

Elaborazione di dati semi-strutturati: ipotesi implementative e casi d'uso tratti da testi in inglese antico

Roberto Rosselli Del Turco

Università di Torino, Italia

roberto.rossellidelturco@unito.it

Abstract

Per quanto EVT si sia pian piano evoluto in uno strumento versatile e potente per la pubblicazione di fonti e testi letterari sulla base del formato XML/TEI, nessuna delle modifiche – e periodiche riscritture – della sua base di codice ha cambiato significativamente il modo in cui vengono trattati i dati dell'edizione. Ancora oggi, infatti, l'obiettivo è quello di visualizzare e navigare tali dati, senza dubbio in maniera sofisticata (gestione dei livelli di edizione, delle *named entities*, del rapporto testo-immagine), ma limitandosi in ogni caso a quelle che potremmo definire funzionalità di base per una DSE. I testi codificati in XML/TEI, viceversa, rappresentano un potenziale tesoro di informazioni che attende solo di essere interrogato e messo a disposizione dell'utente. In questo articolo individuo tre casi d'uso – l'elaborazione dei dati relativi ai caratteri speciali inseriti per mezzo della `<charDecl>` e dell'elemento `<g>`; la gestione di *named entities*, *realia* e altri elementi del testo; l'uso di ontologie all'interno di un documento XML/TEI – relativi a testi letterari in inglese antico per proporre una codifica TEI e una successiva elaborazione che permetta all'utente di ricevere risposte a interrogazioni complesse, *query* trasversali che implicano collegamenti incrociati fra tipi diversi di elementi.

Although EVT has slowly evolved into a versatile and powerful tool for publishing historical sources and literary texts on the basis of the XML/TEI format, none of the modifications - and periodic rewrites - of its code base has significantly changed the way in which edition data is handled. Even today, in fact, the main goal is to visualise and navigate such data, no doubt in a sophisticated way (support for multiple edition levels, support for named entities management, text-image linking), but in any case limited to what we might call “core functionalities” of a DSE. Texts encoded in XML/TEI, on the other hand, are a potential treasure trove of information just waiting to be interrogated and made available to the user. In this article I identify three use cases – the processing of special characters encoded by means of the `<charDecl>` and `<g>` elements; the management of named entities, *realia* and other interesting elements of the text; the use of ontologies within an XML/TEI document – related to literary texts in Old English in order to

propose a TEI encoding and a subsequent processing able to put the user in a position to receive answers to complex, transversal queries involving cross-links between different types of elements.

Introduzione

EVT (Edition Visualization Technology)¹ è un software open source per la visualizzazione di edizioni digitali (diplomatiche o critiche) basate sulla codifica dei testi nel formato XML/TEI.² Grazie al lavoro di questi anni,³ e grazie all'impegno di tutto il team di sviluppo, siamo riusciti a creare uno strumento valido e molto apprezzato, come testimonia il buon numero di edizioni che ne fa uso.⁴ EVT 2, versione basata su AngularJS, ha raccolto l'eredità di EVT 1, aggiungendo molte funzionalità importanti, a cominciare dal supporto per le edizioni critiche.

Se partiamo dal requisito iniziale - la pubblicazione sul web di un'edizione diplomatica, o critica basata su testimone unico, con immagini del manoscritto - la strada percorsa è stata davvero tanta. Nel 2021, tuttavia, il panorama per le DSE (Digital Scholarly Edition) è completamente diverso rispetto a quello di 10-15 anni fa.⁵ C'è stata una notevole attività di riflessione sul piano teorico, con l'intento di affrontare problemi che vanno dalla modellizzazione del testo alla manutenibilità e sostenibilità delle DSE. Anche sul piano tecnico, in relazione abbastanza fluida con gli sviluppi teorici di cui sopra, sono stati fatti notevoli passi in avanti.

I principali punti di svolta per le DSE del futuro, da prendere in considerazione come possibili obiettivi per il successore di EVT 2, sono molteplici: dal supporto a risorse di tipo LOD e protocolli aperti come IIIF e CTS/DTS per realizzare l'edizione distribuita, alla convergenza da un lato verso dispositivi di tipo *mobile*, dall'altro verso le edizioni tradizionali grazie a una conversione del testo in un formato adatto alla stampa (PDF).

Dobbiamo in ogni caso notare che il processo di preparazione di un'edizione digitale si è notevolmente standardizzato nel corso degli ultimi anni, un fenomeno senz'altro positivo, ma che forse ci induce a trascurare aspetti che esulano da un *workflow* ormai relativamente semplice, se non 'semplificato':

1 Home page: <http://evt.labcd.unipi.it/>.

2 TEI Consortium: <https://tei-c.org/>.

3 La prima versione di EVT si è affacciata alla storia nel lontano 2013 grazie al lavoro di Raffaele Masotti, per ulteriori dettagli si veda [15].

4 Si veda, ad esempio, il *Codice Pelavicino Digitale* (<http://pelavicino.labcd.unipi.it/>), al momento basato su EVT 1, mentre un'interessante edizione che in parte aggiunge codice personalizzato a EVT 2 è quella del *Progetto Vasto* (<https://dharc-org.github.io/progetto-vasto/>).

5 E anche prima, i desiderata per il *Vercelli Book Digitale* risalgono al 2003-4, per non parlare di edizioni pionieristiche come l'*Electronic Beowulf* di Kevin S. Kiernan e *The Exeter Anthology of Old English Poetry* di Bernard J. Muir.

- codifica dei testi nel formato XML/TEI;
- pubblicazione con uno strumento di visualizzazione selezionato in base agli obiettivi iniziali;⁶
- fruizione sul web da parte dell'utente finale: lettura e studio dei testi, ricerche testuali, visualizzazione e analisi delle immagini, etc.

Si tratta di un percorso comune a praticamente tutte le DSE più recenti, comprese quelle basate su EVT.⁷ A tal proposito, nel passaggio da EVT 1 a EVT 2 è cambiata radicalmente la struttura del codice, ma il trattamento dei dati è identico:

- obiettivo primario resta la visualizzazione di testi e immagini;
- affiancata da alcuni strumenti di base, come la ricerca testuale e alcuni semplici strumenti relativi alle scansioni dei manoscritti;
- vi sono poi alcuni strumenti supplementari che riguardano le *named entities* (recupero informazioni, liste, navigazione delle occorrenze);
- per l'edizione critica sono disponibili filtri e altre modalità di analisi delle varianti (*heat map*, vista collazione);
- mentre per le edizioni diplomatiche in facsimile un semplice *pre-processing* dei dati permette la visualizzazione della struttura fascicolare del manoscritto grazie al software VisColl.⁸

Risulterà evidente da questa breve lista che le operazioni consentite da EVT riguardano principalmente la selezione del contenuto, la sua visualizzazione e la navigazione fra sezioni diverse dell'edizione digitale. Questo non dovrebbe sorprendere oltre modo, dato che si tratta appunto di una Edition Visualization Technology. Che cosa manca? Come spesso è accaduto nella storia dello sviluppo di questo strumento, la risposta a questa domanda è rappresentata dalle esigenze di ricerca di un progetto di pubblicazione con EVT, nella fattispecie l'edizione sperimentale della *Lira* senese da parte di un giovane ricercatore dell'Università di Siena, Marco Giacchetto.⁹ Commentando il risultato di una prima sperimentazione, Marco mi ha chiesto se, oltre alla visualizzazione di testi e immagini, fosse possibile far effettuare a EVT un processing dei dati in grado di rispondere a interrogazioni del testo necessarie per l'analisi di tali documenti: quanti commercianti erano attivi nelle contrade senesi fra 1200 e 1300? quanti di loro disponevano di fondaci di proprietà? quanta seta di produzione orientale è stata venduta nel 1348

6 Rispetto a dieci anni fa l'ecosistema di strumenti di preparazione e visualizzazione di edizioni basate sul formato XML/TEI è sicuramente molto più ricco.

