Иницијал. Часопис за средњовековне студије 13 (2025) 243–266
Initial. A Review of Medieval Studies 13 (2025) 243–266

UDC: 091=163.41“13“:004.352.22
DOI: 10.29341/IN.13.0.243266

Žarko VUJOŠEVIĆ
University of Graz, Faculty of Humanities, Department of Digital Humanities
Elisabethstraße 59, Graz, Austria, zarko.vujosevic@uni-graz.at
https://orcid.org/0009-0009-7309-3213

Pavle NIKOLIĆ
Belgrade, Serbia, nikolicpavle1419@gmail.com
https://orcid.org/0000-0002-0951-8303

PROSPECTS OF DIGITAL EDITION OF DIPLOMATIC MATERIAL ILLUSTRATED IN THE EXAMPLE OF THE DOCUMENTS ISSUED BY KING STEFAN UROŠ III DEČANSKI (1321–1331)

Abstract: The article presents a prototype of a digital edition of diplomatic material, exemplified through the scarce yet typologically and in terms of textual tradition diverse documentary heritage of King Stefan Uroš III (Dečanski). The methods of encoding metadata and texts of charters and letters in XML format are discussed, and the application of the widely adopted standard Text Encoding Initiative (TEI), supplemented by appropriate concepts from the *Vocabulaire international de la diplomatique* (VID), is demonstrated. The functions of the XPath and XQuery expression languages as well as the application of the Keyword in Context (KWIC) method are examined and shown through selected examples. The preliminary conclusions regarding the contribution of digital methods in diplomatic research are offered, along with proposals for further work on Serbian documentary sources.

Keywords: diplomatic material, digital edition, charters and letters of Stefan Dečanski, XML, Text Encoding Initiative, Vocabulaire International de la Diplomatique, digital research methods.

A digital (scholarly) edition

A scholarly edition entails a critical interpretation of a (historical) text, that is, a research-based processing of the content of one or more source documents, and analytically reproducing and/or reconstructing it.[1]

[1] The presented definition is derived from the one given and thoroughly explained by P. SAHLE, What is a Scholarly Digital Edition?, *Digital Scholarly Editing:*

A digital edition provides the same – basically the reconstruction of a literary text – but in a digital environment, by using available digital technologies in order to reproduce any level or detail of the text in question.[2] It is not a digitized edition (of what was previously printed), nor a database, and it doesn't contain only text. Rather, a digital edition is comprised of text, metadata (description, commentary, images), programming functions (browsing, systematization, statistical processing) and possibly other content such as maps and graphs, with the text and metadata being generated as machine-readable codes. That is why it cannot be printed out without losing (at least a part of) its substance and functionalities.[3] Finally, all of the mentioned elements are customizable, thus, an author can at any time add, delete, or change their content. Consequently, a (slightly extended) definition can be also given as follows: digital edition is a critical interpretation of a text by means of a computer program, which serves as the carrier of both content and various functions, providing the possibility of continuous additions and refinements.

From the very beginning of Digital Humanities as a discipline, interested scholars were envisioning the possibilities that a digital environment would provide, and in the case of a digital edition to overcome the restrictions of a printed edition.[4] Peter Robinson, a researcher behind one of the first digital edition projects, the „Canterbury Tales Project", wrote about the „millennial enthusiasm", as he named it, of his colleagues and

Theories and Practices, eds. M. J. DRISCOLL – E. PIERAZZO, Cambridge 2016, 19–39, pp. 23–26.

[2] G. FRANZINI – M. TERRAS – S. MAHONY, Digital Editions of Text: Surveying User Requirements in the Digital Humanities, *Journal on Computing and Cultural Heritage* 12/1 (2019) 1–23, p. 1; cf. P. SAHLE, What is a Scholarly Digital Edition?, 26–28.

[3] P. SAHLE, What is a Scholarly Digital Edition?, 27.

[4] Ibidem, 20. In fact, the digital approach to historical sources has itself a relatively long history. Considered the pioneering project in the field, the „Index Thomisticus" of Roberto Busa started in 1946, creating concordances to the works of Thomas Aquinas (*Index Thomisticus: Sancti Thomae Aquinatis operum omnium indices et concordantiae*, Computers and the Humanities 18 [1984] 109–137, Review by D. BURTON; see also www.corpusthomisticum.org/it – all cited links accessed on 15 September 2025). For the history of the Digital Humanities in general: S. HOCKEY, The History of Humanities Computing, *A Companion to Digital Humanities*, eds. S. SCHREIBMAN – R. SIEMENS – J. UNSWORTH, Oxford 2004, 1–19; E. VANHOUTTE, The Gates of Hell. History and Definition of Digital | Humanities | Computing 1, *Defining Digital Humanities*, eds. M. TERRAS – J. NYHAN – E. VANHOUTTE, Farnham 2013, 119–156; *Digital Humanities. Eine Einführung*, eds. F. JANNIDIS – H. KOHLE – M. REHBEIN, Stuttgart 2017, 3–12 (M. THALLER).

himself at the beginning of the „digital era“.[5] Impressed by the possible golden path towards a perfect edition, a series of ambitious projects was started with the goal of collecting all of manuscripts and printed texts of various works, authors, or time periods. However, the vast array of possibilities posed a unique challenge for editors attempting to define the framework of their desired digital edition. Many of them didn't know what exactly they had produced: so for instance, „The Book of Duchess“ was categorized as a „hypertext edition“, „The William Blake Archive“ as a „hypertext archive“, and „The Dante Gabriel Rosetti archive“ as a „hypermedia research archive“.[6] As the earliest examples of „successful“ (or „real“) digital editions „The Canterbury Tales Project“ (started in 1989), „Codex Sinaiticus“ (online since 2009), „Briefwechsel Sauer–Seuffert“ (started in 2012) and „Digital Dead Sea Scrolls“ (the digital library launched in 2012) have been recognized.[7] Today, the number of digital editions produced under various frameworks has grown significantly. Their overview is given by two extensive catalogs, providing basic details and relevant links: „The Catalogue of Digital Editions" (started in 2012), compiled by Greta Franzini, which currently lists 352 editions – 91 of which deal with medieval sources – and „Digital Scholarly Editions" (started in 2008), compiled by Patrick Sahle, comprising 714 editions, 114 of which are related to the medieval period.[8]

Markup languages and encoding schemata (XML, TEI, CEI)

To create a digital edition, a markup language is mostly used.[9] Markup languages are syntax systems designed for marking different

[5] P. ROBINSON, What is a Critical Digital Edition?, *Variants: The Journal of the European Society for Textual Scholarship* 1 (2002) 43–62, p. 45: „In place of editorial anxiety about which text, what choices, we can have all the texts, all the versions, and never have to make any decision. We can have images, lots of images, showing all different forms of the text, from manuscript scrawls, through cheap serialisations and deluxe printings. We can add galleries of sound, maps, videos: whole kits of material to equip the reader to taste the air the text breathed.“

[6] P. ROBINSON, What is a Critical Digital Edition?, 47.

[7] G. FRANZINI ET AL., Digital Editions, 2.

[8] https://dig-ed-cat.acdh.oeaw.ac.at/index.html; https://v3.digitale-edition.de/.

[9] For other methods see for instance H. COOLS – R. PADLINA, Formal Semantics for Scholarly Editions, *Graph Data-Models and Semantic Web Technologies in Scholarly Digital Editing*, eds. E. SPADINI – F. TOMASI – G. VOGELER, Schriften des Instituts für Dokumentologie und Editorik 15, Norderstedt 2021, 97–124 (https://kups.ub.uni-koeln.de/55229) or S. ALASSI – L. ROSENTHALER, Semantic precision: crafting RDF-based digital editions for unveiling the layers of historical correspond-

parts of documents with standard codes or tags, which are interpreted by computer programs as instructions for specific actions, such as properly displaying text and extracting particular information. In 1986, SGML (Standard Generalized Markup Language) was adopted as an international standard, in order to provide a set of rules for marking up electronic documents.[10] As a simpler application of the annotation principles defined by SGML, HTML (HyperText Markup Language) was developed and became the dominant language for marking, preparing, and publishing electronic documents on the web. But as HTML was designed primarily to display data on web browsers, there was a need for a markup language that would focus on describing data and specifying the nature of the enclosed information – thus, XML was born.[11]

XML (Extensible Markup Language) represents a set of rules for defining semantic tags that break a document into identifiable parts. It is often categorized as a meta-markup language due to its high level of customization, as it lets the user create their own tags according to their own needs, although, those tags must be organized according to a certain set of principles.[12] For example, XML requires that tags must be created in the form of a quite simple meta-syntax formula `<x>`,[13] where the symbols < and > are needed to begin and end the tag (with <> for opening and </> for closing tags), while `x` isn't specified at all, as XML lets the user input any information they want.

