

Dalla teoria alla prassi: la sperimentazione sul *Vat.Lat. 3880*

Per quanto un linguaggio di codifica sia valido e flessibile, esso ha sempre delle limitazioni, poiché nessun sistema di marcatura è indipendente da un accordo preventivo su ciò che interessa recuperare informaticamente. La codifica, considerata come impresa dalla connotazione scientifica e interpretativa, apre dunque un mondo fatto di decisioni, teoriche e operative. Ciò che si vuole codificare è ciò che si vuole conservare, ovvero ciò che si ritiene importante:

memorizzare e codificare un testo, farne un'edizione – così come in passato – significa introdurre una nostra griglia interpretativa tra le sue pieghe: rompere i legami che in esso ci appaiono naturali per disporre le sue parti (quelle parti che siamo riusciti ad individuare) in un ordine diverso¹.

Simili presupposti indicano chiaramente come, confezionando un'edizione digitale, si creino dei vincoli di consapevolezza fondativi, lontani dalle soluzioni di comodo prospettate dai detrattori delle applicazioni informatiche alle discipline storiche. Un documento storico veicola, ovviamente, una serie di contenuti testuali, di informazioni su eventi, luoghi, persone, oggetti del mondo materiale, concetti, idee: riconoscerli, poterli ritrovare, passare dall'uno all'altro, collegarli fra loro costituisce un esperimento euristico che i linguaggi di marcatura consentono e che evidentemente fa slittare l'operazione di codifica di un testo dai suoi aspetti puramente linguistici, formali o strutturali a quelli relativi al suo contenuto informativo e ai più complessi significati di cui esso è portatore. Introdursi in un universo così complesso – almeno apparentemente – ha significato, pragmaticamente, operare delle scelte.

Nel caso prospettato il primo provvedimento da prendere ha riguardato, ad esempio, l'eventuale adozione o meno di un modello di marcatura preesistente. In effetti, nonostante l'assenza di precise linee-guida per la codifica informatica di fonti storiche, non è recentemente mancata – all'interno della comunità umanistica attenta alle tecnologie elettroniche – una spinta alla definizione di modelli e procedure condivisibili per l'edizione dei testi. In questa direzione, grande diffusione ha avuto la *Text Encoding Initiative (TEI)*, progetto di ricerca internazionale volto allo sviluppo di un sistema di codifica uniforme finalizzato alla rappresentazione dell'informazione testuale e alla gestione dei dati

¹ G. GIGLIOZZI, *Il testo e il computer. Manuale di informatica per gli studi letterari*, Milano, Bruno Mondadori 1997 (Sintesi), p. 209.

umanistico-letterari in formato elettronico, avviato nel 1987² e promosso dalle tre più importanti associazioni nel campo dell'informatica umanistica e della linguistica computazionale: l'*Association for Computers and the Humanities (ACH)*, l'*Association for Computational Linguistics (ACL)* e l'*Association for Literary and Linguistic Computing (ALLC)*³. Nella definizione di uno standard libero finalizzato all'interscambio di documenti umanistici tra piattaforme informatiche differenti, indipendente da una particolare applicazione software e adatto alla rappresentazione di tutte le categorie e caratteristiche testuali possibili, la *TEI* si è giustamente orientata verso una codifica di tipo dichiarativo, adottando inizialmente *SGML*⁴; nel giugno del 2002 il *TEI Consortium* ha provveduto inoltre alla pubblicazione di una nuova versione dello schema di codifica, *TEI P4*, contraddistinto dalla piena conformità con l'*eXtensible Markup Language*, proiettando prepotentemente l'iniziativa – sulla scorta del notevole successo di *XML* e delle tecnologie ad esso correlate – verso nuovi importanti sviluppi nell'ambito delle biblioteche digitali e del World Wide Web stesso⁵. Tuttavia, il modello descrittivo dei testi che sottende la *TEI* si

² L'iniziativa è stata avviata durante il convegno tenutosi nel 1987 al Vassar College di Poughkeepsie a New York, nel corso del quale sono stati redatti i cosiddetti *Poughkeepsie Principles*, ossia i principi base cui ci si sarebbe dovuti attenere nella realizzazione di uno schema di codifica dei testi letterari per la ricerca umanistica. Sulla *TEI* si veda, in particolare, C.M. SPERBERG-MCQUEEN, L. BURNARD, *Guidelines for the Encoding and Interchange of Electronic Texts (TEI P3)*, Chicago and Oxford, The TEI Consortium 1994 (Text Encoding Initiative).