7 Sia EVT 1, v. ad esempio il *Codice Pelavicino Digitale* (<http://pelavicino.labcd.unipi.it/>), sia EVT 2, v. le numerose edizioni prodotte dal gruppo di ricerca guidato da Fulvio Delle Donne (<https://bup.unibas.it/library/DH>).

8 Home page: <https://viscoll.org/>.

9 Il cui contributo su questo argomento, "Problemi e questioni nello studio delle fonti fiscali tardomedievali: la "Lira" senese nel XV secolo", troverete in questo stesso numero di *Umanistica Digitale*.

e con quale margine di guadagno? si possono fare calcoli sulla base dei dati numerici relativi a transazioni commerciali codificati nel documento TEI?

In breve, manca la possibilità di interrogare i dati testuali dell'edizione per ottenere risposte a domande che incrociano il contenuto di elementi diversi o, più semplicemente, che facciano calcoli sugli elementi con contenuto numerico.¹⁰ Questa capacità è sicuramente fondamentale per gli studi storici, ma può essere molto rilevante anche per ricerche di tipo filologico e/o letterario. Il testo codificato in formato TEI costituisce un "tesoretto" di informazioni che restano almeno in parte inutilizzate, probabilmente perché siamo ancora condizionati dal "printed page paradigm" a vedere l'edizione digitale come un punto di arrivo, mentre dal punto di vista informatico si tratta di dati che possono essere ulteriormente elaborati:¹¹ non solo a fini di visualizzazione e navigazione, ma per rispondere a *query* sofisticate come quelle accennate sopra.

Da questa conversazione è nata l'idea di tenere un workshop sull'argomento, coinvolgendo altri studiosi, con l'obiettivo di individuare casi d'uso e problemi comuni da risolvere, e possibili soluzioni all'interno di EVT.¹² Per soddisfare queste esigenze, infatti, si potrebbe pensare a motori di ricerca o, meglio ancora, database di tipo XML Native particolarmente sofisticati e potenti, ma questo tipo di strumenti può coprire tali necessità solo in parte e, soprattutto, dovrebbero essere configurati e calibrati ogni volta in relazione al tipo di contenuto. La proposta emersa nel corso del workshop è quella di creare uno strumento specifico integrato in EVT, altamente flessibile e configurabile, che possa essere usato in una varietà di casi molto ampia.

In questo articolo intendo proporre alcuni esempi concreti basati sulle mie ricerche riguardo testi in inglese antico.

Indagini paleografiche: Exeter Book e Vercelli Book

In un suo articolo del 1986¹³ P. Conner, studioso molto noto nell'ambito delle ricerche sulla lingua e letteratura anglosassone, ha pubblicato i risultati di una sua ricerca su un manoscritto di grande importanza per tale tradizione culturale, l'Exeter Book, redatto nel dialetto tardo sassone occidentale alla fine del X secolo. Conner ha sviluppato una teoria sulla composizione originale di questo manoscritto, a suo avviso composto da tre codicelli (*booklets*) separati, redatti in

10 L'elemento <table> può essere usato in modo da creare dei veri e propri fogli di calcolo, caratteristica utilissima al momento in cui è necessario codificare documenti con dati strutturati quali le denunce fiscali della Lira senese.

11 Su questo punto si veda [12].

12 "Fonti archivistiche medievali nel digitale. La sfida di trattare e visualizzare dati semi-strutturati", 22-23 giugno 2020, workshop tenuto online, disponibili le registrazioni degli interventi a video dei relatori e della successiva discussione: <http://www.labcd.unipi.it/fonti-archivistiche-medievali-nel-digitale/>.

13 "The Structure of the Exeter Book Codex (Exeter, Cathedral Library, MS. 3501)."

momenti diversi e successivamente riuniti a formare un unico codice senza che venisse rispettato l'ordine cronologico della loro produzione. A sostegno della propria teoria, Conner ha effettuato un'analisi delle occorrenze di un certo numero di caratteri speciali e legature usati dallo scriba che ha redatto l'Exeter Book, in modo da individuare particolarità nell'uso che potessero aiutare a raggruppare i fascicoli sulla base delle abitudini scribali. Tutti i calcoli sono stati effettuati a mano, al lettore sono presentati solo i risultati finali (Figura 1).

Quire	S+T		S+P		S+Wynn	
	<i>no lig.</i>	<i>lig.</i>	<i>no lig.</i>	<i>lig.</i>	<i>no lig.</i>	<i>lig.</i>
1	24	17	8	5	2	2
2	21	37	7	9	1	2
3	17	26	6	9	4	0
4	17	30	5	10	3	1
5	15	18	9	4	4	0
6	16	38	15	16	4	5
7	1	70	0	38	0	8
8	0	49	0	42	0	13
9	1	53	0	29	0	10
10	1	59	0	23	0	10
11	0	43	0	33	0	5
12	0	61	0	29	0	4
13	2	86	6	43	2	8
14	1	44	0	39	0	6
15	2	54	2	20	1	4
16	0	48	1	15	1	7
17	1	54	0	13	0	1

Figura 1: La tabella 3 dell'articolo di P. Conner.

Si tratta di un'analisi interessante, anche per le prospettive che apre in merito a simili indagini da portare a termine su altri manoscritti in inglese antico, ma che presenta alcuni problemi metodologici:

- è stata effettuata interamente a mano, sia per quanto riguarda il conteggio dei caratteri sia per la successiva produzione di statistiche;
- è limitata a un set di lettere e legature molto ristretto;
- in generale, non è verificabile in maniera autonoma se non ripetendo in toto o in parte il processo di conteggio manuale;

- alcune affermazioni non sono verificabili perché richiederebbero una ulteriore elaborazione dei dati che non è stata portata a termine.¹⁴

Il metodo usato da Conner, essendo interamente basato su calcoli manuali, pone una serie di questioni che vanno al di là della praticità e onerosità dello stesso: siamo sicuri che tutti i caratteri speciali individuati da Conner per calcolare le sue statistiche siano stati contati correttamente per ciascun *booklet*? è possibile che qualcuno gli sia sfuggito? o, viceversa, che qualcuno sia stato contato due volte per lo stesso foglio di pergamena? dato che il metodo utilizzato è interamente manuale, l'unico tipo di verifica può essere soltanto una ripetizione di entrambe le fasi: conteggio delle occorrenze e calcolo della loro distribuzione su base statistica. Grazie alla codifica XML/TEI, tuttavia, entrambe queste operazioni potrebbero essere automatizzate e ulteriori tipi di elaborazione e confronto dei dati potrebbero essere portati a termine. Vediamo come.