Example:

`<head><title>`Charter XY`</title></head>`

would be the same as

`<заглавље><наслов>`Charter XY`</наслов></заглавље>`.

Due to its flexibility, XML enables individual professions to develop their own domain-specific markup languages, which do not rely on

ence, *Digital Scholarship in the Humanities* 39/3 (2024) 813–835 (https://doi.org/10.1093/llc/fqae027).

[10] *Digital Humanities*, 128–146, p. 132 (G. VOGELER – P. SAHLE).

[11] G. G. CHOWDHURY – S. CHOWDHURY, *Organizing Information: From the Shelf to the Web*, London 2007, 158–163.

[12] On XML see the whole chapter in *Digital Humanities*, 128–146 (G. VOGELER – P. SAHLE), especially pp. 132–134. It is also important to note that XML is much more capable of describing structure and semantics than text formatting, as the XML document mostly contains tags which define the logical organization and meaning of its content rather than its visual representation (E. R. HAROLD, *XML 1.1 Bible*, Indianapolis 2004, 5).

[13] Computer codes, programming syntax, and technical commands are formatted in Courier New font for visual distinction.

support from browser manufacturers or plug-ins. As XML is using only basic text encoding in Unicode, any XML document could be sent to another person without worrying whether the recipient has the adequate proprietary software to open it. It is ideal for large and complex documents because the data is structured, and researchers can specify a standardized and unique syntax that defines the elements of the document structure and relations between those elements.[14] Therefore, XML is established as the standard for digital representation of texts as it is very well capable of supporting their structural analysis.

For marking-up textual contents and features in a standardized way, an appropriate encoding schema is needed. In the last approximately 40 years the guidelines of the TEI (Text Encoding Initiative) are widely used by many of linguistic and literary disciplines for encoding and representing texts in digital form. TEI is an international consortium established in 1987, that maintains the guidelines (the TEI guidelines) and continually revising them according to demands and recommendations from researchers.[15] With the emergence of XML as a dominant markup language in a short time span, the TEI consortium decided to shift from SGML to XML, in order to maintain compatibility with existing standards. This resulted in the widespread adoption of XML-TEI as a standard for encoding and preservation of textual documents by major text archives. The combination means the use of XML markup structure together with defined ontologies for text features, i.e. the TEI elements and attributes.

Example:

```
<scriptNote>Cyrillic
<term type="style">uncial</term>
</scriptNote>,
```

where `<scriptNote>` and `<term>` are elements designing the features of the script (of the text in question) and `type` is an attribute which allows narrower descriptions – in this case, that the style of the script is uncial (for specifications of these and other elements and attributes see the current TEI guidelines at https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html).

The main goals of TEI are to develop the ways of representing the textual features in the digital environment, making them usable for researchers without specific software, and to allow researchers to efficiently

[14] E. R. HAROLD, *XML 1.1 Bible*, 5–8.
[15] https://tei-c.org.

process the texts, providing user-defined extensions.[16] Consequently, the TEI guidelines have established themselves as the major standard in digital processing and presentation of texts.

However, regarding encoding medieval documents there are no common standards yet. Most of the proposals for a standardized XML schema for digital edition of diplomatic material remain project-specific.[17] In the meantime even a domain-specific markup language based on XML was developed for charters – the CEI (Charters Encoding Initiative). It started in 2004 at the University of Munich, proposing an XML markup specially constructed to address research needs regarding digitally presenting of medieval and early modern charters, in particular for expressing formulaic language of diplomatic texts (for which the TEI schema occurs too general).[18] Since its beginnings, CEI was and is still used in some significant digital charter projects such as „Edizione digitale dei documenti dell'abbazia di S. Maria della Grotta di Vitulano" at Monasterium.net,[19] „Regesta Imperii" of the Academy of Sciences and Literature in Mainz,[20] and the prepublication of the charters issued by the Emperor Charles IV from 1357 to 1378 by the MGH working group at the Berlin-Brandenburg Academy of Sciences in Berlin.[21]

In spite of being accepted by at least a part of the community, CEI still remains improvable, in particular as it is too narrowly focused on charters of Western European tradition.[22] Nevertheless the biggest chal-

[16] J. CUMMINGS, The Text Encoding Initiative and the Study of Literature, *A Companion to Digital Literary Studies*, eds. R. SIEMENS – S. SCHREIBMAN, Indianapolis 2013, 451–476, pp. 451–456.

[17] Ž. VUJOŠEVIĆ – G. VOGELER, Encoding Medieval Charters in preparation for a Diplomatics Semantic Web. Combining the TEI with the Vocabulaire International de Diplomatique, *Digital Medievalist* 18 (2025) – forthcoming (preprint at Zenodo: https://doi.org/10.5281/zenodo.14696237) [1–30, p. 4], also naming some exemplary projects.

[18] G. VOGELER, Towards a Standard of Encoding Medieval Charters with XML, *Literary and Linguistic Computing* 20/3 (2005) 269–280 (see also https://www.cei.lmu.de). CEI strongly suggests using XML as the dominant standard for data exchange, also re-using many tags from the previous versions of the TEI guidelines.

[19] https://www.monasterium.net/mom/SMG1200-1250/collection.

[20] https://www.regesta-imperii.de/unternehmen/ri-online.html.

[21] Belonging to MGH series „Constitutiones et acta publica imperatorum et regum – https://telota.bbaw.de/constitutiones/index.html.

[22] A digital editor could find marking-up diplomatic elements not flexible enough in CEI. For instance, for encoding penal clause (sanctio), the schema provides only a generic `<sanctio>` tag with no attrubutes for sub-types (poena materialis, poena spiritualis). From the perspective of the Southeast European diplomatics,

lenge in using CEI is the fact that it is getting somehow isolated by the rising dominance of TEI guidelines in almost all literary disciplines. In the recent years the potential of the TEI as the main reference for the humanities was recognized also among diplomatists.[23] By using TEI, digital diplomatics and charter encoding would not be a „lone island" anymore, as its digital editions would be easily readable and understood by researchers from other disciplines. Diplomatists would take a step closer towards a standardized digital edition of texts, bridging disciplinary divides.

The challenge which diplomatists are facing in using TEI relates to describing certain diplomatic forms in charter encoding, as TEI is too general and lacks concepts such as invocation, notary subscription, logos-formula, and many others. Therefore, a meaningful and creative adaption of the respective ontologies is necessary, which means integrating existing frameworks (non TEI-vocabularies) tailored to this specific category of documents. One solution is the application of *Vocabulaire international de la diplomatique* (VID), a multilingual thesaurus attempting at defining most of the diplomatic phenomena.[24] In 2013, it became available on the semantic web (www.cei.lmu.de/VID). The terms and headings in VID can be easily converted into structures made available by the W3C SKOS vocabulary, which allows linking its entries with respective XML tags.[25] As VID is (still) not complete, lacking some concepts which may be crucial for charter description (such as self appellation, recording note, deperdita and others), new terms and definitions were added to its

the lack of concepts belonging to Byzantine tradition, such as logos-formula or menologem, can present a significant obstacle.

[23] Most recently by G. VOGELER, Die Text Encoding Initiative (TEI) als Werkzeug des Urkundeneditors – Erfahrungen und Desiderate, *Papsturkundenforschung zwischen internationaler Vernetzung und Digitalisierung. Neue Zugangsweisen zur europäischen Schriftgeschichte*, eds. I. FEES – B. HOTZ – B. SCHÖNFELD, Göttingen 2015 (res doctae) – https://doi.org/10.26015/adwdocs-441. For charters projects already using TEI see Ž. VUJOŠEVIĆ – G. VOGELER, Encoding Medieval Charters, [5].