³ Le tre associazioni sono raggiungibili ai seguenti indirizzi internet: <http://www.ach.org/> (*AHC*), <http://www.cs.columbia.edu/~acl/home.html> (*ACL*), <http://www.kcl.ac.uk/humanities/cch/allc/> (*ALLC*). Il principale supporto finanziario al progetto è fornito dall'*U.S. National Endowment for the Humanities (NEH)*: <http://www.neh.fed.us/>), dal *XIII Direttorato della Commissione Europea (Information Society Directorate-General CEC/DG-XIII)*: http://europa.eu.int/comm/dgs/information_society/index_en.htm), dalla *Andrew W. Mellon Foundation* (<http://www.mellon.org/>) e dal *Social Science and Humanities Research Council of Canada* (<http://www.sshrc.ca/>). Sull'onda del successo ottenuto dal linguaggio *TEI*, adottato in numerosi progetti accademici a livello internazionale, nel dicembre del 2000 i membri della *Text Encoding Initiative* hanno deciso di rendere permanente il progetto, costituendo il *TEI Consortium* (<http://www.tei-c.org>), una organizzazione internazionale senza scopo di lucro fondata al fine di sostenere e mantenere lo sviluppo e la diffusione dello standard *TEI*. Il *TEI Consortium* ha i suoi uffici esecutivi a Bergen in Norvegia ed è ospitato da quattro università nel mondo: University of Bergen, Brown University, Oxford University, e University of Virginia. Il Consorzio è diretto da un Consiglio di Direzione e la sua attività tecnico-scientifica è controllata da un Consiglio elettivo. Le attività sono svolte tipicamente da piccoli gruppi internazionali di esperti, coordinati da due curatori, uno in Nord America e uno in Europa.

⁴ Le prime specifiche provvisorie sono state pubblicate nel 1991 con il titolo *Guidelines for Electronic Text Encoding and Interchange, TEI P1*. Successive revisioni del linguaggio si sono avute nel 1992, *TEI P2*, nel quale la struttura della DTD è stata notevolmente rimaneggiata e nel 1994, *TEI P3*.

⁵ L'indirizzo da cui è possibile prelevare o consultare una copia delle direttive per la codifica è: <http://www.tei-c.org/P4X/>; ne esiste anche un'edizione a stampa, per i tipi della University of Virginia Press, scaricabile dal sito: <http://www.upress.virginia.edu/books/tei.html>. La *Document Type Definition* della *TEI* prevede oltre 500 elementi, ripartiti in uno schema caratterizzato da una forte modularità (*DTD Fragments*). I *DTD Fragments* definiscono tre gruppi di *tag* (*tag sets*), che vengono combinati fra di loro per produrre uno schema adatto alla tipologia del testo da sottoporre a codifica: un *Core tag set*, che contiene il nucleo minimo di elementi comuni ad ogni tipo di testo, un *Base tag set*, che contiene diversi *tag set* specifici per vari tipi di documenti fondamentali (lirica, prosa, dramma) e un *Additional DTD tag set*, che include elementi per la rappresentazione di caratteristiche utili ad evidenziare determinate peculiarità del testo (codifica di apparati di varianti, rappresentazione di strutture ipertestuali, codifica di fonti primarie). Al fine di aiutare gli utenti nella costruzione di versioni personalizzate della *DTD* nel sito del *TEI Consortium* si trova una applicazione Web, chiamata scherzosamente *Pizza Chef* con cui è possibile senza grande sforzo "cucinarsi", per

basa su una formalizzazione delle consuetudini nella produzione di documenti testuali definite a partire dalla diffusione della stampa, riproponendo una sostanziale trasposizione delle strutture e funzioni convenzionalmente vigenti nei testi letterari, siano essi in prosa o in poesia. Nonostante infatti nella *TEI* siano previsti marcatori di indubbia utilità nella descrizione delle trascrizioni, della tradizione manoscritta e dell'apparato critico⁶, si ricava infatti l'impressione che questo approccio, orientato a identificare e sovrapporre le regole di costruzione del testo con la logica funzionale delle macchine, sia consapevolmente destinato a lasciare in ombra la storicità dell'oggetto testuale, risultando dal punto di vista contenutistico e semantico assolutamente inadeguato e insoddisfacente. Se si concepisce la codifica del testo come una rappresentazione di quell'informazione che si aggiunge al testo vero e proprio, e che eccede da esso, allora la *TEI*, che si appropria al testo come rappresentazione grafica di un oggetto documentario o letterario, resta essenzialmente uno strumento limitato, e limitante. Il discorso vale in egual misura per l'adattamento della *TEI* alla documentazione storica medievale e moderna, denominato *Charters Encoding Initiative (CEI)* e avviato nell'aprile del 2004 su iniziativa di Georg Vogeler⁷. Anche in questo caso infatti, sebbene si tratti di un disegno assolutamente orientato alla trascrizione elettronica di documentazione dai forti connotati storici e basato sul *Vocabulaire International de la Diplomatie*⁸, attraverso il quale è stato possibile normalizzare gli elementi del discorso diplomatico, il suo eventuale utilizzo non riesce ad abbracciare i dati semi-strutturati, quelli cioè che interessano l'analisi storica, non rilevabili né tantomeno inseribili all'interno delle canoniche etichette diplomatiche e archivistiche.