Al momento di definire in dettaglio la codifica dei testi per il progetto *Digital Vercelli Book* ([14]), trattandosi in primo luogo di una edizione diplomatica ho deciso di marcare tutti i caratteri speciali seguendo il metodo TEI così come descritto nel capitolo 5 *Characters, Glyphs, and Writing Modes* delle *Guidelines*.¹⁵ Caratteri e glifi particolari sono descritti in una sezione specifica dell'intestazione TEI (<teiHeader>), nell'elemento <charDecl> (lett. "dichiarazione dei caratteri") a sua volta strutturato in ulteriori sotto-elementi (Figura 2), dopo di che all'interno del testo si ricorre all'elemento <g> con attributo @ref per collegare ogni occorrenza al rispettivo <char> o <glyph>. Con la codifica <g ref="#amacr"/> usata nel testo, ad esempio, si fa riferimento all'elemento <char> corrispondente, identificato in base al valore amacr per l'attributo @xml:id.

L'obiettivo iniziale, tuttora valido, è quello di gestire versioni diverse dei caratteri speciali¹⁶ da inserire nei rispettivi livelli di edizione (diplomatica e interpretativa), quindi anche in questo caso l'elaborazione dei dati della codifica viene effettuata per fini di visualizzazione e ricerca:

- tutti i caratteri e glifi speciali sono mostrati nel livello di edizione diplomatica, e traslitterati in quello di edizione interpretativa, in maniera automatica;
- questo rende possibile, in combinazione con altre particolarità della marcatura,¹⁷ la visualizzazione di livelli di edizione diversi a partire da un unico documento TEI;
- grazie a una tastiera virtuale per l'inserimento nel campo di ricerca, infine, è possibile fare ricerche testuali che includano i caratteri speciali.

14 Ad esempio: "I have been able to determine no linguistic correlations with the use of these forms; there seems to be no particular word nor combination of letters where one is more likely to get one form of the Y rather than another." ([2]: 240).

15 Disponibile online: <https://tei-c.org/release/doc/tei-p5-doc/en/html/WD.html>.

16 Facendo ricorso a tanti elementi <mapping> quanti sono i glifi e le varianti desiderati (Figura 2).

17 In particolare l'uso dell'elemento <choice> per gestire coppie di elementi editoriali.

```

<char xml:id="amacr">
  <charName>LATIN SMALL LETTER A WITH MACRON</charName>
  <mapping type="codepoint">U+0101</mapping>
  <mapping type="diplomatic">ā</mapping>
  <mapping type="normalized">a</mapping>
</char>
<char xml:id="aeligmacr">
  <charName>LATIN SMALL LETTER AE WITH MACRON</charName>
  <mapping type="codepoint">U+01E3</mapping>
  <mapping type="diplomatic">ǣ</mapping>
  <mapping type="normalized">æ</mapping>
</char>
<char xml:id="AEligmacr">
  <charName>LATIN CAPITAL LETTER AE WITH MACRON</charName>
  <mapping type="codepoint">U+01E2</mapping>
  <mapping type="diplomatic">Ǽ</mapping>
  <mapping type="normalized">Æ</mapping>
</char>
<char xml:id="emacr">
  <charName>LATIN SMALL LETTER E WITH MACRON</charName>
  <mapping type="codepoint">U+0113</mapping>
  <mapping type="diplomatic">ē</mapping>
  <mapping type="normalized">e</mapping>
</char>
<char xml:id="imacr">
  <charName>LATIN SMALL LETTER I WITH MACRON</charName>
  <mapping type="codepoint">U+012B</mapping>
  <mapping type="diplomatic">ī</mapping>
  <mapping type="normalized">i</mapping>
</char>

```

Figura 2: Una piccola parte della <charDecl> del *Digital Vercelli Book*.

Una marcatura di questo tipo, tuttavia, consentirebbe allo studioso di eseguire in maniera automatica i calcoli e le statistiche effettuate manualmente da Conner. Sulla base della codifica XML/TEI si possono elaborare i dati relativi ai caratteri speciali codificati nella <charDecl> con i seguenti obiettivi:

- analisi statistica delle occorrenze di ciascun carattere;
- analisi di altri fenomeni per i quali è stata effettuata una codifica, ad es. le caratteristiche generali dei capolettera, le iniziali miniate, le legature;
- ricerca di eventuali correlazioni fra specifici caratteri;
- individuazione delle abitudini scribali, in modo da permettere la formulazione di ipotesi riguardo il rapporto fra il documento analizzato e il suo apografo.

Anche la verifica può essere automatizzata: un metodo molto semplice e diretto potrebbe consistere in uno script XSLT, o in altro linguaggio simile, in grado di effettuare un *parsing* di un documento XML al fine di individuare ogni carattere speciale inserito direttamente nel testo

e non per mezzo di un elemento <g>. In questo modo si potrebbe essere sicuri al 100% di non aver dimenticato di marcare nessuno dei caratteri desiderati nel testo della trascrizione.

Resta il dubbio sull'accuratezza del conteggio, che nel nostro caso riguarda la trascrizione e la conseguente marcatura: in questo caso è possibile una verifica indiretta, in quanto la visualizzazione dei caratteri speciali, in particolare quando sono disponibili a lato le scansioni del manoscritto, permette di notare immediatamente quelli che sono resi in maniera diversa da quanto prescritto dai rispettivi <char> e <glyph>. Nella fase di controllo della trascrizione, imprescindibile per ogni progetto di questo tipo, tali mancanze risulterebbero immediatamente evidenti.

Sul piano metodologico, infine, si conferma il vantaggio della DSE come strumento di ricerca sia per quanto riguarda la preservazione e condivisione di quel processo di analisi dei dati che in un'edizione tradizionale normalmente andrebbe perduto in maniera irrimediabile, sia pure per la possibilità di verifica e ripetizione delle operazioni su tali dati che, di nuovo, è del tutto impossibile in una pubblicazione a stampa. La documentazione del metodo di lavoro e la sua inclusione nell'edizione, gli strumenti di verifica dei risultati presentati all'utente finale dall'editore (anche sul piano più strettamente filologico), la condivisione degli stessi dati dell'edizione con la comunità accademica al fine di un loro riuso,¹⁸ costituiscono vantaggi significativi e uno sviluppo che avvicina gli studi umanistici al metodo scientifico propriamente detto, pur con tutte le precauzioni relative a discipline che non sono certo "scienze dure", rispetto alla prassi tradizionale.

Named entities e altri elementi di interesse nella Cronaca Anglosassone

Molte DSE marcano le *named entities* presenti nel testo, alcune anche i *realia*¹⁹ e altri oggetti d'interesse, grazie agli strumenti di codifica messi a disposizione dagli schemi TEI: le informazioni riguardo ogni entità, nominata o meno, sono raccolte in una delle liste dedicate (<listPerson>, <listPlace>, etc.) o in una lista generica <list> (facilmente tipizzabile per mezzo dell'attributo @type) usando gli elementi specifici di ogni lista (<person>, <place>, etc.; semplicemente <item> per la lista generica), dopo di che all'interno del testo si rimanda a tali elementi per mezzo dell'attributo @ref, come abbiamo visto per i caratteri speciali. Tale attributo è disponibile per gli elementi che permettono di marcare le singole occorrenze di un nome

18 Parallelamente all'adozione del metodo *open source* per quanto riguarda lo sviluppo di strumenti software nell'ambito delle Digital Humanities, infatti, si va sempre più affermando il principio dell'*open data* che riguarda appunto la condivisione delle varie componenti (testo, paratesto, immagini, dataset di altro tipo) di un'edizione digitale.

19 Si veda, ad esempio, il progetto di edizione di *Dei viaggi di Messer Marco Polo, gentiluomo veneziano* a cura di Marina Buzzoni, Eugenio Burgio e altri studiosi dell'Università Ca' Foscari Venezia (<http://virgo.unive.it/ecf-workflow/books/Ramusio/main/index.html>).

