsl:output method="html"/>`, esso definisce il formato del documento di output. È un elemento di primo livello e deve apparire come nodo figlio di `<xsl:stylesheet>` o `<xsl:transform>`

All'interno della pagina xsl è presente anche l'istruzione `xsl:result-document`, con *href* uguale all'*id* del div presente nel html. Se l'*href* fornito non corrisponde a un elemento esistente nella pagina HTML, l'istruzione viene ignorata. Inoltre sono presenti le template rules, cioè delle regole, da applicare per generare il corpo della pagina. Lo schema di base è il seguente:

```
<xsl:template match="/">
  <xsl:result-document href="">
    <!--varie istruzioni-->
  </xsl:result-document>
</xsl:template>
```

```
<xsl:template match=""> <!--altri template che vengono richiamati
all'interno di result-document con <xsl:apply-templates select=""> -->
```

**Listato 7:** frammento di pagina xsl, che genera frammenti di html con SaxonJs2.

Successivamente, dopo aver creato la pagina xsl, essa deve essere trasformata nel formato SEF. Per fare ciò viene utilizzata la seguente istruzione da riga di comando:

```
xslt3 -t -xsl:stylesheet.xml -export:stylesheet.sef.json -nogo
```

- *-t* fa sì che il comando visualizzi i messaggi di avanzamento;
- *xsl:stylesheet.xml* identifica il foglio di stile di origine;
- *export:stylesheet.sef.json* identifica il file in cui verrà scritto il file SEF generato;
- *-nogo* sopprime qualsiasi tentativo di eseguire il foglio di stile;

In seguito, vengono create le funzioni di trasformazione che generano ciascuna una porzione di html differente.

```
SaxonJS.transform(options [, execution])
```

L'argomento delle *options* è un oggetto JavaScript che fornisce le opzioni di trasformazione. La funzione utilizzata nel sito è la seguente:

```
function() {
  saxonJS.transform({
    stylesheetLocation: "xsl/stylesheet.sef.json",
    sourceLocation: "source.xml"
  }, "async");
}
```

**Listato 8:** funzione SaxonJs che generare frammenti di html

Nel listato 8, il primo argomento della funzione è *SaxonJS.transform* che presenta due parametri, *stylesheetLocation* è un URI che può essere utilizzato per individuare il foglio di stile compilato, un file SEF in formato JSON. Il secondo parametro è *sourceLocation*, anch'esso un URI che può essere utilizzato per individuare il documento XML o JSON, nel nostro caso prende il documento XML che contiene l'intero corpus delle cartoline codificate.

Il secondo argomento invece è una stringa, e le due opzioni per il suo valore sono 'sync' e 'async'. Se l'argomento è omissso, o è la stringa 'sync', la funzione di trasformazione è sincrona, cioè ritorna quando la trasformazione è completa e l'argomento delle opzioni sarà stato aggiornato per includere i risultati della trasformazione. Se invece è presente, come nel nostro caso, la stringa 'async', la funzione attiva la trasformazione in modo asincrono, restituendo una *promise*<sup>14</sup>. Quest'ultima si realizza quando la trasformazione si completa, con successo o senza successo. Al completamento con successo, il valore passato alla funzione di risoluzione è l'oggetto opzioni, aggiornato per contenere le informazioni sui risultati.

Un altro parametro opzionale utilizzato è *stylesheetParams*, un oggetto JavaScript contenente i valori dei parametri globali del foglio di stile. La chiave è il nome del parametro sia come nome locale semplice, sia come nome completo URI. Il valore corrispondente è un valore JavaScript. Questi non possono essere utilizzati per fornire parametri statici, i valori di questi ultimi erano già vincolati quando il foglio di stile è stato compilato. Nel foglio di stile xsl corrispondente, è inserito il

---

<sup>14</sup> La promise è un oggetto che produce e consuma un risultato. Essa invece di attendere che una funzione lenta restituisce un risultato, viene definita un'altra funzione che gestisce il suo risultato, continuando con l'elaborazione che non dipende da esso.

comando `<xsl:param name = "id" />`, che viene utilizzato per dichiarare un parametro locale o globale. L'attributo *name* specifica il nome del parametro.