La scelta, in fin dei conti, si è rivelata obbligata. Seguendo la *TEI* – o la *CEI* – l'edizione del codice Vat.Lat.3880, pur rispettando un modello internazionale e condiviso, si sarebbe

adottare l'ironica definizione usata nel sito dell'applicazione, la propria *DTD* (<http://www.tei-c.org/pizza.html>). Sempre con l'obiettivo di favorire la diffusione della *TEI*, la commissione editoriale del progetto ha inoltre prodotto una versione ridotta dello schema di codifica battezzata, *TEI-Lite*, ovvero una *DTD*, che contiene soltanto un sottoinsieme degli elementi (circa 170) dell'intero sistema *TEI*, pur mantenendo la piena compatibilità e conformità con esso. *TEI-Lite* è praticamente una versione semplificata dell'intero schema di codifica definito dalla *TEI*, creato per soddisfare le esigenze più comuni nella codifica dei testi e facilitare la realizzazione di testi in formato elettronico compatibili con l'intero schema, senza richiedere lo studio di tutta la *DTD*. Al momento *TEI-Lite* è il sottoinsieme di *TEI* più diffuso e adottato in numerosi progetti di codifica testuale e nella creazione di archivi documentali.

⁶ Ad esempio: per le abbreviazioni (<abbr> ed <expan>), le correzioni (<add>, <corr>,), i danni (<damage>), i cambi di mano (<handshift>); tra gli elementi dell'apparato (<app>): le lezioni (<rdg>) e i testimoni (<wit>), prevedendo anche la possibilità di implementare una lista (<witlist>) dei singoli testimoni (<witness>).

⁷ Le *Guidelines* all'indirizzo: <http://www.cei.uni-muenchen.de/index.htm>.

⁸ Opera collettiva intrapresa dal più autorevole organismo scientifico internazionale del settore, la *Commission Internationale de la Diplomatie. Comité International des Sciences Historiques*, ed. MA MILAGROS CÀRCEL ORTÍ, Valencia 1994.

configurata esclusivamente come la riproposizione elettronica dell'aspetto tipografico della fonte materiale di riferimento, lasciando in ombra quegli elementi logici e funzionali che in verità costituivano l'interesse primario di questa ricerca. Identificare un autonomo *set* di elementi e attributi, calibrato sulle caratteristiche specifiche della documentazione storica utilizzata, ma tale comunque da essere in grado di generare risultati significativi in fase di *information retrieval*, è sembrata dunque l'opzione più consona ai fini del progetto attuato e della stessa documentazione edita, la cui ricchezza informativa ha quasi costretto al superamento delle regole internazionali in favore di una marcatura forgiata sulla semantica e le specificità storico-territoriali dei documenti⁹. Nell'evidente diversità di prospettive, finalità ed esiti fra il sistema di marcatura rigidamente proposto dalla *TEI* e quello sperimentato in questa sede, dove si predilige con evidenza una prospettiva logica, lasciando volontariamente da parte gli aspetti tipografici del testo, si è tentato di proporre uno strumento ulteriore rispetto agli standard consolidati, ma con essi integrabile, e forse anche un modo nuovo, oltre che più elastico e semplice, per affrontare e ridurre in formato elettronico testi finalizzati alla ricerca e alla pratica storiografica.

L'idea a partire dalla quale si è poi costituito il modello di codifica adottato, ha previsto – sin dall'inizio – di attingere ai dati attraverso lo spoglio fisico dei documenti, secondo un approccio che ha contribuito in maniera determinante a fissare in partenza gli elementi descrittivi e analitici delle informazioni storiche perseguite. Parallelamente però, è avanzata la consapevolezza che la struttura di un documento ideale su cui basare l'operazione di marcatura non dovesse ricalcare esclusivamente le informazioni editoriali ma al contrario, dovesse essere finalizzata ad evidenziare tutta una serie di elementi aggiuntivi, integrabili con gli attributi stabiliti dai canoni editoriali e in grado di approfondirne il significato: «l'*hors-texte*, ciò che è fuori dal testo, è anche *dentro* il testo, si annida tra le sue pieghe: bisogna scoprirlo, e farlo parlare»¹⁰, ha giustamente detto Carlo Ginzburg ricordando come, nel valutare le proprie prove, gli storici dovrebbero sempre spazzolare la storia contropelo, imparando a leggere le testimonianze al di là delle intenzioni di chi le ha prodotte. In questo senso, il lavoro ha previsto il recupero di dati di livello superiore rispetto alla tradizionale

⁹ Una classificazione semantica si fonda sull'idea che essa non serva soltanto a produrre una fra le tante modalità di accesso ai documenti, ma offra invece uno strumento potente e pervasivo per l'organizzazione integrata di un'edizione digitale: gli strumenti semantici costituiscono infatti una guida alla formulazione efficace dei bisogni conoscitivi, all'impostazione delle ricerche e all'esame dei materiali a disposizione.