(<persName>, <placeName>, etc.) o di altre entità (ad esempio monete, cariche etc.), in questo secondo caso vanno bene anche elementi generici come <rs> e <seg> purché dotati di attributi di *linking* come @ref o @target. Questo meccanismo è semplice e potente, e permette di trasformare i collegamenti interni a un documento TEI²⁰ in link ipertestuali all'interno di una edizione digitale. Negli esempi che seguono (Figura 3 e Figura 4) gli *screenshot* riguardano l'edizione del *Codice Pelavicino Digitale* perché è probabilmente l'edizione basata su EVT più ricca e dettagliata per quanto riguarda le *named entities*,²¹ ma è evidente che l'applicazione di questo metodo a un testo come la *Cronaca Anglosassone*, opera ricchissima di riferimenti concreti a personaggi storici, luoghi, etc., potrebbe essere assai produttivo ai fini di una edizione digitale.

```

<!-- ##### Lista Persone ##### -->
<listPerson>
  <head>Lista Persone</head>
  <!-- ##### A ##### -->
  <person xml:id="AbraçaleneBarbazano">
    <persName>
      <forename>Abraçalene</forename>
      <surname>de Barbazano</surname>
    </persName>
    <sex>M</sex>
  </person>
  <person xml:id="AbraciatorArmannini">
    <persName>
      <forename>Abraciator</forename>
      <surname>quondam Armannini</surname>
    </persName>
    <sex>M</sex>
  </person>
  <person xml:id="AbraciatiusPonzano">
    <persName>
      <forename>Abraciatius</forename>
      <surname>de Ponzano</surname>
    </persName>
    <sex>M</sex>
  </person>
  <person xml:id="Acaritus">
    <persName>
      <forename>Acaritus</forename>
    </persName>
    <sex>M</sex>
  </person>

```

Figura 3: Codifica delle named entities nel Codice Pelavicino Digitale (<listPerson>).

20 Se si desidera utilizzare una collezione di liste per più documenti TEI è possibile spostare questa parte della codifica in un file esterno al quale tali documenti possono accedere in maniera indipendente.

21 Nel *Digital Vercelli Book* sono state introdotte alcune named entities a livello sperimentale per testare il meccanismo, ma la natura stessa dei testi (23 omelie e 6 componimenti poetici) si presta solo in parte a una marcatura di questo tipo.

```

<body>
  <div xml:id="CI_div">
    <pb n="214v" xml:id="fol_214v"/>
    <p xml:id="CI_p_001" n="1">De <placeName ref="#Sarzanello">castro
      Sarzane</placeName><note place="foot" n="a" xml:id="CI_n_001"><term
        xml:lang="latin">De castro Sarzane</term>
      <emph>in inchiostro rosso.</emph></note>.</p>
    <p xml:id="CI_p_002" n="2"><ptr target="#CI_st_001" facs="#st_214v_001"/> In eterni
      Dei nomine, amen. Quemadmodum in instrumento publico manu <persName
        ref="#Bartholomeusnot">Bartholomei <roleName>notarii</roleName></persName>
      confecto, quod ego <persName ref="#Confortusnot">Confortus
        <roleName>notarius</roleName></persName> vidi et legi, continebatur,
      dominus <persName ref="#PipEpLuni">Pipinus, Dei gratia olim <roleName>Lunensis
        episcopus</roleName></persName>, renovellavit, fecit et firmavit per se
      et successores suos, consilio curie, tale pactum et talem constitutionem quale
      et qualem fecerat dominus <persName ref="#AlbertusEpluni">Albertus, bone memorie
        Lunensis <roleName>episcopus</roleName></persName>, predecessor suus, cum
      operariis omnibus de <placeName ref="#Sarzana">curte Sarzanei <note place="foot"
        n="b" xml:id="CI_n_002"><term type="incerto">Sarzanei</term>: <emph>cosi
        nel testo.</emph></note></placeName>. Uterque enim eorum, velut in
      eodem instrumento legebatur, condonavit et remisit per se et suos successores
      omnibus suprascriptis operariis dona et opera atque quod nullus eorum cogatur ad
      canevas recipiendam nec ad castaldiam neque ad iscariam, nisi per voluntatem,
      excepto quod omni anno debet unus eorum esse <roleName>iscarius</roleName>, si
      fuerit voluntas episcopi, et debet suum feudum habere. Placita, districta,
  
```

Figura 4: Codifica delle named entities nel *Codice Pelavicino Digitale*: uso di @ref nel testo.

EVT e altri programmi per la navigazione di edizioni digitali permettono di gestire questi dati come segue:

- visualizzazione delle informazioni relativa a ciascuna *named entity* per ogni occorrenza nel testo;
- visualizzazione delle liste dove sono raccolte le informazioni;
- accesso diretto alle occorrenze dalle liste.

Si tratta quindi di una elaborazione, ma limitata: i dati sono gestiti per inserire le informazioni relative allo stesso individuo (o luogo, oppure organizzazione ecclesiastica, etc.) in maniera che siano disponibili per ogni occorrenza dello stesso su richiesta dell'utente (Figura 5).

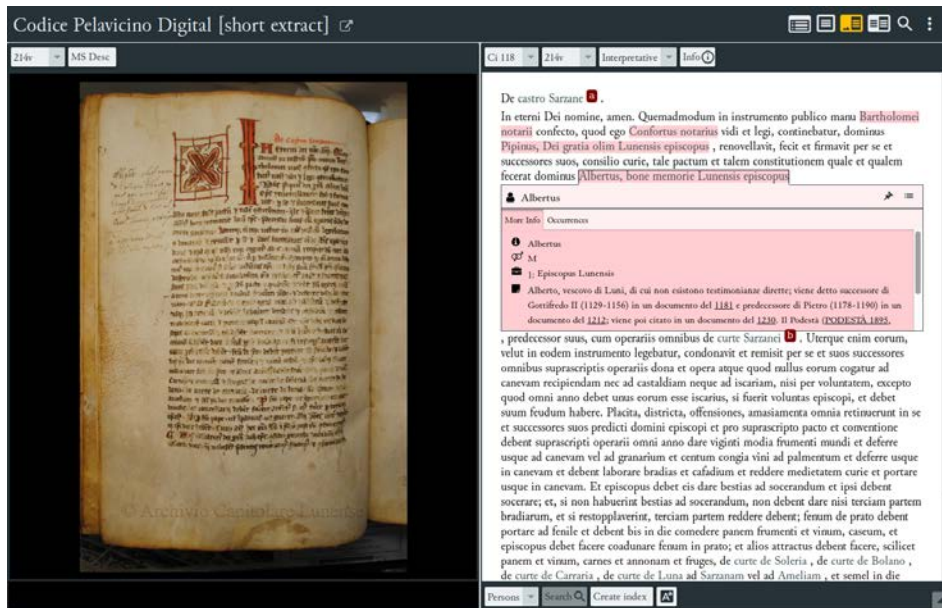


Figura 5: Visualizzazione delle informazioni nel testo per Alberto, vescovo di Luni.

È inoltre possibile scorrere direttamente le liste in modo da reperire quelle stesse informazioni e, se così desiderato, spostarsi rapidamente nel punto del testo in cui sono presenti le relative occorrenze (Figura 6).



Figura 6: La voce relativa ad Alberto, vescovo di Luni, nella lista di persone menzionate nel CPD.