Per esempio, nella funzione che visualizza l'immagine della cartolina (si veda listato 9), è stato utilizzato *stylesheetParams* per prendere l'attributo *id* di ciascuna immagine della cartolina, in modo tale che al click dell'icona corrispondente a tale funzione, venisse visionata esattamente l'immagine con identificativo corretto.

```
function displayImg(e1) {
  let id = e1.id;
  SaxonJS.transform({
    stylesheetLocation: "xsl-sef/infoCartolina.sef.json",
    sourceLocation: "corpusTot.xml",
    stylesheetParams: {
      "id" : id
    }
  }, "async");
}
```

**Listato 9:** esempio di funzione SaxonJS 2 che genera la vista dell'immagine della cartolina

### 3.3.1 Funzionalità di ricerca

La pagina che permette di interrogare il corpus delle cartoline, all'interno del sito, in base ad un criterio di ricerca, è stata sviluppata in maniera differente rispetto alle altre. Allo scopo non è stato utilizzato Saxon-JS 2.

Innanzitutto, è stato creato in html un `<div id = "cerca">` che presenta all'interno tre `<input type="radio">` che indicano i tre criteri con cui l'utente può effettuare la ricerca, essi sono: *il titolo*, *il mittente e la parola chiave*. Nella parte sottostante, è presente una casella di testo, inizialmente vuota, dove l'utente può inserire la stringa di interrogazione.

Al fine di implementare la funzionalità di ricerca, sono stati creati tre file xsl, uno per ciascun criterio di ricerca. È stata realizzata una trasformata che prende in input il 'corpusTot.xml', che contiene le codifiche XML-TEI di tutte le cartoline, il file di ricerca xsl, e in output viene generato un indice, inizialmente in formato xml. Successivamente, tale indice viene convertito in un file JSON. La



struttura dell'indice risultante è un array, dove ogni elemento è strutturato in tre valori: *value*, *id*, *descrizione* (si veda listato 10):

- “*value*”: identifica il valore di ricerca. Nell'indice per la ricerca attraverso il titolo, il valore di questa proprietà rappresenta il titolo della cartolina, nell'indice per la ricerca attraverso il mittente, il valore di questa proprietà rappresenta il nome del mittente, e nell'indice della ricerca attraverso la parola chiave, il valore di questa proprietà rappresenta il lemma;
- “*id*”: nell'indice per la ricerca attraverso il titolo, il valore di questa proprietà rappresenta l'identificatore univoco della cartolina corrispondente a quel titolo, mentre negli indici per la ricerca attraverso mittente e parola chiave, il valore di questa proprietà rappresenta il numero totale delle cartoline che contengono quel determinato mittente o parola chiave;
- “*descrizione*”: identifica, con il tag <img>, l'immagine della cartolina che viene visualizzata come risultato di una ricerca.

```
[
  {
    "value": "Serenata a Lola",
    "id": "7694_001_001/002",
    "descrizione": "<img src='img-cartoline/cartolina_7694-001F.jpg' alt='cartolina_7694-001' id='cartolina_7694-001' class='imgCerca' onclick='displayMeta(this.id)'onclick='displayMeta(this.id)'/>"
  },
  {
  },
  <!--altri elementi -->
]
```

**Listato 10:** schema dell'indice utilizzato per la ricerca tramite titolo della cartolina

Successivamente, l'indice viene inserito all'interno del sito grazie a JQuery, con il seguente codice:

```
var xobj = new XMLHttpRequest();
xobj.overrideMimeType("application/json");
xobj.open('GET','indiceTitolo.json', false);
xobj.send("");
var idx = xobj.responseText;
```

**Listato 11:** frammento JQuery per l'inserimento dell'indice json

In seguito, grazie ad una funzione JQuery, al click viene effettuato un `JSON.parse()` dell'indice, così i dati che inizialmente sono delle stringhe, vengono trasformati in un oggetto JavaScript.