[24] *Vocabulaire international de la diplomatique*, ed. M. M. CÁRCEL ORTÍ, Valencia 1997 (1st ed. 1994). The approach was proposed by Georg Vogeler and Žarko Vujošević at the Next-Gen TEI Conference in 2021: https://doi.org/10.5281/zenodo.6523304 („Using non-TEI vocabularies in TEI descriptions"). As a comprehensive proposal it is (to be) published in Ž. VUJOŠEVIĆ – G. VOGELER, Encoding Medieval Charters, [1–30].

[25] SKOS (Simple Knowledge Organization System) belongs to the Semantic Web standards, developed by the World Wide Web Consortium (W3C) – https://www.w3.org/2004/02/skos/.

SKOS version.[26] Based on these preconditions, the XML-TEI-VID combination can be very productive in encoding medieval charters. The application of the approach will be presented in the example of a corpus of documents belonging to Serbian diplomatic material.

Implementation of the proposed encoding schema (XML-TEI-VID) in medieval Serbian diplomatic material

The approach described above is consequently employed in the project MeSeCA (Medieval Serbian Charters Action), being currently realised at the University of Graz, Department of Digital Humanities.[27]

As an interim result, the documents issued by King Stefan Uroš III. Dečanski (1321–31) have been processed and prepared to be published as a digital edition. The corpus in question is relatively scarce, comprising of (only) 22 diplomatic units, which are preserved in 27 copies (diplomatic items), even if two of them, the both versions of the charter for the Monastery of Dečani, belong to the largest documents of the Serbian Middle Ages.[28] On the other hand, it is very diversified regarding typology and tradition, as contains 12 donation or confirmation charters (of which 10 for church addressees and two for the Municipality of Dubrovnik), seven letters serving as notifications or instructions, and three quittances for Dubrovnik citizens. Most of the documents are originals (with several ecclesiastical charters being still disputed), while six are preserved only as register copies (in Italian translation, except the letter for Venice, originally written in Latin). The corpus bears a rich content relevant for different fields of historical research, as for instance the names of numerous places and persons mainly in the charters issued

[26] Z. VUJOŠEVIĆ – G. VOGELER, Encoding Medieval Charters, [6–10]. (For all proposed additions see VID/skos/VID.skos.ttl at FSGervers · GVogeler/VID · GitHub, lines 13199–13398.)

[27] https://gewi.uni-graz.at/de/unsere-forschung/drittmittelprojekte/drittmittelprojekte-2024/medieval-serbian-charters-action.

[28] For „diplomatic unit" and „diplomatic item" see Ž. VUJOŠEVIĆ, Diplomatarium Serbicum Digitale. Eine Wiederaufnahme zwischen Datenbank und digitaler Edition, *From Digital to Distant Diplomatics*, eds. D. LUGER – G. VOGELER, Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde, Beiheft 22 (2025), 255 –275, p. 256 and footnote 5. The first Dečani charter is preserved as a 5,2 m long scroll of parchment at the State Archives of Serbia, while its extended version as a parchment book of 71 folia gone lost after being published in 1880 (for references see Listing 1).

for monasteries.[29] Having such general characteristics, the documentary production of King Stefan Dečanski gives the opportunity to researchers to deal with many different diplomatic phenomena, and especially to survey particular encoding approaches.

Listing 1: Overview of documents issued by Stefan Uroš III[30]

Date	Recipient	Tradition	Applicable printed edition
(1308–14) – issued by Stefan as King of Zeta and the heir to the Serbian throne	Vranjina monastery	Original	Ž. VUJOŠEVIĆ, Povelja prestolonaslednika Stefana Uroša (Dečanskog) manastiru Vranjini, *Initial* 11 (2023) 191–213, pp. 200–203.
27 December 1321	Dubrovnik	Original	N. PORČIĆ, *Dokumenti srpskih srednjovekovnih vladara u dubrovačkim zbirkama. Doba Nemanjića*, Beograd 2017, 192–193.
Before 17 June 1323	Dubrovnik citizen Andrija Peštić	Register copy in Italian translation	N. PORČIĆ, *Dokumenti*, 193–194.
Before 17 June 1323	Dubrovnik	Register copy in Italian translation	N. PORČIĆ, *Dokumenti*, 194–195.
September 1324 (?)	St. Apostles church (Bishopric of Hum)	Original	V. TRIJIĆ, Povelja kralja Stefana Dečanskog manastiru Svetih apostola Petra i Pavla na Limu, *Stari srpski arhiv* 5 (2006) 51–64, pp. 53–56.

[29] Most of this content has been already processed in printed editions accompanied by commentaries and toponomastic and prosopographic analyses (see Listing 1).

[30] The metadata presented here and in the material below resulted from research within the project Diplomatarium Serbicum Digitale (DSD) at the Audiovisual Archive and Digitization Center of the Serbian Academy of Sciences and Arts (Ž. VUJOŠEVIĆ, Diplomatarium Serbicum Digitale, 255–275; as „work-in-progress project“ – https://ava.sanu.ac.rs/o-avi/mustra2/3 – DSD up to now has only a „beta-version“). That is why they may differ from the metadata given in the listed printed publications. Included in this survey are also new editions, which were made by Žarko Vujošević within the DSD and MeSeCA projects for all the charters for ecclesiastic recipients (except those for Dečani monastery).

Date	Recipient	Tradition	Applicable printed edition
Before 31 October 1325	Dubrovnik	Original	N. PORČIĆ, *Dokumenti*, 197.
11 January 1326	Dubrovnik citizen Luka Lukarević	Register copy in Italian translation	N. PORČIĆ, *Dokumenti*, 198–199.
Before 9 February 1326	Dubrovnik	Register copy in Italian translation	N. PORČIĆ, *Dokumenti*, 199.
25 March 1326	Dubrovnik	Original	N. PORČIĆ, *Dokumenti*, 199–200.
25 March 1326 (?)	Vranjina monastery	Original	F. MIKLOSICH, *Monumenta Serbica spectantia historiam Serbiae, Bosnae, Ragusii*, Wien 1858, 112-113, n. 93.
Before 19 April 1326	Dubrovnik	Original (alongside a register copy in Italian translation)	N. PORČIĆ, *Dokumenti*, 201–202.
April – July 1326 (?)	St. Mother of God church (Bishopric of Prizren)	Original	S. MIŠIĆ, Hrisovulja kralja Stefana Uroša III Prizrenskoj episkopiji, *Stari srpski arhiv* 8 (2009) 11–36, pp. 12–19.
15 July 1326	Dubrovnik citizen Andrija Peštić	Register copy in Italian translation	N. PORČIĆ, *Dokumenti*, 202–203.
15–28 July 1326	Dubrovnik	Original (alongside a register copy in Italian translation)	N. PORČIĆ, *Dokumenti*, 203–204.
8 February 1327	Hilandar monastery	Original (alongside a copy from the 15th century)	S. MIŠIĆ, Povelja kralja Stefana Uroša III manastiru Hilandaru, *Stari srpski arhiv* 5 (2006) 65–81, pp. 66–68; 70–72 (both versions).
9 July 1327	Hilandar monastery	Original	S. MIŠIĆ, Hrisovulja kralja Stefana Uroša III Hilandaru, *Stari srpski arhiv* 2 (2003) 29–40, pp. 31–33.

Date	Recipient	Tradition	Applicable printed edition
5 September 1327	Hilandar monastery	Original (alongside a copy from 1843)	S. MIŠIĆ, Hrisovulja kralja Stefana Uroša III Hilandaru o sporu oko međa Kruševske metohije, *Stari srpski arhiv* 3 (2004) 3–17, pp. 5–7.
Before 3 December 1328	Dubrovnik	Original	N. PORČIĆ, *Dokumenti*, 204–205.
1 May 1330	Venice	Register copy	N. PORČIĆ – N. ISAILOVIĆ, *Dokumenti srednjovekovnih vladara Srbije i Bosne u venecijanskim zbrikama*, Beograd 2019, 222–224.
9 September 1330	Monastery of St. Nicholas in Orehovo	Original	S. MIŠIĆ, Povelja kralja Stefana Uroša III Dečanskog manastiru svetog Nikole Mračkog u Orehovu, *Stari srpski arhiv* 1 (2002) 55–68, pp. 56–59.
After September 1330	Dečani monastery	Original (alongside a copy from 1762)	M. GRKOVIĆ, *Prva hrisovulja manastira Dečani - The First Charter of the Dečani Monastery*, Beograd 2004, 64–81.
Autumn 1330 – August 1331	Dečani monastery	Original (lost)	M. S. MILOJEVIĆ, Dečanske hrisovulje, *Glasnik Srpskog učenog društva* 12 (1880) 1–142, pp. 1–68.