Una possibile evoluzione di questo metodo è il ricorso ai LOD (Linked Open Data), risorse condivise pubblicamente sul web, già messo in pratica da alcune DSE.²² Il limite di questo approccio è che migliora le due caratteristiche fondamentali di una DSE pubblicata con uno strumento come EVT:

- la *visualizzazione* di informazioni, estesa grazie al collegamento con risorse disponibili sul web;
- la *selezione e navigazione* di informazioni, in parte disponibili *in situ*, in parte distribuite sul web.

Si tratta sicuramente una prospettiva molto più ampia rispetto alla singola DSE, un arricchimento importante sul piano dei contenuti e della capacità di dialogare con altre fonti e risorse condivise sul web, ma a livello concettuale non cambierebbe quasi niente.

Un ulteriore sviluppo sarebbe invece la possibilità di annotare le relazioni fra le varie entità, a quel punto sarebbe possibile aggiungere un livello di ricerca e interrogazione dei dati del tutto nuovo. Supponendo di avere annotato con XML/TEI uno o più testimoni della *Cronaca Anglosassone*, ad esempio, sarebbe uno strumento di analisi formidabile una funzionalità di interrogazione del testo in maniera tale da ottenere una risposta alle seguenti domande:

- con quali altri personaggi storici ha interagito re Alfredo il Grande nel periodo che va dall'871 all'899?
- in quali eventi è stato coinvolto re Edwin di Northumbria?
- l'area geografica del regno del Wessex che fenomeni storici (battaglie, visite pastorali, etc.) ha conosciuto nel X secolo?
- etc.

La ricchezza di informazioni della *Cronaca Anglosassone* è tale, del resto, che sin dalla prima edizione a stampa del 1692 a cura di E. Gibson sono stati inseriti in appendice degli indici dei nomi di persona e dei luoghi, in modo da poter rintracciare rapidamente le loro occorrenze nel testo (Figura 7).²³ La gestione delle *named entities* con il metodo sopra descritto rappresenterebbe già un significativo passo in avanti rispetto agli indici a stampa, soprattutto considerando che la generazione delle liste e il loro collegamento con le occorrenze del testo sarebbe del tutto automatico. Una funzionalità che permetta di collegare fra loro le singole entità e altri oggetti d'interesse, viceversa, metterebbe a disposizione dello studioso uno strumento molto più potente.

22 Si veda in proposito [8].

23 Cfr. anche l'edizione di Thorpe ([19]) e, in tempi decisamente più recenti, i volumi de *The Anglo-Saxon Chronicle. A collaborative edition*, general editors David Dumville and Simon Keynes, 1983-.

N O M I N U M L O C O R U M
 Q U Æ
 I N C H R O N I C O S A X O N I C O m e m o r a n t u r ,
 E X P L I C A T I O .

Literæ, quæ in hoc Glossario Historicorum nomina designant.

<p>B. Bedæ Al. Alfrædus M Ac. Alfrædus Et. Ethelwærdus I. Ingulphus F. Florentius M. Will. Malmesburienf. Hunt. Henr. Huntingdonienf. D. Sin. Duxelmensis E. Ethelredus, Abbas Riccallis</p>	<p style="text-align: center;">1090 } clauit } circa } Ann.</p> <p style="text-align: center;">720 890 900 1109 1114 1140 1148 1164 1166</p>	<p>H. Hagustaldensis Prior, Joannes R. Ricardus, Prior Hagustald. G. Gervasius Hovd. Hordenus B. Bromtonus S. Stubbes W. Matt. Westmonasterienf. T. Thorn. K. Knibton</p>	<p style="text-align: center;">1180 1184 1200 1204 1330 1360 1377 1380 1395</p>
---	---	---	---

A Bban-dun. Abben-dun. p. 124, l. 9. 125, 31. 150, 38. 151, 22. 157, 20, 34. 158, 29. 161, 32, 33. 181, 20. 220, 29. i. e. Abbatia mons sive oppidum. Abbandonia, F. Abbandune, E. Abbingdon, B. Abendon, Abyndon, K. Nomen hoc paulatim induit, postquam Cissa, Occidentalium Saxonum Rex, Monasterium sive Abbatiam ibi construxerat. Quo autem nomine antea vocabatur, non ita liquidò constat.

tur. Æthelwulfi prælium cum Danis, anno dccclii. memoratum, non alibi commissum fuit, quam ubi hodie conspicitur Drelep, in agri Sutbreiensis parte Sussexiensis contermina; alibi certe Concilia annis dccclxxxii. & dccclxxxix. celebrabantur. Nec tamen scio qua in parte Angliæ locus iste sit querendus. Missò enim quòd in singulis penè gentis hujus Provinciis istius appellationis loca reperiuntur; desiderantur ea omnia quibus ad

Figura 7: L'indice dei nomi di luogo nell'edizione di Gibson della *Cronaca Anglosassone*.

Questo tipo di funzionalità ha dunque un terreno di applicazione naturale nel caso di documenti e testi di ambito storico, ma potrebbe essere usato con profitto anche nel caso di testi letterari, ad esempio testi poetici della tradizione epico-eroica germanica medievale, come il *Beowulf* o le saghe in norreno: la ricchezza di riferimenti a personaggi, reali o mitologici, luoghi ed eventi invoca a gran voce una marcatura semantica effettuata con rigore e in maniera sistematica, in modo da preparare il terreno a una successiva esplorazione di tutte le relazioni che verrebbero determinate automaticamente.

Gli strumenti di marcatura utilizzabili sono già disponibili grazie agli schemi TEI: gli elementi per la codifica di *named entities* menzionati sopra (<person>, <place>, <org> etc.) in connessione con elementi di tipo <event>, organizzati in una <listEvent>, in modo da stabilire relazioni fra le entità codificate nel testo in base sia agli eventi che le riguardano, sia ai luoghi in cui questi si sono verificati, sia anche alle date in cui tali eventi hanno avuto luogo (elemento <date>).²⁴

²⁴ Un'altra possibile applicazione per questo tipo di elementi è infatti la creazione di timeline con relativa geolocalizzazione degli eventi descritti nel testo.

Ontologie basate su codifica TEI

La metodologia necessaria per implementare questa funzionalità nella maniera più efficace, tuttavia, richiederebbe, se non una *full fledged ontology*, per lo meno la possibilità di definire delle triple in maniera simile a quanto è possibile fare usando il linguaggio RDF.²⁵ La TEI offre già gli elementi fondamentali per descrivere le entità di una ontologia basilare, infatti oltre a `<listPerson>` e altre liste per *named entities* possiamo contare sull'elemento `<relation>`, con una dotazione di attributi sufficiente per stabilire relazioni fra oggetti diversi sotto forma di triple, e la relativa `<listRelation>`, tuttavia il supporto alla creazione di triple non è del tutto soddisfacente.

La tripla "Re Alfredo ha sconfitto Guthrum a Edington nell'878" potrebbe essere codificata così:

```
<relation name="defeatInBattle" active="#kingalfred" passive="#guthrum"
when="0878"/>
```

dove si nota che nell'elemento `<relation>` abbiamo un soggetto, identificato con `@active`, che compie un'azione, espressa con `@name`, che riguarda un oggetto, a sua volta identificato con `@passive`. In questo modo è possibile esprimere la tripla `King alfred - defeated - Guthrum` arricchita da ulteriori informazioni sulla data (`@when`) e volendo dal collegamento a una descrizione più dettagliata dell'evento in questione (cfr. *infra* l'uso di `@corresp` e `<event>`).