Successivamente, la funzione di completamento automatico fornisce suggerimenti durante la digitazione nella casella di testo. In questo modo verrà effettuata la ricerca correttamente, che porterà al risultato di una o più cartoline, in base al criterio. Di ciascuna cartolina presente tra i risultati verrà visionata l'immagine, che al click porterà alla pagina di presentazione delle informazioni della stessa.

Nello specifico, per quanto riguarda la ricerca attraverso la parola chiave, per creare l'indice è stata effettuata una lemmatizzazione del testo. I testi delle cartoline sono stati lemmatizzati grazie al tool `linguA`, una pipeline di annotazioni linguistiche, che permette di analizzare i testi e di scaricarne l'analisi in formato CoNLL. In seguito, tali analisi, vengono registrate all'interno di un documento XML (si veda listato 12).

```
<analisi id="cartolina_7694-031">
  1   Ti   ti   P   PC   num=s|per=2|gen=n   2   clit
  2   desidero   desiderare V   VM   num=s|per=1|mod=i|ten=p   0
  -
  3   fortemente   fortemente B   B   _   2   mod

  1   Giovannino   Giovannino S   SP   _   0   ROOT
</analisi>
```

**Listato 12:** lemmatizzazione testo della cartolina 7694-031

Successivamente, grazie ad una trasforma XSLT, viene inserito il tag `<back>` all'interno della codifica di ciascuna cartolina (si veda listato 13).

```
<back>
  <div type="linguistic">
    <ab>
      <seg lemma="ti">
        <w type="form">Ti</w>
      </seg>
```

```
<seg lemma="desiderare">
  <w type="form">desidero</w>
</seg>
<seg lemma="fortemente">
  <w type="form">fortemente</w>
</seg>
<seg lemma="Giovannino">
  <w type="form">Giovannino</w>
</seg>
</ab>
</div>
</back>
```

**Listato 13:** frammento di codifica della cartolina 7694-031 che contiene la lemmatizzazione del testo

Infine, grazie ad un ulteriore trasformata XSLT, viene creato l'indice che contiene i lemmi delle cartoline. Ogni lemma corrisponde a un elemento dell'array, ed è presente nell'oggetto "value".

Per esempio, nel caso visto precedentemente della cartolina 7694-031, essa sarà presente nell'oggetto "descrizione" con tag:

```
<img src='img-cartoline/cartolina_7694-031F.jpg' />
```

in tutti gli elementi con "value" corrispondenti ai lemmi: "ti", "desiderare", "fortemente", "Giovannino".

## 4. Risultato finale: Archivio digitale

Il sito web è stato sviluppato con l'obiettivo di garantire all'utente finale un'esperienza amichevole attraverso una fruizione semplice del corpus digitale. L'accesso facile e veloce permette di approfondire in maniera puntuale il repertorio di cartoline.

Al fine di rilasciare pubblicamente i codici sorgenti del progetto, tutto il materiale è stato caricato su una repository GitHub: <https://github.com/DomiElia/SitoCartoline.git>.

La *home page* dell'archivio digitale si presenta come illustrato in fig. 11:



**Fig.12** home page sito web

Nella parte superiore è presente il menù principale, che si conserva in tutte le altre pagine, per garantire una navigazione efficiente tra le varie sezioni del sito. Al menu segue il titolo dell'archivio, e successivamente le immagini in anteprima di tutte le cartoline presenti nella raccolta. Ogni riquadro, a sua volta, presenta alcune opzioni che permettono all'utente di accedere agli approfondimenti relativi a ciascuna cartolina.


Al click della *preview* dell'immagine, viene visualizzata la sezione dei metadati:

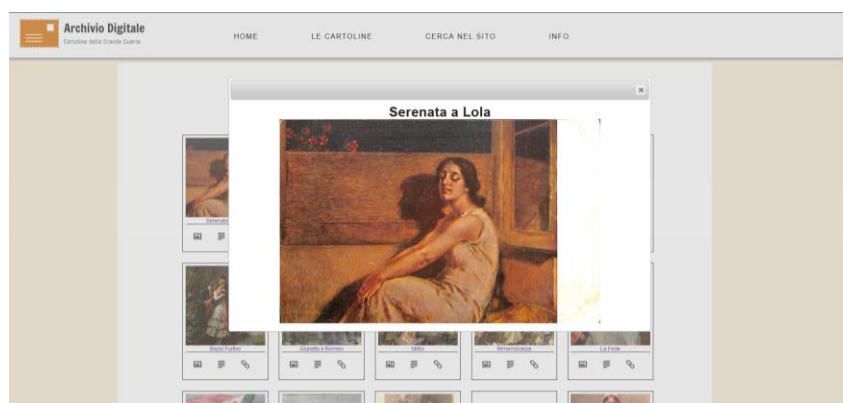


**Fig.13** metadati sito web


Nel contesto di questa vista (si veda Fig. 13), viene rappresentata nuovamente l'immagine della cartolina posta sul fronte, accanto il nome della cartolina e le informazioni sull'opera d'arte, come per esempio il nome dell'autore e la descrizione e le informazioni proprie della cartolina, come l'editore, il luogo e la data di pubblicazione, il codice identificativo e la collocazione. Nella parte inferiore sono poste due frecce, per scorrere il repertorio avanti e indietro, che consentono di cambiare visuale, e navigare tra le altre cartoline, sempre nella sezione delle informazioni riguardanti i metadati.

Nella home, così come per le altre pagine del sito, ci sono ulteriori tre opzioni di navigazione:

-  Visualizzazione in primo piano dell'immagine della cartolina. Al click si apre una finestra di dialogo, che presenta l'immagine, e che può essere personalizzata dall'utente oppure chiusa con l'apposito pulsante in alto a destra.




**Fig. 14** icona foto sito web

-  Visualizzazione del testo della cartolina. Sono presenti le frecce, avanti e indietro, per cambiare la visuale sul testo di altre cartoline.



**Fig. 15** testo cartolina sito web

-  Visualizzazione contemporanea del testo e dell'immagine, in questo modo è possibile mettere a confronto la scrittura manuale presente sul retro della cartolina, e la codifica digitale. Grazie a questa vista è possibile cogliere anche l'indirizzo del destinatario, posto accanto al testo, eventuali descrizioni di timbri e francobolli e il testo tipografico se presente. Sono presenti le frecce, avanti e indietro, per spostare la visuale su altre cartoline.



**Fig.16** testo-immagine sito web

Dal menù principale si può continuare la navigazione sulle altre pagine del sito.

Al click della dicitura 'LE CARTOLINE' viene mostrata una descrizione dettagliata della raccolta di cartoline, da dove provengono, chi le ha donate e altre informazioni sul repertorio. Al click della dicitura 'INFO' vengono presentate le informazioni riguardanti le persone che hanno contribuito alla

realizzazione dell'archivio digitale, in tutte le sue fasi, dalla codifica, alla supervisione alla codifica, direttore dei servizi culturali, funzionario responsabile e responsabile scientifico.

Infine, l'ultima funzionalità del sito, è la sezione 'CERCA':