¹⁰ C. GINZBURG, *Rapporti di forza. Storia, retorica, prova*, Milano, Feltrinelli 2000 (Campi del sapere), p. 46; lo storico ha però ripreso alcune suggestive immagini proposte da Walter Benjamin nel suo *Sul concetto di storia*, Torino, Einaudi 1993 (Biblioteca Einaudi, 15).

descrizione archivistico-diplomatica, riformulandone la funzione di indicatori/qualificatori del contenuto dei documenti censiti. Si è pertanto provveduto alla costruzione di una struttura relazionale sovrimposta alla registrazione dei documenti, che valorizzasse il carattere selettivo e mirato dell'universo documentario contenuto nel *Liber Privilegiorum* e contemporaneamente esplicitasse e garantisse una serie ampia di collegamenti fra entità interne e oggetti esterni. In altre parole, il modello di codifica proposto, raccogliendo l'invito lanciato a più riprese da Michele Ansani¹¹, è stato calibrato sulle specificità storiche e territoriali della documentazione, introducendo una marcatura semantica con palesi – ma palesati – elementi di interpretazione personale, nella convinzione che la possibilità offerta dall'utilizzo di *XML* di restituire un testo comunque integrale fosse comunque in grado di ridurre al minimo i pericoli connessi nella dichiarata soggettività dell'operazione.

La metodologia adottata ha proceduto per gradi successivi di raffinamento, passando necessariamente da uno schema di codifica iniziale a maglie larghe che solo attraverso lo studio analitico e storico-diplomatistico dei documenti si è andato arricchendo di sfumature descrittive. In particolare, allo scopo di esplicitare essenzialmente le caratteristiche contestuali, compositazionali e interpretative:

- è stato esplorato il codice di partenza della fonte, di cui sono state delineate funzioni e regolarità;
- sono stati individuati i dati pertinenti, ovvero tutti gli elementi significativi della fonte, sciogliendo le ambiguità di entità e relazioni interne;
- è stato infine progettato un sistema di codifica dei dati adattato all'oggetto, al canale e ai destinatari.

La codifica progettata si articola in due macro-blocchi: quello delle metainformazioni di apparato e quello delle metainformazioni di testo, relative alla struttura e al contenuto del documento.

Le metainformazioni di apparato, poste in apertura ad ogni documento, sono relative a:

- elementi dell'edizione (<EDITIO>): posizione del documento all'interno del cartulario (<NUMDOC>), documenti precedenti e successivi che contrassegnano l'ordine del documento all'interno dell'edizione, permettendo la navigazione sequenziale al suo interno

¹¹ Lo studioso ha più volte ribadito l'importanza di una costruzione analoga; si veda in particolare il saggio *Diplomatica e nuove tecnologie. La tradizione disciplinare fra innovazione e nemesi digitale* cit.

(<DOCPREC>, <DOCSUCC>), informazioni editoriali (<INFOED>) quali la data topica e cronica del documento (<DATA>), il numero delle carte di riferimento nel manoscritto (<NUMCARTE>: es. <NUMCARTE cc="1v A-3v A">);

- elementi dell'apparato: la tradizione del documento (<TRADITIO>) in relazione agli originali (<ORIGIN>: es. <ORIGIN entecons="BCRS" fondo="FM" segn="Balsamo 29" dim="mm. 850+36590">) e le copie (<COP>: es. <COP entecons="BSEM" segn="XX.E.8">), le eventuali edizioni critiche passate (<ED>: es. <ED editio1="Pirro, Sicilia Sacra, I, p. 453" editio2="Del Giudice, Descrizione, p.2" editio3="Lo Monaco Aprile, Le Decime, pp.8-16">), i registi (<REG>: es. <REG regesto1="Behring, Regesten, 2, p. 201" regesto2="Garufi, Catalogo, n.15, pp.10-11">), le notizie bibliografiche (<BIBLIOGRAFIA>: es. <BIBLIOGRAFIA cit1="Barberi, Beneficia, I, p.34" cit2="Celauro, Il territorio, p.26" cit3="Concetti, Memorie, p.115ss." >);
- il regesto (<REGESTO>) ed eventuali commenti al documento (<OSSERVAZIONI>).

Attraverso questa sezione della marcatura vengono dunque delineati gli elementi del profilo del documento, che accompagnano ogni entità documentaria a prescindere dal supporto specifico di memorizzazione, includendo tutte le informazioni utili alla sua identificazione e interpretazione. Si tratta, propriamente, di metadati archivistici, finalizzati a documentare la formazione e sedimentazione dei singoli documenti, le loro reciproche interrelazioni, i legami con le strutture che li hanno prodotti, conservati o successivamente acquisiti e l'eventuale tradizione di studi che li ha accompagnati.