Come si può vedere da questo semplice esempio, per quanto tale codifica possa essere inserita nel meccanismo di collegamento prospettato sopra per le *named entities*,²⁶ risulta piuttosto faticosa da gestire in un quadro chiaro dal punto di vista metodologico: da un lato, infatti, rappresenta un'efficace estensione del meccanismo di linking fra oggetti diversi (`<person>`, `<place>`, etc.) al fine di rappresentare relazioni dirette fra gli stessi; dall'altro, invece, si sovrappone all'uso di `<event>` per quanto riguarda gli avvenimenti; gli attributi, infine, non solo non sono esattamente intuitivi, ma sono anche molto limitati se confrontati con la grande potenzialità espressiva del linguaggio RDF. Una caratteristica che aumenta di molto la flessibilità di questo markup, viceversa, è la possibilità di usare l'attributo `@name` premettendo al contenuto l'etichetta di un *namespace* che appartiene a una ontologia specifica per un determinato ambito, ad esempio in questo elemento `<relation>`

```
<relation name="saws:formsPartOf" active="#nowellcodex"
passive="#cottonvitelliusAXV"/>
```

si dichiara che il Nowell Codex, il manoscritto che conserva l'unica copia esistente del *Beowulf*, fa parte del codice noto come Cotton Vitellius A. XV²⁷ grazie alla relazione `formsPartOf` che

25 Sull'opportunità di usare direttamente RDF con i documenti TEI si veda [20].

26 I valori `#kingalfred` e `#guthrum` rimandano ad altrettanti elementi `<person>` che hanno `kingalfred` e `guthrum` come valori dell'attributo `@xml:id`.

27 Questo manoscritto composito è probabilmente opera dello stesso Sir Robert Cotton (1571-1631).

è stata definita all'interno dell'ontologia SAWS.²⁸ In questo modo si può far ricorso a ontologie esistenti, già ben strutturate e testate, senza che sia necessario reinventare la ruota.

In ogni caso, se `<event>` rappresenta una parziale duplicazione dal punto di vista del markup, permette anche di espandere il contenuto con molti più dettagli rispetto a quello che si potrebbe fare in una `<relation>`, elemento basato sul solo uso di attributi. L'ipotesi a questo punto è quella di gestire non solo i vari oggetti (`<person>`, `<place>`, etc.), le rispettive liste e le relazioni fra di essi, ma anche il collegamento fra almeno alcuni specifici tipi di `<relation>` con elementi di tipo `<event>`, in modo da garantire una descrizione accurata di avvenimenti storici e altri fenomeni descritti nel testo. Questo collegamento è facilmente ottenibile facendo uso dell'attributo `@corresp` all'interno di `<relation>`:

```
<relation      corresp="#battle_of_Edington"      name="defeatInBattle"
active="#kingalfred" passive="#guthrum" when="0878"/>
```

mentre altrove nel documento avremo una `<listEvent>` che contiene un `<event>` relativo alla battaglia di Edington:

```
<listEvent>
  <event source="https://en.wikipedia.org/wiki/Battle_of_Edington" xml:id="battle_of_edington">
    <head>The Battle of Edington</head>
    <p>At the Battle of Edington, an army of the Anglo-Saxon kingdom of <placeName ref="#wessex">
      Wessex</placeName> under <persName ref="#kingalfred">Alfred <addName>the Great</addName></persName>
      defeated the Great Heathen Army led by the Dane <persName ref="#guthrum">Guthrum</persName> on a date
      between <date from="0878-05-06" to="0878-05-12">6 and 12 May 878</date>, resulting in the <rs
      ref="#treaty_of_wedmore">Treaty of Wedmore</rs> later the same year. Primary sources locate the battle
      at "Eðandun". Until a scholarly consensus linked the battle site with the present-day town of Edington in
      Wiltshire, it was known as the Battle of Ethandun. This name continues to be used.</p>
  </event>
</listEvent>
```

Figura 8: La voce relativa alla battaglia di Edington nel documento TEI.

Il prezzo da pagare è una certa complessità del markup, ma il risultato finale è comunque soddisfacente per casi d'uso non particolarmente complessi.

I limiti degli schemi TEI nell'ambito delle ontologie,²⁹ tuttavia, sono evidenti e tali da rendere alquanto complicata la definizione di ontologie potenti e ben strutturate, motivo per cui una prima richiesta di miglioramento risale già al 2011,³⁰ a testimonianza del fatto che una

²⁸ Come evidenziato dal prefisso `saws`: che dichiara il namespace di appartenenza della relazione `formsPartOf`. Progetto *SAWS: Sharing Ancient Wisdoms*, home page: <https://ancientwisdoms.ac.uk/>; riguardo l'ontologia SAWS si veda la pagina <https://ancientwisdoms.ac.uk/method/ontology/index.html>

²⁹ Non sarà sfuggito al lettore, ad esempio, il fatto che `<relation>` offre molti attributi per indicare *quando* un evento si è verificato (infatti appartiene a ben quattro classi di attributi create a tale scopo: `att.dateable`, `att.dateable.w3c`, `att.dateable.iso` e `att.dateable.custom`), ma nessuno per quanto riguarda il *dove*. Dal punto di vista semantico usare a tal fine l'unico elemento consentito al suo interno, ovvero `<desc>`, non sembra corretto.

³⁰ Issue 311: <https://github.com/TEIC/TEI/issues/311>.

implementazione più potente e flessibile di questi strumenti di codifica rappresenta un'esigenza condivisa all'interno della comunità TEI. La discussione sulla strategia da adottare per raggiungere questo obiettivo è tutt'ora in corso,³¹ al momento la soluzione sperimentale proposta dagli editor TEI prevede l'integrazione diretta degli attributi RDFa³² negli schemi TEI, aggiunta che rappresenterebbe un salto di qualità notevole per quanto riguarda la creazione di ontologie in un documento TEI senza ricorrere direttamente all'uso di RDF.

Una volta codificata all'interno di un documento TEI, l'annotazione delle relazioni per mezzo di triple dovrebbe essere successivamente elaborata dal software di visualizzazione per poter rispondere alle *query* dell'utente. Anche in questo caso si tratta di porre le fondamenta per un'analisi successiva, ma un aspetto molto importante da prendere in considerazione è la progettazione di un'interfaccia utente adeguata, sia per quanto riguarda l'inserimento delle *query*, sia per la presentazione dei risultati. Se esaminiamo il sito *PASE (Prosopography of Anglo-Saxon England)*,³³ infatti, possiamo notare che l'interfaccia utente di questo database di informazioni prosopografiche e relazioni fra individui, con riferimenti ai testi del periodo, è alquanto farraginoso. Si tratta indubbiamente di uno strumento molto potente: la ricerca delle informazioni desiderate viene effettuata attraverso *query* strutturate, oppure navigando direttamente sul sito, con la possibilità di seguire link ipertestuali fra oggetti diversi (Figura 9). Peccato che i contenuti non siano resi accessibili come LOD o attraverso una API adeguata.

31 Issue 1860: <https://github.com/TEIC/TEI/issues/1860>. L'ultimo messaggio visibile al momento in cui è stato redatto il presente articolo recita "The current draft of the customisation ODD is added in the branch issue-1860 at 151136c. It simply adds all RDFa attributes to a new class `att.global.analytic.rdfa` and hooks this class into `att.global.analytic`." Purtroppo non ci sono ulteriori aggiornamenti su tale sperimentazione.