In the following listing (Listing 2), an example of a charter encoded in XML by using TEI guidelines and VID concepts will be presented. Although proprietary software is not needed for XML encoding, this survey employs Oxygen XML Editor (latest version 27.1) for its enhanced functionalities and features. The model document is the charter issued probably

in September 1324 to the Church of St. Apostles upon Lim, the seat of Hum diocese.[31]

The XML document starts with the tag `<teiHeader>`, an element of almost every XML-schema for presenting metadata – here including title, document description with particular external and internal features as well as commentaries on provenance and tradition (if applicable), list of editions and relevant literature (here reduced to each one title) and abstract (regest).

Listing 2: XML-TEI encoding of metadata
with references to the VID terminology

```
<teiHeader>
 <fileDesc>
  <titleStmt>
   <title>1324-09-xx Стефан Дечански - Повеља краља
Стефана (Дечанског) цркви Св. апостола на Лиму, седишту
Хумске епископије, о даровању три црквена властелинства
у Полимљу (септембар 1324 ?)</title>
  </titleStmt>
  <sourceDesc>
   <msDesc>
   <msIdentifier>
    <settlement>Света Гора</settlement>
    <repository>Манастир Хиландар</repository>
    <idno>Хил. 12 (А 1/3)</idno>
   </msIdentifier>
   <msContents>
    <textLang>Српски</textLang>
   <msItem>
    <filiation><term
ana="VID:42">Оригинал</term></filiation>
   </msItem>
   </msContents>
   <physDesc>
   <objectDesc>
   <supportDesc>
```

[31] The presented encoding is a simplified, extracted version of the completely coded XML file (parts of which are presented as screenshots – see Figure 1). The text content is bolded to improve clarity and visibility. In order to keep „validity" of the XML-document in this figure, the VID concepts used in the example are named here: VID:42=„L'original"; VID:16=„L'auteur d'un acte écrit"; VID:17=„Le destinataire d'un acte"; VID:540=„Un sceau pendant"; VID:add_27=„lost" (VID/skos/VID.skos.ttl at FSGervers · GVogeler/VID · GitHub, lines 13186–13393 – "Any object relevant to diplomatics research, that is expected but not existant is 'lost'.")

```
    <support><material>Пергамент (два првобитно слепљена
комада, накандно спојена конопцем)</material></support>
     <extent><dimensions unit="mm">
     <height>639-642</height><width>285-307</width>
     </dimensions></extent>
     <condition>Оштећења услед влаге чине поједине делове
документа, посебно горњу половину (око 30 редова), тешко
читљивим. Пошто је документ био више пута пресавијен,
уочљиви су и трагови пресликаног мастила.</condition>
    </supportDesc>
    </objectDesc>
    <scriptDesc><scriptNote>Ћирилица<term
type="style">устав</term></scriptNote>
    </scriptDesc><decoDesc>
    <decoNote ana="red">Крст симболичке инвокације,
иницијал и потпис</decoNote>
    </decoDesc>
    <sealDesc ana="VID:540 VID:add_27"><condition
ana="traces">Трагови. Видљива су три прореза у средини
првог реда потписа, два изнад, један испод, предвиђена
за висећи печат.</condition></sealDesc>
     </physDesc>
    <history>
     <origin><origPlace>Непознато</origPlace><origDate>
септембар 1324 (?)</origDate></origin>
     <provenance>Заједно са осталим сачуваним повељама
за <orgName key="Св. Петар и Павле на Лиму"
type="church">Св. апостоле на Лиму</orgName> (1254-xx-xx
tpq Uroš I, 1317-05-xx tpq Milutin, 1343-10-25 Dušan),
уз још две које су припадале њиховом корпусу докумената
(1243-xx-xx taq Vladislav, сачувана као трансумпт уз
1243-xx-xx tpq Uroš I), и ова је у непознато време и под
непознатим околностима доспела у <orgName key="Хиландар"
type="monastery">манастир Хиландар</orgName></provenance>
    </history>
    </msDesc>
    <listBibl type="edition">
     <bibl>С. Новаковић, Законски споменици српских
држава средњега века, Београд 1912), 598 (према Љ.
Ковачевићу, извод)</bibl>
     <bibl>[...]</bibl>
    </listBibl>
    <listBibl type="literature">
     <bibl>Д. Синдик, Српска средњовековна акта у
манастиру Хиландару, Хиландарски зборник 10 (1998) № 15,
26-27 (опис, коментар: концепт или препис)</bibl>
     <bibl>[...]</bibl>
```

```
    </listBibl>
   </sourceDesc>
  </fileDesc>
  <profileDesc><abstract><p><persName key="Стефан
Дечански" ana="VID:16">Краљ Стефан Урош III
(Дечански)</persName> даје <orgName key="Св. Петар и
Павле на Лиму" type="church" ana="VID:17">цркви Св.
Апостола Петра и Павла на Лиму</orgName>, седишту
Хумске епископије, три цркве у Полимљу са њиховим
властелинствима на молбу <persName key="Стефан, епископ
хумски">епископа Стефана</persName>: <orgName key="Св.
Никола на Сиротињи" type="church">Св. Николу на
Сиротињи</orgName>, <orgName key="Св. Богородица
Бистричка" type="church">Св. Богородицу
Бистричку</orgName> и <orgName key="Св. Богородица
Кутањска" type="church">Св. Богородицу
Кутањску</orgName>.<ref>(Жарко Вујошевић, Српски
дигитални дипломатар)</ref>
  </p></abstract></profileDesc>
</teiHeader>
```

The example shows how an editor can deal with the limitations of the TEI elements for marking up of charter's metadata. For instance, since TEI lacks a dedicated element to explicitly identify the auctor of a charter, which information can be part of regest expressed within `<profileDesc><abstract>`, the respective concept defined in VID as nº 16 („L'auteur d'un acte écrit est la personne au nom de qui cet acte est intitulé"), can be employed. It is done through combining of the `<persName>` element, which identifies personal names, and the `@ana` attribute, which indicates any elements containing interpretation, in this case the definition of auctor in the VID. Similar steps were made for the concepts of document tradition (TEI: „filiation") and document recipient. On the other hand, when VID is lacking a proper definition, as with the lost object of diplomatic relevance such as the seal of the charter, the semantic link is established by indicating to the proposed VID additions – in this case VID:add_27 – `<sealDesc ana="VID:540 VID:add_27">`.

The next segment of the XML document, initiated with the element `<text>`, consists of the charter text itself, with a possibility for direct marking-up its different external and internal features (for instance red letters, graphical signs or specific diplomatic structure and formulas). For representing the basic charter structure in the `<body>`, the child element of `<text>`, the three-level hierarchy of `div/ab/seg` is applied, as proposed

by Vogeler and Vujošević.[32] The tag is used in order to display the basic structure of (ideal) charter – differentiating the introductory section (protocol), the „main" text (corpus, context), and the closing section (eschatocol). The <ab> tag represents a logical unit similar to a paragraph – therefore, it can be used to mark specific diplomatic clauses such as invocatio, salutatio, arenga, dispositio, sanctio, and others. The <seg> tag is used to markup shorter sections (segments) that are not completely semantically divided from other diplomatic forms. For example, this tag can be used to mark an intitulatio section in an expositio clause (as shown in the example below). During textual and paleographical markup, editors can choose between simplified and intense markup strategies, depending on their research objects and desired level of semantic granularity. Simplified markup may prioritize structural clarity, while intense markup allows for deeper analysis through detailing paleographical features and semantic enrichment.

In the following example (Listing 3), an extracted version of XML encoded (and shortened) text of the same charter as above is given. The textual content is divided into three main sections using the element (alongside corresponding diplomatic clauses). By appropriate further elements, markup is presented for personal names (<persName>), institutions („organizations") such as churches and monasteries (<orgName>), geographical regions (<region>) and figures (<figure>), such as crosses. In some cases, for narrower explanation the three attributes @ana, @type and @key are used. The attribute @ana, as already mentioned, is used to indicate elements containing interpretation, @key is used to provide standardized form of names, and @type is used to subcategorize elements, in this case to differentiate types of organizations (churches, monasteries, and political entities – such as „Municipality of Dubrovnik" as addressee of King's several charters and letters belonging to the corpus processed) and geographical places (hamlets, villages and towns).[33]

[32] Ž. VUJOŠEVIĆ – G. VOGELER, Encoding Medieval Charters, [16]; cf. Ž. VUJOŠEVIĆ, Diplomatarium Serbicum Digitale, 273–275.