Il blocco delle metainformazioni di testo, come si è detto sopra, è stato bipartito in due sottoinsiemi distinti ma interagenti. Sono stati infatti previsti dei marcatori di struttura, attenti alla disposizione materiale del testo dei documenti e alla sua articolazione logica, riconoscendone l'importanza fondamentale in documentazione di natura giuridica, emanata da cancellerie pubbliche, all'intero della quale era d'obbligo individuare una definita struttura diplomatica finalizzata a corroborarne il valore legale. I marcatori definiscono lo svolgimento del discorso documentario, circoscrivendone anzitutto le articolazioni principali (<PROTOCOLLO>, <TESTO>, <ESCATOCOLLO>): ciascuno di questi segmenti include poi ovviamente delle sotto-partizioni.

- Tenore del documento (<TENOR>)
- Protocollo (<PROTOCOLLO>), coi suoi sottoelementi tradizionali di invocazione (<INVOCATIO>), intitolazione (<INTITULATIO>), iscrizione (<INSCRIPTIO>), data cronica e topica (<DTCRON> e <DTTOP>), preghiera (<APPRECATIO>), formula di perpetuità (<FORMPERP>)

- Testo (<TESTO>) e sottoelementi relativi: preambolo (<ARENKA>), narrazione (<NARRATIO>), dichiarazione (<PROMULGATIO>), disposizione (<DISPOSITIO>), sanzione (<SANCTIO>), autenticazione (<CORROBORATIO>)
- Escatocollo (<ESCATOCOLLO>), comprendente la rota della quale viene riportata la legenda esterna e interna e la citazione biblica (<ROTA>: es. <ROTA legendaest="Dextera Domini fecit virtutem. Dextera Domini exaltavit me" legendaint="WILLELMUS DIVINA FAVENTE CLEMENCIA REX SICILIE, DUCATUS APULIE ET PRINCIPATUS CAPUE." citbib="Ps 117(118), 16">), la data topica e cronica (<DTTOP>, <DTCRON>), la ricognizione cancelleresca (<RECOGNITIO>), le sottoscrizioni (<SUBSCRIPTIO>) espresse attraverso il *signa manus* dei sottoscrittenti (<SMS>) e l'elenco dei testimoni (<IT>), la *completio* del notaio (<COMPLETIO>).

Questa sezione della codifica, abilitata a definire l'articolazione del discorso documentario, è stata tuttavia strutturata secondo una griglia che, riflettendo in linea di massima la tradizionale analisi formale dei documenti medievali, non ha prescritto però regole di inclusione rigidamente preordinate e ha quindi consentito numerose eccezioni nel rispetto della prassi documentaria analizzata. Oltre ai metadati relativi alle partizioni documentarie, il blocco delle meta-informazioni di testo prevede l'annidamento di ulteriori *tag*, che identificano i contenuti e in particolar modo:

- i toponimi (<TOP>), con attributi obbligatori per la normalizzazione del nome ("nm") e l'identificazione del toponimo ("id"); ove non sia stato possibile procedere all'identificazione è stato previsto il valore "non-identificato". I toponimi identificati prevedono inoltre ulteriori attributi: gli attributi "tipo", attraverso il quale stabilire la tipologia del toponimo, se diverso dai toponimi indicanti località (in questo caso sono previsti i valori "microtoponimo"¹² o "el geografico"¹³) e "subtipo", che fornisce una definizione sintetica; l'attributo "ub" è invece relativo all'attuale ubicazione del toponimo. Un esempio di marcatura effettuata su un toponimo è il seguente: <TOP nm="Iati, castellum" id="Castello di Iato" tipo="microtoponimo" subtipo="castello" ub="Monte Jato, Comune di San Cipirello, Pa">;
- persone (<PERSONA>), con attributi obbligatori "nm" (normalizzazione del nome) e "id" (identificazione della persona) e, ove non sia possibile con valore "non-identificato". Qualora fossero presenti, sono stati segnalati gli attributi di parentela: "fil" (figlio), "pat" (padre), "mat" (madre), "sor" (sorella), "fr" (fratello), "vir" (marito), "ux" (moglie), "par" (familiare); gli attributi "nameattr" e "tit" identificano invece il titolo, la carica, la qualifica o il mestiere esercitati. Un ruolo a parte hanno invece i *tag* <SCRIPT> e <TT> che marcano rispettivamente il redattore del documento e gli eventuali testimoni: si tratta infatti di etichette che identificano le funzioni reali esercitate dalle persone nell'ambito dell'azione

¹² Rientrano in questa categoria: cappelle, casali, castelli, cimiteri, coltivazioni, cortili, croci, fortificazioni, giardini, guadi, mulini, porte, porti, recinti, scale, spelonche, strade, torri, vigne, etc.