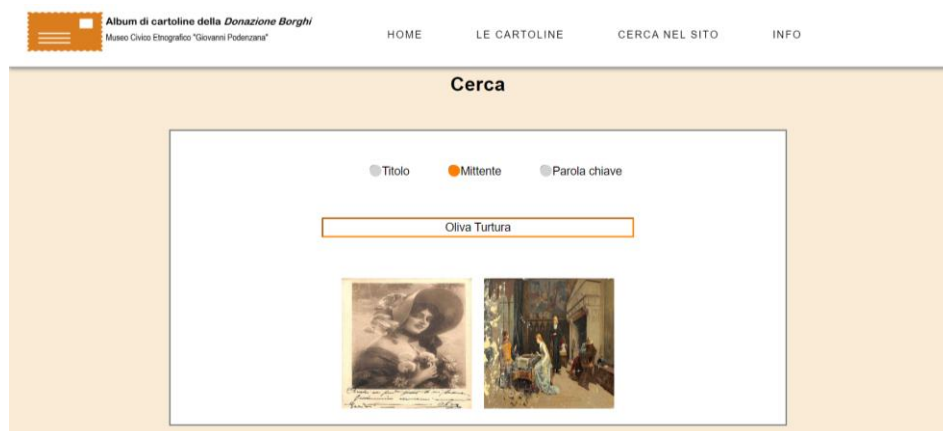
**Fig.17** cerca nel sito

Trattandosi di una vasta raccolta di cartoline, nello specifico il corpus consta di 227 cartoline, il modo più semplice ed immediato per individuare una o più cartoline specifiche, è la sezione cerca. Vengono presentati tre criteri di ricerca, il primo in base al titolo della cartolina, il secondo in base al mittente, e il terzo tramite una parola chiave. L'utente può selezionare il criterio che preferisce, e in seguito scrivere all'interno della casella di testo la parola con cui intende interrogare la raccolta. Grazie alla funzione di autocompletamento, verranno dati dei suggerimenti (si veda Fig. 18).



**Fig.18** autocompletamento cerca

Il risultato della ricerca si presenta in questo modo:



**Fig.19** risultato ricerca

Verranno visualizzate le immagini delle cartoline che corrispondono alla ricerca effettuata dall'utente. Al click dell'immagine stessa, l'utente verrà portato alla sezione metadati di tale cartolina, e da lì potrà navigare tra le informazioni della stessa.



## 5. Conclusioni

Il lavoro di tesi si è incentrato sulla creazione di un archivio digitale, in grado di contenere una vasta raccolta di cartoline della Grande Guerra. L'ambiente web è stato progettato affinché sia accessibile da un più ampio pubblico possibile, grazie ad un'interfaccia utente semplice e funzionale.

Lo sviluppo si è articolato in diverse fasi: il primo passo è stata la trascrizione e codifica dei testi, successivamente, grazie a trasformate XSLT, è stato ottenuto un corpus XML-TEI che comprendesse le codifiche di tutte le cartoline che costituiscono la raccolta. Queste trasformate XSLT sono state sviluppate con principi di genericità e quindi applicabili anche ad un qualsiasi corpus di cartoline che presenta la stessa struttura minimale.

In seguito, il corpus è stato utilizzato per creare il sito web di fruizione e approfondimento, grazie soprattutto all'innovativo processore XSLT Saxon-JS 2, scritto in JavaScript e in grado di funzionare sia sul browser, lato client, sia su Node.js, lato server. L'aggiornamento principale della versione 2.0 rispetto alla precedente, è l'inclusione all'interno del pacchetto di un nuovo compilatore in JavaScript, per la conversione dei fogli di stile XSLT nel formato di esportazione riconosciuto da Saxon-JS, il formato SEF (stylesheet export file).

La scelta di utilizzare Saxon-JS 2 nasce principalmente dalla semplicità d'uso della libreria e dalla possibilità di astrazione rispetto alle singole tecnologie dei vari browser. Infatti, i browser attraverso Javascript possono eseguire trasformazioni XSLT tramite istanze dell'oggetto XSLTProcessor. Una volta che è stato istanziato, il processore XSLTP implementa un metodo `XSLTProcessor.importStylesheet()` che accetta come argomento il foglio di stile XSLT da utilizzare nella trasformazione. Tale foglio di stile, a sua volta, deve essere caricato nella pagina web attraverso l'oggetto `XMLHttpRequest` prima di invocare il metodo precedentemente descritto. Invece, per quanto riguarda Saxon-JS 2, viene utilizzata semplicemente una funzione di trasformazione che prende in input un file SEF (in origine file xsl), ed un file XML, generando le porzioni di html che compongono il sito web.