32 "RDFa is a way of expressing RDF-style relationships using simple attributes in existing markup languages such as HTML." RDFa Core 1.1, <https://www.w3.org/TR/rdfa-core/>.

33 Home page: <http://www.pase.ac.uk/>.

The screenshot shows the PASE website interface. At the top, there is a navigation menu with links for Home, About PASE, Database, Domesday, Reference, Contact, and Help. Below the navigation, there is a search bar and a 'Reset Constraints' button. The main content area is divided into three columns:

- Left Column:** A search filter section with options for 'Persons' and 'Factoids'. It includes checkboxes for 'Include: coin data' and 'Domesday data'. Below this are several filter categories: 'Contains text', 'Persons', 'Source', 'Possession', 'Location', 'Event', and 'Date', each with a radio button.
- Middle Column:** A 'Criterion' section with a 'Reset Constraints' button. Below it is a 'Results' section showing a list of names and their reigns, including CEAWLIN 1, EGBBERT 10, EDWIN 2, OSWALD 1, OSWU 1, REDWALD 1, ÆLLE 1, and ÆTHELBERHT 3.
- Right Column:** A detailed profile for 'Edwin 2 (Male)', King of the Northumbrians, 616-633. It includes a 'FACTOID LIST' with various categories and counts, such as 'Recorded Name (55)', 'Personal Information (16)', 'Education (2)', 'Office (54)', 'Status (6)', 'Personal Relationship (70)', and 'Event (73)'. Each category is followed by a list of specific events or roles.

Figura 9: Visualizzazione delle informazioni relative a re Edwin sul sito PASE.

Come osservato in merito alle indagini paleografiche, l’implementazione di queste strategie di codifica rappresenta senza dubbio un arricchimento di una DSE e l’estensione delle sue capacità in quanto strumento di ricerca. Anche in questo caso è quanto meno auspicabile che l’editore metta a disposizione della comunità accademica i dati codificati nel formato XML/TEI, in modo che possano essere riutilizzati per ulteriori analisi o per formare la base di nuove edizioni digitali. Nella prospettiva di un ecosistema digitale sempre più aperto e interconnesso, basato su risorse pubblicamente accessibili come i Linked Open Data e le numerose ontologie offerte agli studiosi per un uso del tutto libero, avrebbe poco senso curare un proprio orticello chiuso a ogni sguardo esterno.

Conclusioni

Una delle tendenze più recenti è quella dell’edizione distribuita, cui si accennava nell’*Introduzione* al presente articolo: si tratta di uno sviluppo importante, che considero fondamentale per l’evoluzione della DSE 2.0,³⁴ in quanto permette di stabilire relazioni fra

34 Si veda a tal proposito l’intervento “Nuove prospettive per la filologia digitale: la sfida delle edizioni distribuite” tenuto in occasione del convegno *Cicero Digitalis. Cicero and Roman Thought in the Age of Digital Humanities*, International Conference (Turin-Vercelli 25-26 February 2021), home page: <https://cicerodigitalis.uniupo.it/>.

oggetti di varia natura, testuali e non, che sono condivisi in modi diversi sulla rete, in modo da realizzare edizioni e/o risorse digitali altrimenti impensabili. Come si è osservato in relazione ai LOD, tuttavia, si aumentano le informazioni proposte all'utente senza che ci sia un progresso paragonabile dal punto di vista ermeneutico: è necessario pensare a un incremento della conoscenza basato non solo su un ulteriore accumulo di informazioni da proporre al lettore, ma sulla capacità di elaborare i dati esistenti in maniera flessibile e potente.³⁵ L'obiettivo è dunque quello di affiancare alla tradizionale ermeneutica delle edizioni, digitali e non, nuovi metodi euristici che possano portare a un incremento della conoscenza dei testi pubblicati sotto forma di DSE.

Gli ambiti di ricerca all'interno dei quali può risultare preziosa una ulteriore elaborazione dei dati dell'edizione, come si può arguire dai casi d'uso brevemente presentati sopra, sono molto numerosi. Altri possibili esempi in cui una elaborazione dei dati XML/TEI potrebbe portare a miglioramenti significativi nella capacità di analisi di un testo sono i seguenti:

- metrica: nelle lingue germaniche antiche, ad esempio, il verso allitterativo segue schemi fissi (i tipi codificati da Sievers, cfr. [17]) che in tempi più recenti sono stati messi in discussione ([7]) e che comunque sarebbe interessante investigare in maniera sistematica e verificabile; altre *query* potrebbero riguardare, ad esempio, la presenza o meno di doppie allitterazioni, il tipo di allitterazioni nei versi ipermetrici e la distribuzione di formule nel primo e nel secondo semiverso;
- analisi lessicale: dipende strettamente dal tipo di informazioni codificate, ma non è difficile immaginare un'analisi dei campi semantici, delle co-occorrenze, o di altri tipi di relazione fra lessemi diversi scelti in base alle caratteristiche del testo;
- analisi stilistica: qui tipicamente il metodo di codifica è affidato al singolo studioso, di conseguenza anche capacità di interrogazione sono limitate e disomogenee; sarebbe invece interessante offrire un meccanismo generale che, basandosi su un impianto generale simile a quanto visto per le *named entities* e i *realia*, permetta di indagare figure retoriche, formule, temi e altre caratteristiche di componimenti poetici; interrogazioni possibili: quali sono le *kenningar* più frequenti nei poemi anglosassoni? in quali *fitt* (sezioni) del *Beowulf* sono usate formule che indicano il mare e la navigazione? in quali componimenti si rintraccia il tema delle 'fiere della battaglia'?
- apparato critico e varianti: anche qui abbiamo una ricchezza di informazioni, soprattutto se si specificano una tipologia (attributo *@type*) e le possibili cause delle varianti (attributo *@cause*), che potrebbe costituire la base per un approccio olistico

35 "The scholarly edition has traditionally been conceived of as hierarchically ordered downwards from a text, buffered and augmented by apparatuses as subordinate editorial paratexts. Of old, the paratexts used to stand in a hermeneutic relationship – broadly, a commentary relationship – to the edition text. Increasingly, however, the hermeneutic dimension of the scholarly edition gave way to modes of positivist accumulation of materials, in support not so much of the interpretive reading but of the editorial establishing of the edited texts." [5]: 43.

all'apparato critico, con la formulazione di domande (e risposte) che vedano incrociate e collegate le diverse tipologie di fenomeni dovuti alla trasmissione del testo.

Una tale varietà di casi d'uso impone accuratezza nella progettazione e implementazione degli strumenti necessari per offrire le funzionalità desiderate. In particolare, è necessario intervenire su due livelli: in primo luogo, quello degli strumenti di codifica, intesi come strategie di markup,³⁶ che costituiscono il fondamento nella preparazione dei dati da elaborare; sia, altrettanto importante, su quello della modalità di programmazione di uno strumento per l'elaborazione dei dati testuali di un'edizione digitale. I programmi come EVT devono diventare più "intelligenti" per soddisfare meglio le esigenze di chi interroga i testi, e questo obiettivo non è facile da raggiungere perché le variabili in gioco sono molte.

Al momento in cui è stato redatto il presente articolo è in corso la progettazione per EVT di una libreria JavaScript che contenga tutte le funzioni necessarie per portare a termine tale elaborazione. La scelta di implementare questa funzionalità sotto forma di libreria esterna a EVT ci è sembrata la migliore anche nell'ottica di condividere il codice con altri progetti.³⁷

Il lavoro per la creazione di tale libreria è cominciato all'inizio di marzo 2021, non mancheremo di mettere la comunità accademica al corrente degli sviluppi.