¹³ Nella categoria di elemento geografico rientrano: campi, canali, conche, fiumi, fonti, fosse, grotte, isole, laghi, monti, colline, paludi, piane, prati, rupi, sorgenti, valli, etc.

giuridica e della documentazione, e che non necessariamente coincidono con le qualifiche denunciate. Un esempio di marcatura effettuata su un nome di persona è: <PERSONA nm="Silvester" nameattr="comes" id="Silvestro, conte di Marsico" tit="comes" fil="Guillelmus, comes Marsici">;

- istituzioni ecclesiastiche (<ECCL>), anche in questo caso con attributi obbligatori “nm” (normalizzazione del nome) e “id” (identificazione dell’istituzione, con possibile valore “non-identificato”). Sono inoltre previsti un attributo “tipo”, che identifica la tipologia dell’istituzione se dichiarata (ad esempio di una chiesa, un’arcivescovato, un’abbazia e via dicendo) e “ub”, secondo lo schema di “Città, Provincia” e la possibilità di implementare il valore “incerto”: es. <ECCL nm="Sanctae Mariae de Ammirato, ecclesia" id="Chiesa di S. Maria dell'Ammiraglio" tipo="chiesa" ub="Comune di Palermo, Pa">;
- Infine, sono stati introdotti anche i *tags* per gli elenchi e le descrizioni di beni sia immobili che mobili (<BENIMM> e <BENMOB>) – molto utili sebbene nella documentazione analizzata in verità scarseggino – e il *tag* <EVENT> per segnalare un fatto storico all’interno del documento.

Le informazioni così marcate sono, ovviamente, ricomprese entro strutture di livello superiore, che definiscono l’articolazione formale dei documenti, secondo i tre livelli di marcatura precisati, descrivendo ordinatamente la struttura ideale di ciascun testo contenuto nel cartulario, con le dovute eccezioni:

<EDITIO>

<NUMDOC/>

<DOCPREC/>

<DOCSUCC/>

<INFOED>

<DATA/>

<NUMCARTE/>

<APPARATO>

<TRADITIO>

<ORIGIN/>

<COP/>

</TRADITIO>

<ED/>

<REG/>

<BIBLIOGRAFIA/>

<REGESTO/>

</APPARATO>

</INFOED>

<TENOR/>


```

<PROTOCOLLO>
  <DTCRON> data cronica </DTCRON>
</PROTOCOLLO>
<TESTO>
  <DISPOSITIO> <PERSONA informazioni indicizzate> nome</PERSONA> azione <TOP
  informazioni indicizzate> luogo </TOP>
  </DISPOSITIO>
</TESTO>
<ESCATOCOLLO>
  <DTTOP> <TOP> data topica </TOP> </DTTOP> <SUBSCRIPTIO>
  <PERSONA informazioni indicizzate> sottoscrizioni autografe
  </PERSONA> </SUBSCRIPTIO> <IT> <PERSONA informazioni indicizzate> elenco dei
  testi </PERSONA> </IT>
  <COMPLETIO> <SCRIPT> <PERSONA informazioni indicizzate> nome del notaio
  </PERSONA> </SCRIPT>
  </COMPLETIO>
</ESCATOCOLLO>
</EDITIO>

```

La marcatura *XML* di tutti documenti editi del codice Vat.Lat. 3880 è liberamente accessibile a chiunque desideri consultarla: per visualizzare il file sorgente basta infatti scegliere dal menu del browser *Visualizza>Sorgente pagina* oppure *Visualizza>HTML* in Internet Explorer. A titolo esemplificativo si riproduce qui un documento interamente codificato secondo lo schema progettato, con i marcatori principali evidenziati in grassetto e le prime due righe, contenenti la dichiarazione dello *standard XML* e del *set* di caratteri impiegati e il percorso relativo del foglio di stile *XSL*, utilizzato per la trasformazione del documento in *HTML*.

```

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-stylesheet type="text/xsl" href="trasformazione.xsl"?>
<EDITIO>
  <NUMDOC>I.7</NUMDOC>
  <DOCPREC><a class="link" href="I.6.xml">Precedente</a></DOCPREC>
  <DOCSUCC><a class="link" href="I.8.xml">Successivo</a></DOCSUCC>
  <INFOED>

```

<DATA luogo="Palermo" anno="1183" mese="maggio" giorno="non-indicato" ind="I">Palermo 1183, maggio, I ind.</DATA>

<NUMCARTE>[cc. 13v A-13v B]</NUMCARTE>

<APPARATO>

<TRADITIO><ORIGIN>Orig.: Palermo, Bibl. Centr. della Regione Siciliana, Tabulario di S. Maria Nuova di Monreale, perg. nr. Balsamo 61, di mm. 365+52×322.</ORIGIN><COP>Cop.: Bibl. del Seminario Arcivescovile di Monreale, ms. XX.E.8.; Palermo, Bibl. Centr. della Regione Siciliana, Liber Pandectarum, ms. F.M. 14, f. 100</COP></TRADITIO>