[33] Same as in the footnote 31 above, the VID concepts used in the example are named here: VID:182=„Le protocole"; VID:146=„Une invocation figureé"; VID:196 bis=„Le préambule"; VID:181=„Le texte" (context); VID:197=„L'exposé" (narratio); VID:187=„La suscription" (intitulatio); VID:319=„La requête" (petitio); VID:198=„Le dispositif"; VID:182a=„L'eschatocole"; VID:244=„La clause des menaces spirituelles" (poenae spirituales); VID:246=„Corroboration"; VID:251=„L'apprécation"; VID:259 =„La signature"; VID:347=„Les notes dorsales ou marginales".

Listing 3: XML-TEI encoding of charter text with references to the VID terminology

```
<text>
 <body>
  <div ana="VID:182">
   <ab ana="VID:146"><figure><desc>+</desc></figure></ab>
   <ab ana="VID:196bis">Понєже оубо мнози начеше чинити
повѣсти ѡ извѣстованыхь вь нась вештехь, словеси
оутврьждениє, изволи се и мнѣ послѣдьствовати реченномоу
[...] светыє же мошти ихь мѵ̑ро изливають притѣкающтиимь
сь вѣрою к ныимь на исцѣлениє. Имьже кто быше
начельници, нь вьи, светыи апостоли?</ab>
  </div>
  <div ana="VID:181">
   <ab ana="VID:197">К вам же и азь вь полоучьшихь чести
избраньнихь благодѣти Светаго Доуха, рекоу же <persName
key="Стефан Немања">господина моєго светаго
Симеѡна</persName>, вамь ретника, смѣрени по милости
твоєи, владыко мои Христе, <seg ana="VID:187"><persName
key="Стефан Дечански">вседрьжавныи господинь все
срьбьскиє землє и поморьскыє, Стѣфань Оуошь
краль</persName>, сынь <persName key="Стефан
Милутин">великаго господина крала Стѣфана Оуроша, •г•
Оурошь</persName></seg>, вама се молоу, врьховнима
апостолома <persName key="Петар,
апостол">Петьроу</persName> и <persName key="Павле,
апостол">Павлоу</persName>, ѡбласть приємьша ѡть Бога
прѣблаженаıа везати и раздрѣшати грѣхи, [...] вьньмѣта ми
риданиє срьдчьно єже ѡ грѣсѣхь, некли бихь ваю ради
милосрьдиємь не имѣти никединоє скврьни ни враки
грѣховьниє на себѣ.</ab>
   <ab ana="VID:319">Сповѣдѣ ми <persName key="Стефан,
епископ хумски">єпископоупь хльм`ски Стѣфань</persName>,
ıако <orgName key="Хумска епископија"
type="diocese">єпископоупиѣ</orgName> єгова запоустѣла. Ни
врьховине не има, ни бири, никединаго доходька не има
ниѡтькоудере.</ab>
   <ab ana="VID:198">Да и ıа видехь ѡбништаниє и
запоустѣниє <orgName key="Св. Петар и Павле на Лиму"
type="church">црькьвь Светихь апостоль</orgName>, да
приложихь Светимь апостоломь: <orgName key="Св. Никола на
Сиротињи" type="church">црькьвь Светаго Николоу на
Сиротинѣ</orgName> и сь всѣми сели и сь лоудми и
заселиıами и сь мегами и сь влахи и планинами, како є ѡть
прѣднихь светихь господь било; <orgName key="Св.
Богородица Бистричка" type="church">црькьвь Госпождиноу оу
Бистрици</orgName> сь вьсѣми сели и сь лоудьми и
```

```
заселиıами и с мегами; црькьвь Госпождиноу оу <orgName
key="Св. Богородица Кутањска" type="church">Коутьньскои
Светоу Богородицоу</orgName> сели и сь лоудьми и сь всѣми
медами и сь всѣми медами и сь влахи. Приложи <persName
key="Стефан Дечански">кралевьство ми</persName> весе да
соу вь ѥдиньство подь писпоупию, ѥре <persName key="Стефан,
епископ хумски">ѥписпоупь хльмьски</persName> не има
помена доуховнаго, ни кроуга, ни бири.</ab>
  </div>
  <div ana="VID:182a">
   <ab ana="VID:244">Даште ли тко вьсхоштеть разорити и
wтнети, w горе томоу, [...] и wт богодарована ми вѣньца
да ѥсть проклеть.</ab>
   <ab ana="VID:246">Сего ради писавь и подьписахь вь
свѣдѣниѥ всѣмь.</ab><ab ana="VID:251">Аминь.</ab>
<figure><desc>+</desc></figure>
   <ab ana="VID:259"><persName key="Стефан
Дечански">СТЕФАНЬ ОУРѠШЬ</persName>ПО МИЛОСТИ БОЖИЕИ
КРАЛЬ •Г• САМОДРЬЖЬЦЬ И ВСѢХЬ СРЬБСКЫХЬ ЗЕМЛѢИ И
ПОМОРЬСКЫХЬ</ab></div>
  <div><ab ana="VID:347">Хрисовулъ. За села оу <region
key="Захумље">Захолмїю</region>, дарованнаѧ тамошнеи
соборнои церкви.</ab>
   <ab ana="VID:347">Сеи хрисовуль светаго <persName
key="Стефан Дечански">Стефана Оуроша
Дечанскаго</persName>, дань <persName key="Стефан,
епископ хумски">епископу Стефану захолмъском</persName>,
церкви его wбништавшеи и оупустѣвшїи wт безбожнихь
разбоиникь. И wписуетъ всѧ дарованнаѧ <orgName key="Св.
Петар и Павле на Лиму" type="church">церкви тои светои
светыхь апостоль</orgName>, всѧ по границами.</ab>
  </div>
 </body>
</text>
```

The displayed XML-coded content as a „raw“ dataset serves as the „backend“ of the digital edition. Its „frontend“ can be created in different programs, depending on editor's affinities and/or needs of his/her target audience. In the current phase of the MeSeCA project, a preliminary frontend is made with eXist-db, a document-oriented database system for building applications. Its simple and intended user-friendly version as well as some functionalities can be shown through screenshots of different charters (see Figures 2–5).[34]

[34] Many thanks to the project employee Thea Schaaf, University of Graz, Department of Digital Humanities, for creating the appropriate frontend version.

Research potential and aplication of expression languages (XPath and XQuery)

The research potential of digital editions is mainly related to improved accessibility, clarity and searchability of material, as well as to the possibility of linking and comparing it with other text corpora which are processed in a similar or same standardised XML format. In that sense, the presented edition in progress should help Serbian diplomatics, which still lacks analyses from a comprehensive overview and comparison. In particular, this can simplify the establishment of general criteria of charters' authenticity as well as the reconstruction of the process of diplomatic production, respectively the role and working methods of royal „chancery(/-ies)".[35]

Serving as a tool, a digital edition is primarily improving work efficiency, but its „added value" for diplomatic research lies in using digital methods, such as (but not only) different search queries. The digital environment provides many tools regarding intra-textual search and cross-document analysis, aiding the researcher in navigating the content of a digital edition. This can lead to diplomatic insights that otherwise would be difficult or even impossible to obtain, as for example in the field of dictation analysis.[36] Searching for author-specific words or phrases could help to identify persons or at least a milieu responsible for producing charter texts in particular historical context. As examples from the Serbian diplomatics here can be named the well-known phrase „participator of the Greek lands/people" („честник Грчким странам / Грком"), which was recognised by the research not only as a chronologic clue for dating charters of King Stefan Dušan during his conquest of Byzantine Macedonia and North Albania between 1343 and 1345, but also as a specific

[35] Regarding these goals, the XML-based edition is a step forward comparing to the DSD database, being already presented as a prospect (Ž. VUJOŠEVIĆ – N. PORČIĆ – D. ŽIVOJINOVIĆ, Das serbische Kanzleiwesen. Die Herausforderung der digitalen Diplomatik, *Digital diplomatics. The Computer as a Tool for the Diplomatist?*, eds. A. AMBROSIO – S. BARRET – G. VOGELER, Archiv für Diplomatik, Schriftgeschichte, Siegel- und Wappenkunde, Beiheft 14 [2014] 133–147) as well as an interim result (Ž. VUJOŠEVIĆ, Diplomatarium Serbicum Digitale, 255–275), significantly improving its research potential and functionalities, which are limited to application of basic digital methods, such as search of full-text or predefined entries/categories.