References

- [1] *Codice Pelavicino. Edizione digitale*, a cura di E. Salvatori, E. Riccardini, R. Rosselli del Turco, L. Balletto, C. Alzetta, C. Di Pietro, C. Mannari, R. Masotti, A. Miaschi, 2a ed., 2020. URL: <http://pelavicino.labcd.unipi.it>. ISBN 978-88-944430-2-8 DOI: <https://doi.org/10.13131/978-88-944430-2-8>.
- [2] Conner, Patrick W. 1986. "The Structure of the Exeter Book Codex (Exeter, Cathedral Library, MS. 3501)." *Scriptorium*, vol. 40, no. 2: 233–42.
- [3] Delle Donne, Fulvio (ed.). 2020. *Petrus de Ebulo. De rebus Siculis Carmen*. Potenza: BUP - Basilicata University Press (*Digital Humanities* 1). ISBN: 978-88-31309-02-8. Versione online: <http://web.unibas.it/bup/evt2/pde/index.html>.

36 Il dato di partenza, come menzionato nell'*Introduzione* al presente articolo, è il fatto che i documenti TEI contengono già molte informazioni che tuttavia sono sotto-utilizzate, inoltre aggiungendo un layer di markup mirato si potrebbero ottenere risultati ancora migliori da uno strumento di elaborazione dei dati XML.

37 Una libreria indipendente dal codice sorgente di EVT può essere impiegata, direttamente o in un applicativo di tipo stand-alone, all'interno del sistema interattivo *Engineering Historical Memory* (<http://www.engineeringhistoricalmemory.com>) creato da Andrea Nanetti, si veda in proposito l'articolo di A. Nanetti e D. Benvenuti "Engineering Historical Memory and the Interactive Exploration of Archival Documents" in questo numero di *Umanistica Digitale*.

- [4] Dumville, David, Simon Keynes, Janet B. Bately, Patrick W. Conner, Simon Taylor, Michael Lapidge, G. P. Cubbin, Peter S. Baker, O’Keeffe, and Susan Irvine. 1983. *The Anglo-Saxon chronicle: a collaborative edition*. Cambridge: D.S. Brewer.
- [5] Gabler, Hans Walter. 2010. “Theorizing the Digital Scholarly Edition”. *Literature Compass* 7 (2): 43–56. <https://onlinelibrary.wiley.com/doi/10.1111/j.1741-4113.2009.00675.x>
- [6] Gibson, Edmund. 1692. *Chronicon Saxonicum, Seu Annales Rerum in Anglia Præcique Gestarum, a Christo Nato Ad Annum Usque MCLIV. Deducti, Ac Jam Demum Latinitate Donati. Cum Indice Rerum Chronologico. Accedunt Regule Ad Investigandas Nominum Locorum Origines. Et Nominum Locorum Ac Virorum in Chronico Memoratorum Explicatio*. Oxford: Theatro Sheldoniano.
<https://archive.org/details/chroniconsaxonic00gibs>.
- [7] Hoover, David L. 1985. *A New Theory of Old English Meter*. New York: Peter Lang.
- [8] Monella, Paolo, e Rosselli Del Turco, Roberto. 2020. “Extending the DSE: LOD support and TEI/IIIF integration in EVT”, in Marras, Cristina; Passarotti, Marco; Franzini, Greta; Litta, Eleonora (eds.), *Atti del IX Convegno Annuale AIUCD. La svolta inevitabile: sfide e prospettive per l’Informatica Umanistica*, Università Cattolica del Sacro Cuore, Milano 2020, pp. 148-155 (numero speciale di *Umanistica Digitale*). ISBN: 978-88-942535-4-2; DOI: [10.6092/unibo/amsacta/6316](https://doi.org/10.6092/unibo/amsacta/6316); URL: <http://amsacta.unibo.it/6316/>.
- [9] Muir, Bernard J. 2006. *The Electronic Exeter Anthology of Old English Poetry: An Edition of Exeter Dean and Chapter MS 3501*. DVD. Exeter: University of Exeter Press.
- [10] Nanetti, Andrea. *About Engineering Historical Memory*. Retrieved from <https://engineeringhistoricalmemory.com/About.php>.
- [11] O’Donnell, Daniel Paul, Singh, Gurpreet, Porter, Dot, Rosselli Del Turco, Roberto, Callieri, Marco, Dellepiane, Matteo, & Scopigno, Roberto. 2018. “Publishing (and Forgetting) the Small or Medium-sized Scholarly Edition or Cultural Heritage Collection as Linked Open Data: Using Zenodo and Github to Publish the Visionary Cross Project”. Abstract: <https://zenodo.org/record/3338482>. Presentation: <https://zenodo.org/record/3338458>.
- [12] O’Donnell, Daniel Paul, & Rosselli Del Turco, Roberto. 2020. “Good things come in small packages: designing distributed editions and tools for the age of FAIR data (Version Delivered)”. Presented at the *Assemblée Générale du consortium Cahier 2020*, France. Zenodo: <http://doi.org/10.5281/zenodo.4293723>.
- [13] PASE: Prosopography of Anglo-Saxon England. <http://www.pase.ac.uk/>.
- [14] Rosselli Del Turco, Roberto (ed.). 2017. *The Digital Vercelli Book. A facsimile edition of Vercelli, Biblioteca Capitolare, CXVII*. Transcription and encoding by Roberto Rosselli Del Turco, Raffaele Cioffi, Federica Gorla. EVT software created by Chiara Di Pietro, Julia Kenny, Raffaele Masotti, Roberto Rosselli Del Turco. *Collane@unito.it*. URL: <https://www.collane.unito.it/oa/items/show/11>. ISBN 9788875901073.

- [15] Rosselli Del Turco, Roberto. 2019. “Designing an advanced software tool for Digital Scholarly Editions: The inception and development of EVT (Edition Visualization Technology)”. *Textual Cultures* 12.2: 91–111. URL: <https://scholarworks.iu.edu/journals/index.php/textual/article/view/27690>; DOI: <https://doi.org/10.14434/textual.v12i2.27690>.
- [16] SAWS: Sharing Ancient Wisdoms. <https://ancientwisdoms.ac.uk/>.
- [17] Sievers, Eduard. 1893. *Altgermanische Metrik*. Halle: Max Niemeyer.
- [18] TEI Consortium, eds. 2021. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 4.2.1. Last updated on 1st March 2021. TEI Consortium. <http://www.tei-c.org/Guidelines/P5/>.
- [19] Thorpe, Benjamin. 1861. *The Anglo-Saxon Chronicle, According to the Several Original Authorities*. London : Longman, Green, Longman, and Roberts. <http://archive.org/details/anglosaxonchroni02thoruoft>.
- [20] Vogeler, Georg. 2019. “The ‘Assertive Edition’: On the Consequences of Digital Methods in Scholarly Editing for Historians”. *International Journal of Digital Humanities* 1, n. 2 (luglio 2019): 309–22. DOI: <https://doi.org/10.1007/s42803-019-00025-5>.
- [21] World Wide Web Consortium (W3C), eds. 2015. *RDFa Core 1.1 - Third Edition. Syntax and processing rules for embedding RDF through attributes*. W3C Recommendation, 17 March 2015. <https://www.w3.org/TR/rdfa-core/>

Last access URLs: 18/03/2021