<ED>Ed.: Pirro R., Sicilia Sacra disquisitionibus et noiiiis illustrata, 4 voll., Palermo 1644-47; nuova ed. riveduta e ampliata in 2 voll., Palermo 1733; rist. anast. con uno scritto di F. Giunta, 2 voll., Bologna 1987, I, p. 460</ED>

<REG>Reg.: Lello G.L., Historia della Chiesa di Monreale, Palermo 1702, n. 46</REG>

<BIBLIOGRAFIA>Cit.: Concetti C., Monreale e i suoi dintorni, Palermo 1912; rist. anast. Bologna 1991, p. 29</BIBLIOGRAFIA>

<REGESTO>Regesto: Guglielmo II, su richiesta di Ruggero di Tarsia, acconsente al suo matrimonio con Maria, figlia di Roberto Malconvenant; riconoscendo costoro che la tenuta di Bisacquino non gli appartiene legalmente, la rimettono al demanio reale, da cui viene concessa a Monreale.</REGESTO>

</APPARATO>

</INFOED>

<TENOR>Tenor privilegii prefati <PERSONA nm="Willelmus Secundus" nameattr="rex" id="Guglielmo II d'Altavilla (1166-1189)" tit="rex" pat="Willelmus Primus" mat="Margarita">regis Guillelmi</PERSONA> in quo concedit et donat <ECCL nm="Montis Regalis, ecclesia" id="Chiesa di S. Maria Nuova di Monreale" subtipo="chiesa" ub="Comune di Monreale, Pa">ecclesie Montis Regalis</ECCL> <TOP nm="Busackini, casalis" id="Casale di Bisacquino" tipo="microtoponimo" subtipo="casale" ub="Comune di Bisacquino, Pa">casale quod dicitur Busackinum</TOP> cum omnibus tenimentis et pertinentiis suis. Capitulum VII.</TENOR>

<PROTOCOLLO>

<INVOCATIO>In nomine dei eterni et salvatoris nostri Iesu Christi. Amen.</INVOCATIO><INTITULATIO><PERSONA nm="Willelmus Secundus" nameattr="rex Siciliae, ducatus Apuliae et principatus Capuae" id="Guglielmo II d'Altavilla (1166-1189)" tit="rex" pat="Willelmus Primus" mat="Margarita">Willelmus divina favente clemencia rex Sicilie, ducatus Apulie et principatus Capue.</PERSONA></INTITULATIO>

</PROTOCOLLO>

<TESTO>

<NARRATIO>Per hoc presens scriptum notum facimus quod postquam ad instanciam, precum et suplicacionum <PERSONA nm="Rogerius de Tarsia" id="Ruggero di Tarsia" tit="dominus" ux="Maria Malconvenant">Rogerii de Tarsia</PERSONA> fidelis nostri maiestas nostra matrimonium contrahi concessit inter eum et <PERSONA nm="Maria" nameattr="filia quondam Roberti Malconvenant" id="Maria Malconvenant" pat="Robertus Malconvenant" fr="Guillelmus Malconvenant" vir="Rogerius de Tarsia">Mariam filiam</PERSONA> quondam <PERSONA nm="Robertus Malconvenant" id="Roberto Malconvenant, barone di Racalmuto e signore di Bisacquino" tit="dominus" fil="Guillelmus,