[36] For an example of applying digital methods in dictation analysis of medieval charters see E. DE PAERMENTIER, Le *dictamen* de la chancellerie comtale de Flandre et de Hainaut (1191-1244): méthode d'analyse assistée par ordinateur, résultats et potentialités, *Bibliothèque de l'École des chartes* 169/2 (2011) 385–425.

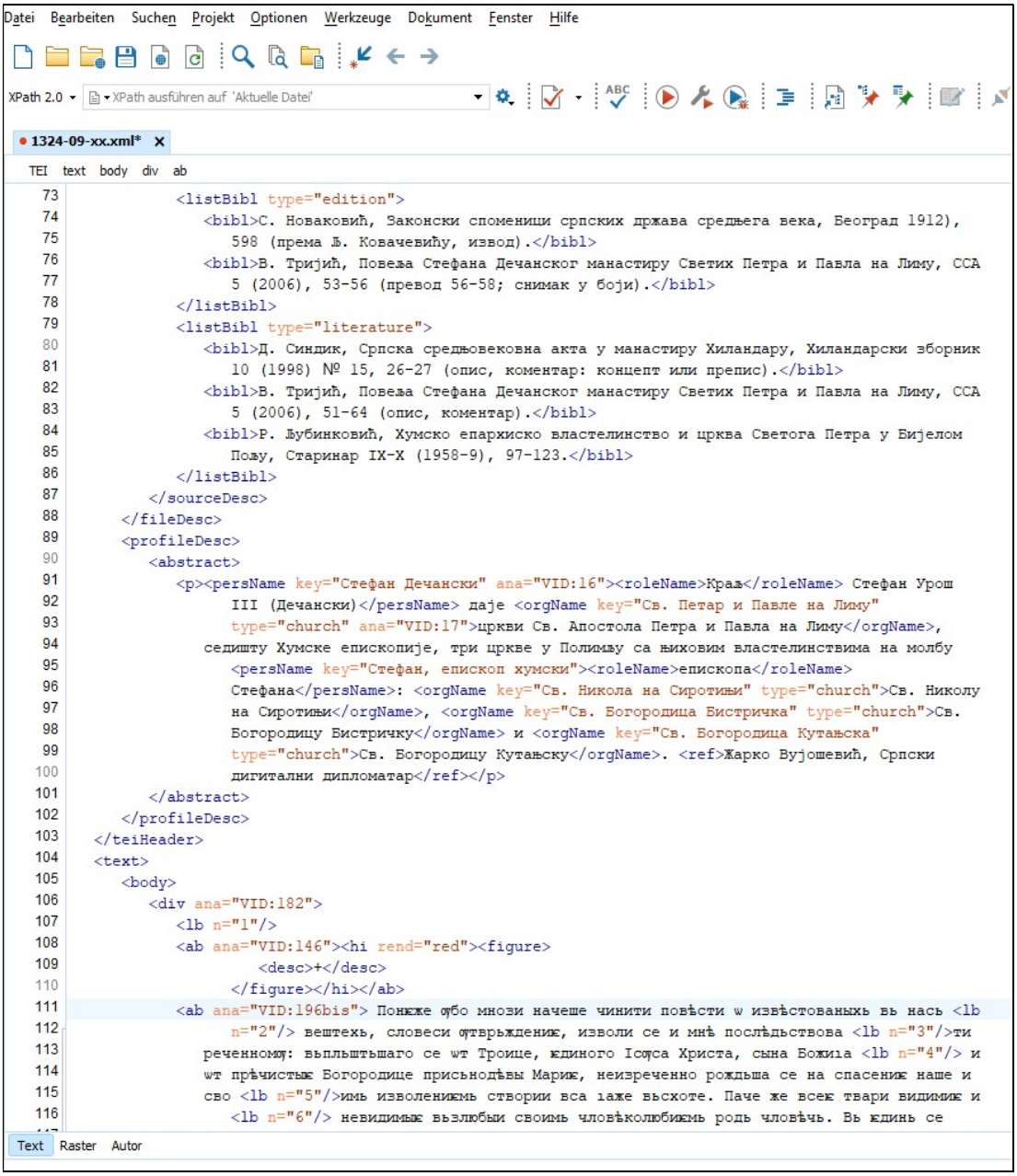

Figure 1: Part of the completely coded XML file of the charter presented in Listing 2 and 3

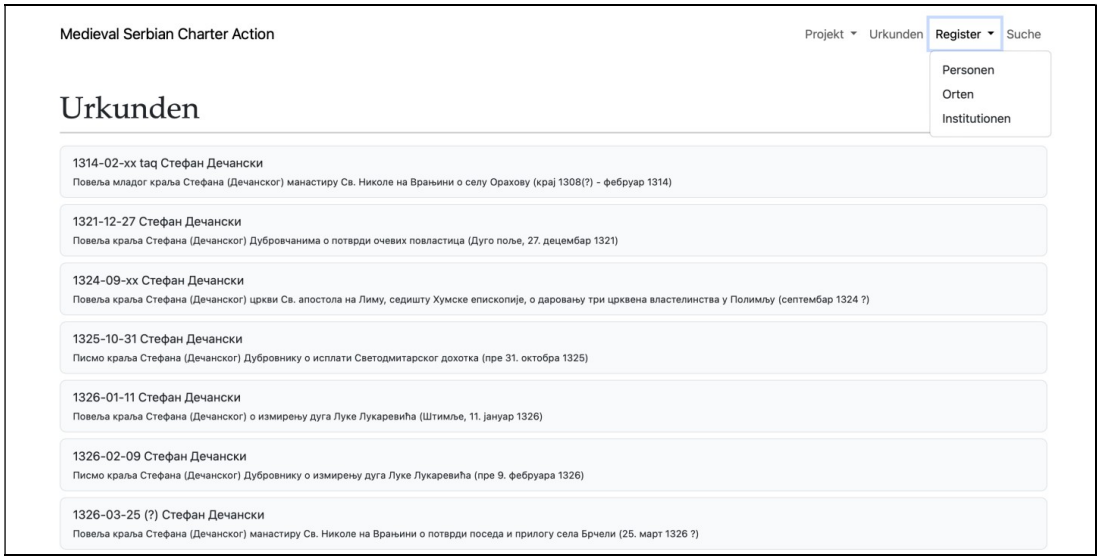

Figure 2: List of charters with short abstracts (register button above shows types of registered entities)