Eventuali possibili sviluppi di tale progetto, potrebbero puntare ad un maggior utilizzo di Saxon-Js 2, applicando altre funzionalità fornite dalla libreria. Per esempio, dato che l'obiettivo è implementare un'interfaccia utente interattiva senza manipolare gli oggetti nativi del processore

XSLT in JavaScript, si potrebbe analizzare meglio la gestione degli eventi di interazione dell'utente sugli elementi del DOM. Ad esempio, nel nostro caso, invece di utilizzare la funzione di *onclick* all'interno dell'html, si potrebbe ricorrere al meccanismo disponibile in Saxon-JS grazie all'uso dell'attributo *mode="ixsl:onclick"* all'interno della pagina XSLT stessa. Inoltre, per migliorare funzionalmente il sito da parte dell'utente, si potrebbero inserire più criteri di ricerca all'interno della sezione 'cerca nel sito', come per esempio in base al luogo di pubblicazione, all'editore, alla data di pubblicazione etc. In questo modo l'utente avrà più libera scelta, e una ricerca più ottimale.

In conclusione, l'archivio digitale realizzato nel presente progetto di tesi, non solo ha combinato più tecnologie e strumenti: XML, TEI, XSLT, Saxon-JS 2, JavaScript e JQuery; ma ha anche fornito all'utente uno strumento utile per navigare e acquisire informazioni di carattere storico. Attraverso la storia d'amore tra Giovanni Coliola e Oliva Turtura, possiamo comprendere la speranza e la paura vissuta dalla popolazione in quel periodo storico che ha coinvolto il mondo intero ovvero la Grande Guerra avvenuta tra il 1914 e il 1918; ma soprattutto, l'importanza delle cartoline, uno strumento di comunicazione ormai obsoleto, che però ad oggi è una fonte essenziale, ed una testimonianza da non sottovalutare.

# Bibliografia

Tomasoni G. e Nuvoli C. 2004. *La grande guerra: raccontata dalle cartoline*, Arca Edizioni.

Fognolo, Fabio, Piero Ambrosini, e Enrico Meliaddò. 2012. *La grande guerra: il fronte italiano nelle cartoline e nelle stampe degli artisti*. Cierre.

Noiret. 2018. Chapter Seven. *Digital Public History*. Dean.

Noiret. 2019. *An overview of Public History in Italy: No longer a field without a name*. Oldenbourg.

F. Boschetti, A.M. Del grosso, E. Salvatori. 2021. *DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale*. AIUCD.

Salvatori E., Boschetti F., Del Grosso A. M. 2019. *From collaborative transcription to interdisciplinary education: the postcards of the Great War case*. In: AIUCD Book of Abstract. Udine.

Salvatori E., Ricci L., Melighetti F., Boschetti F., Del Grosso A. M. 2019. *DH as an ideal educational environment for participatory learning: the Ethnographic Museum of La Spezia*. In: AIUCD Book of Abstract. Milano.

Williams, I. 2009. *Beginning XSLT and XPath: Transforming XML Documents and Data*. Wiley.

Kay, M. 2011. *XSLT 2.0 and XPath 2.0 Programmers's Reference*. Wiley.

Joakim Nivre, Johan Hall, Sandra Kubler, Ryan McDonald, Jens Nilsson, Sebastian Riedel, Deniz Yuret. *The CoNLL 2007 Shared Task on Dependency Parsing*.

Attardi G., Dell'Orletta F. "Reverse Revision and Linear Tree Combination for Dependency Parsing". In: NAACL-HLT 2009 – North American Chapter of the Association for Computational

Linguistics – Human Language Technologies (Boulder, Colorado, June 2009). Proceedings, pp. 261 – 264. Association for Computational Linguistics, 2009.

## Sitografia

*La cartolina illustrata*, <<https://fototecatrieste.it>>

*Parole in trincea. La memoria della grande guerra nelle testimonianze scritte dai soldati*, <<http://www.novecento.org> >

*Museo Civico Etnografico, “G. Podenzana” della Spezia*, < <http://etnografico.museilaspezia.it> >

*Servizio postale militare italiano*, < <http://www.postaesocieta.it> >

*Attenti alla censura*, <<http://www.centenario1914-1918.it/it/2018/10/15/attenti-alla-censura>>

*TEI (text encoding initiative)*, <<http://www.tei-c.org/index.xml>>

*LinguA*, <<http://www.italianlp.it/demo/linguistic-annotation-tool>>

*SaxonJS2*, < <https://www.saxonica.com> >

*Repository del sito su GitHub*, <<https://github.com/DomiElia/SitoCartoline.git> >