Maria">Roberti Malconvenant</PERSONA> cum terra, quam iamdictus <PERSONA nm="Robertus Malconvenant" id="Roberto Malconvenant, barone di Racalmuto e signore di Bisacquino" tit="dominus" fil="Guillelmus, Maria">Robertus</PERSONA> iuste tenuerat, idem <PERSONA nm="Rogerius de Tarsia" id="Ruggero di Tarsia" tit="dominus" ux="Maria Malconvenant">Rogerio</PERSONA> et predicta <PERSONA nm="Maria" nameattr="uxor Rogerii" id="Maria Malconvenant" pat="Robertus Malconvenant" fr="Guillelmus Malconvenant" vir="Rogerius de Tarsia">Maria uxor eius</PERSONA> recognoscentes in curia nostra, <TOP nm="Busackini, casalis" id="Casale di Bisacquino" tipo="microtoponimo" subtipo="casale" ub="Comune di Bisacquino, Pa">Busackinum</TOP> non fuisse iuris ipsius <PERSONA nm="Robertus Malconvenant" id="Roberto Malconvenant, barone di Racalmuto e signore di Bisacquino" tit="dominus" fil="Guillelmus, Maria">Roberti</PERSONA> et se nullum ius habere in eo sed ad nos pertinere spontanea voluntate sua dimiserunt et perpetuo renunciaverunt illud in manus et demanium nostrum, quatinus nec ipse Rogerius nec prefata uxor eius nec heredes eorum, nec aliquis pro parte sua aut heredum eorum ullo unquam futuro tempore, ius aliquod in ipso <TOP nm="Busackini, casalis" id="Casale di Bisacquino" tipo="microtoponimo" subtipo="casale" ub="Comune di Bisacquino, Pa">Busackino</TOP> possint exigere.</NARRATIO> <DISPOSITIO>Et quoniam in proposito gerimus et incomutabili voluntate ad incrementum et commodum <ECCL nm="Montis Regalis, archiepiscopatus" id="S. Maria Nuova di Monreale" tipo="arcivescovato" ub="Comune di Monreale, Pa">archiepiscopatus Montis Regalis</ECCL> tota mentis affectione semper intendere, credentes illud thesauris nostris accrescere quod ei nostra fuerit liberalitate collatum, donamus et perpetuo concedimus eidem ecclesie predictum <TOP nm="Busackini, casalis" id="Casale di Bisacquino" tipo="microtoponimo" subtipo="casale" ub="Comune di Bisacquino, Pa">Busackinum</TOP> cum omnibus tenementis et pertinenciis suis. Quatinus ecclesia ipsa a modo et omni futuro tempore predictum <TOP nm="Busackini, casalis" id="Casale di Bisacquino" tipo="microtoponimo" subtipo="casale" ub="Comune di Bisacquino, Pa">Busackinum</TOP> cum omnibus tenementis et pertinenciis suis ad usum et utilitatem ipsius ecclesie et fratrum qui ibidem Deo servierunt libere et sine servicio teneat et sine alicuius molestatione possideat.</DISPOSITIO> <CORROBORATIO>Ad huius autem concessionis et donationis nostre memoriam et inviolabile firmamentum presens privilegium nostrum per manus <SCRIPT><PERSONA nm="Alexander" nameattr="notarius" id="non-identificato" tit="notarius regius">Alexandri nostri notarii</PERSONA></SCRIPT> scribi et bulla plumbea nostro tipario impressa iussimus roborari. Anno, mense et indictione subscriptis.</CORROBORATIO>

</TESTO>

<ESCATOCOLLO>

<DTTOP>Data in urbe <TOP nm="Panormus" id="Palermo" tipo="toponimo" subtipo="località" ub="Comune di Palermo, Pa">Panormi</TOP> felici</DTTOP><RECOGNITIO>per manus <PERSONA nm="Gualterius" nameattr="venerabilis Panormitani archiepiscopus" id="Gualtiero Offamilio, arcivescovo della chiesa di Palermo (1168-1191)" tit="archiepiscopus">Gualterii venerabilis Panormitani archiepiscopi</PERSONA> et <PERSONA nm="Matheus" nameattr="regis vicecancellarius" id="Matteo Aiello, vicecancelliere reale" tit="vicencancellarius regius" fil="Riccardus">Mathei regii vicecancellarii</PERSONA>, <PERSONA nm="Riccardus" nameattr="venerabilis Messanensis archiepiscopus" id="Riccardo Palmer, vescovo della chiesa di Siracusa (1157-1183)" tit="episcopus">Riccardi venerabilis Messanensis archiepiscopi</PERSONA> domini regis familiarium,</RECOGNITIO><DTCRON anno="1183" mese="maggio" giorno="non-indicato" ind="I">anno dominice incarnationis M^oC^o octogesimo tercio, mense madii, indictionis prime, regni vero domini nostri <PERSONA nm="Willelmus Secundus" nameattr="rex Siciliae, ducatus Apuliae et principatus Capuae" id="Guglielmo II d'Altavilla (1166-1189)" tit="rex" pat="Willelmus Primus" mat="Margarita">Willelmi Dei gratia magnifici et gloriosissimi regis Sicilie,

ducatus Apulie et principatus Capue</PERSONA> anno XVIII°.</DTCRON>
<APPRECATIO>Feliciter. Amen.</APPRECATIO>

</ESCATOCOLLO>

</EDITIO>

L'introduzione di una marcatura siffatta, obbligando all'analisi approfondita del documento su cui opera e grazie all'inserimento di attributi specificanti le proprietà di ogni singolo lemma etichettato, è stata in grado di produrre ritorni significativi in termini di ricerca automatica di informazioni all'interno del cartulario monrealese. L'atteggiamento seguito è stato quello di adeguare via via il lavoro ai problemi nuovi che sorgevano nel corso dell'analisi, predisponendo una codifica iniziale debole che si riservasse la possibilità di intervenire più e più volte sul testo, man mano che emergevano nuovi interrogativi o che la documentazione studiata andava acquisendo maggiore consistenza. Tali presupposti si sono rivelati efficaci ancor più nell'ambito di un progetto di ricerca che, per definizione, non ha voluto prevedere fin dall'inizio i suoi possibili sviluppi, rimandando – anche nel modello di codifica proposto – aperto a successive ridefinizioni, alla luce di nuovi dati o di una diversa interpretazione dei dati esistenti.