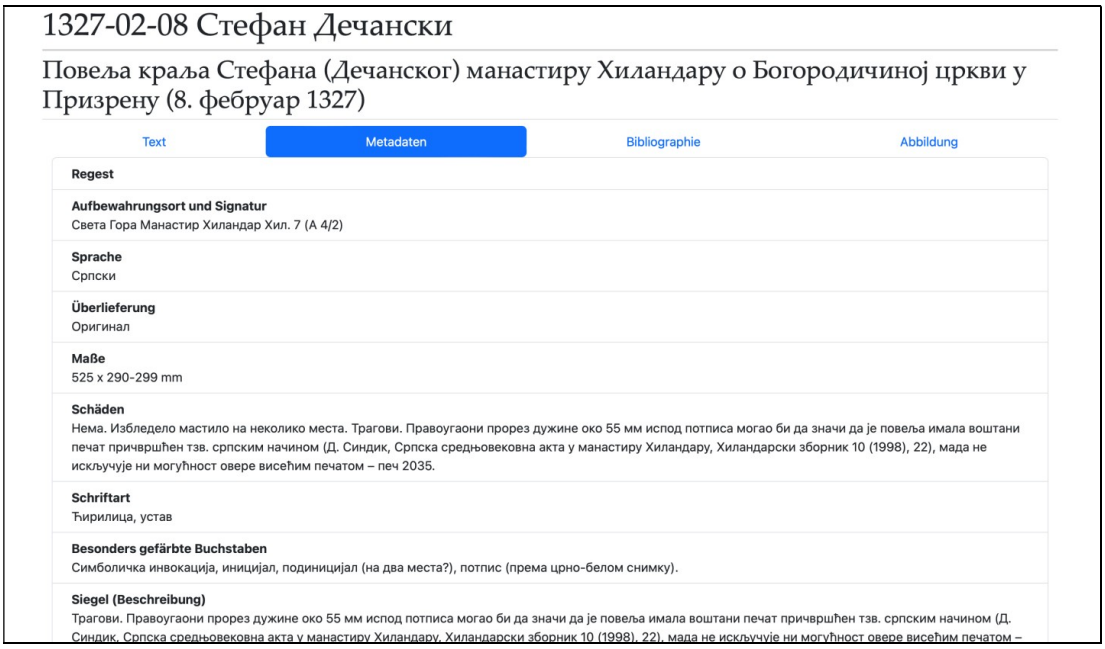

Figure 2: Metadata of a charter derived from the XML-header

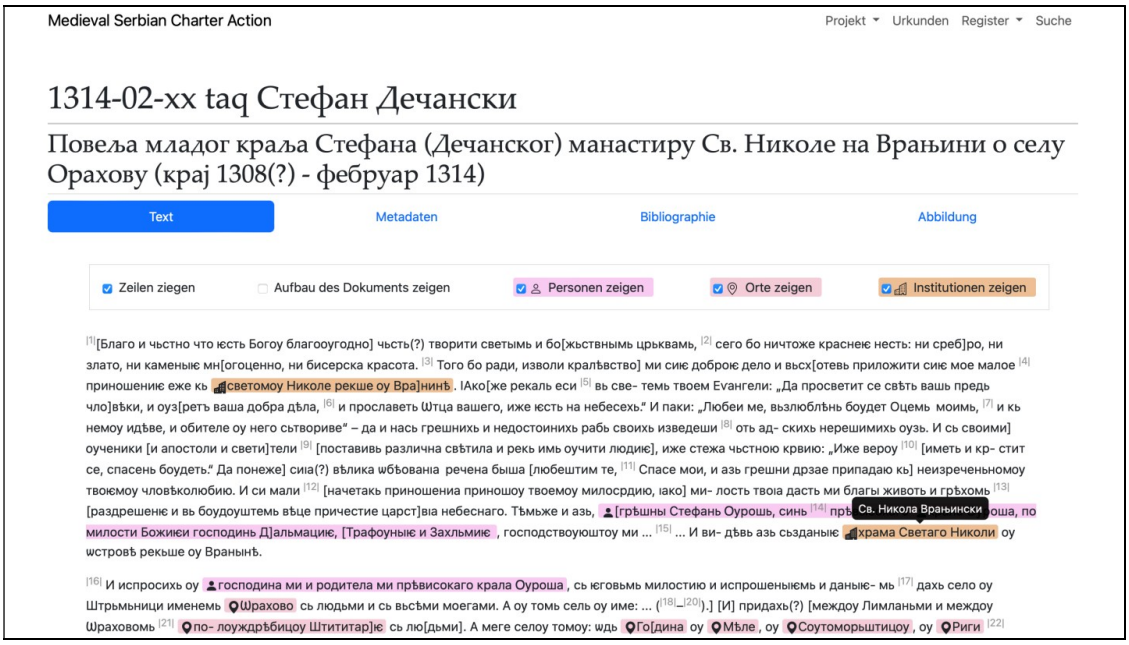

Figure 4: Charter text based on direct markup, with displayed line numbers, persons, places and „institutions“, which entities can also be shown as hypertext in form they appear in the register

Figure 5: Charter text showing diplomatic structure with VID-entries, that can be linked to the VID online presentation

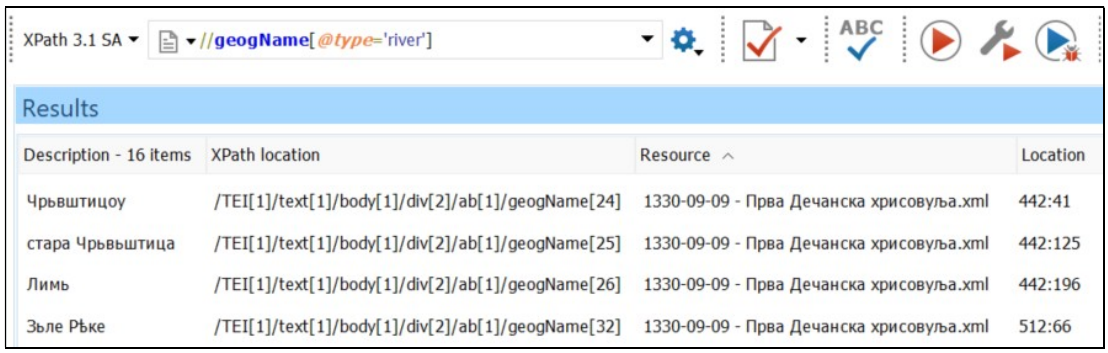

Figure 6: Excerpt from the list of rivers mentioned in the first Dečani charter

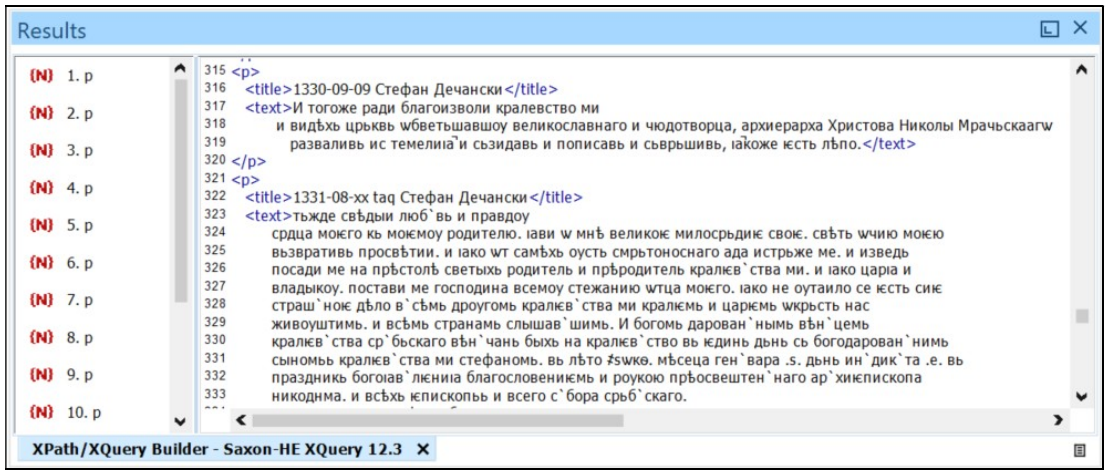

Figure 7: Extracted *narratio* from encoded charters of Stefan Dečanski (excerpt)

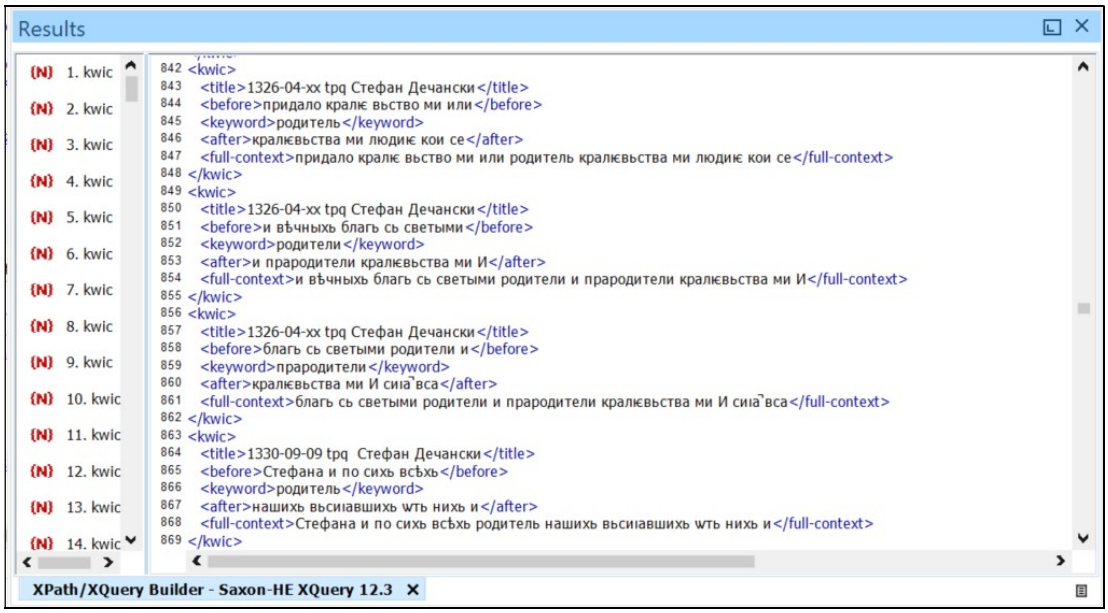

Figure 8: Implementing Key Word in Context (KWIC) on the example of “родител“ (parent/s) in the documentary corpus of Stefan Dečanski (excerpt)

formula used by Hilandar monks taking part in the diplomatic production in King's ambulant chancery; or the phrase „the first Emperor" („први цар") related also to Stefan Dušan (after his imperial coronation in 1346), which according to our preliminary research can be ascribed to the abbot of Hilandar and afterwards Patriarch of the Serbian Church Sava (IV).[37] In the following, two expression languages for querying different parts or contents of XML documents will be presented with their potentials as digital tools: XPath, an XML pointer language, and XQuery, which extends XPath for processing XML documents.

XPath, short for XML Path Language, defined by the W3C in 1999, is designed in order to operate on the hierarchical, abstract, and logical structure of XML documents, addressing their nodes (logical objects processed from tags) and node properties, such as identity, typed value, string value, etc.). Expression, a string of characters constructed from keywords, symbols, and operands, is the basic building block of XPath. Simply put, XPath uses path expressions in order to select nodes or node-trees in an XML document.[38]

For example, the simple expression `//persName` will result in selecting all `<persName>` nodes that mark personal names in an XML-TEI document, regardless of their position within its structure. Similarly, the expression `//settlement[@type='city']` targets all `<settlement>` elements with the attribute `type='city'`, demonstrating XPath's capability to filter nodes based also on attribute values. A researcher fluent in XPath can create their own expressions according to his/her needs, interests and the limitations of the path language. Proprietary software, such as Oxygen XML Editor, the digital environment of this survey, provides an integrated XPath browser, enabling its users to apply their XPath expressions across multiple XML files at the same time, facilitating inter-textual data exploration. The example in Figure 6 shows an excerpt of the list of rivers mentioned in the first charter of Dečani, based on their annotation in the document text.

In order to apply more complex expressions, another language for finding and extracting elements and attributes from XML documents can

[37] For the first example see Ž. VUJOŠEVIĆ, O datumu i tradiciji hilandarske povelje o strumičkom vlastelinu Rudlu (Prilog poznavanju „putujuće kancelarije" Stefana Dušana), *Initial* 9 (2021) 85–109, pp. 95–97. The phrase „the first Emperor" was found in four charters issued by Stefan Dušan and one issued by his successor Stefan Uroš (but also there related to Dušan), in which the Abbot or Patriarch Sava is mentioned. Based on this first insight, further research should follow.

[38] For more details on XPath see https://www.w3.org/TR/xpath20/.

be used – XQuery. It builds upon XPath, supplementing its syntax with expressions that can join, filter and transform data from multiple XML sources.[39] For instance, XPath can split text into segments, but cannot create new XML structures from derived results. On the other hand, XQuery can tokenize every word in the node string, regardless of the node hierarchy in an XML document, and construct new XML elements from the tokenized content. The aforementioned XQuery's abilities, which result in a readable and accessible output, represent a key element in constructing expressions that can aid researchers in their intra- and inter-textual analysis of charters.

In Oxygen XML Editor, the XPath/XQuery Builder is an integrated tool which enables the user to construct his/her own complex expressions. For example, if a researcher wishes to single out a certain diplomatic form from XML-encoded charters, he/she can act with the following XQuery expression:

```
declare namespace tei = "http://www.tei-c.org/ns/1.0";
for $doc in
collection('file:/C:/Users/Administrator/Documents/Maste
r%20studije/Digitalna%20diplomatika/XML/')
let $title := $doc//tei:titleStmt/tei:title[1]
let $doc-uri := document-uri($doc)
for $elem in $doc//*[contains(concat(" ", normalize-
space(@ana), " "), " VID:197 ")]
return
 <p>
  <title>{string($title)}</title>
  <text>{string($elem)}</text>
 </p>
```

This expression is constructed in order to extract the string content from every XML node with the attribute `@ana='x'` from every XML document in a folder. In this case, every narratio (VID:197) from digitally encoded charters of King Stefan Dečanski has to be extracted. As our digital environment can produce a visualization of the result, the expression lays out the format consisting of the document title and string value (text) of every node with the attribute `@ana='VID:197'`. (Figure 7)

XQuery can also be used in order to implement Key Word in Context (KWIC), an indexing method that displays a searched word with adjoining text in order to perceive it in the textual context. The following

[39] https://www.w3.org/TR/xquery-31/.

XQuery expression can extract the object of search with the n-number of words before and after the searched word:

```
declare namespace tei = "http://www.tei-c.org/ns/1.0";
declare function local:get-context($div as
element(tei:div), $keyword as xs:string) as
element(kwic)* {
 let $doc := root($div)
 let $doc-title :=
$doc//tei:teiHeader/tei:fileDesc/tei:titleStmt/tei:title
[1]/string()
 let $all-text := string-join($div//text(), ' ')
 let $words := tokenize(normalize-space($all-text),
'\P{L}+')[normalize-space()]
 for $pos at $index in $words
 where contains($pos, $keyword)
 let $start := max(($index - 5, 1))
 let $end := min(($index + 5, count($words)))
 return
  <kwic>
   <title>{$doc-title}</title>
   <before>{string-join(subsequence($words, $start,
$index - $start), ' ')}</before>
    <keyword>{$pos}</keyword>
    <after>{string-join(subsequence($words, $index + 1,
$end - $index), ' ')}</after>
    <full-context>{string-join(subsequence($words,
$start, $end - $start + 1), ' ')}</full-context>
  </kwic>
};
let $directory :=
"file:/C:/Users/Administrator/Documents/Master%20studije
/Digitalna%20diplomatika/XML/"
let $keyword := "родител"
for $file in collection(concat($directory,
"?select=*.xml"))
for $div in
$file//tei:div[descendant::text()[contains(.,
$keyword)]]
return local:get-context($div, $keyword)
```

As example here the word „родител“ (parent/s) is selected.[40] In order to extract the searched word (`$keyword`), from the XML-encoded

[40] It is recommended to use a string of characters that remain the same in every form of the word (if possible); in this case, the search for „родител“ will include various grammatical forms, but also extended ones, such as „прародител“.

corpus of documents issued by King Stefan Dečanski, every part of the node string divided by „white spaces“, creating new XML structures formatted as `$words` must be first tokenized. This expression also has functions which define how many words before (`$start`) and after (`$end`) the keyword should be extracted – in this example, it will be five (Figure 8). Contextual words will not be extracted outside the `<div>` element in which the keyword is indexed, therefore, for example, disabling the possibility of extracting a word from eschatocol to contextualize the keyword indexed in the corpus.

Conclusion

Converting results of diplomatic research into digital form has already a remarkable history, showing many benefits of using information technologies in terms of accesibility, searchability and interoperability of the processed material, as well as its long-term preservation. As XML has been proven to be a standard technology for creating digital editions, an important task in the field of diplomatics is to choose an appropriate and standardized vocabulary in order to express diplomatic phenomena within this system. The approach accepted and applied in the presented survey is not to „invent" new ontologies (which mostly remain isolated), but to use the TEI guidelines, well established as the main reference for encoding textual sources. Enhancing them with the terminology of the *Vocabulaire international de la diplomatique* as part of the semantic web is recommended as the best solution for diplomatists.

The method seems to be very well applicable to document corpora of different origin and traditions, includinig Serbian medieval diplomatic production. The given example of the digital edition of documents issued by King Stefan Uroš III (Dečanski) may serve as a „roadmap" also for encoding charters in general, as it allows an analytic and comprehensive presentation of research results, bearing functionalities that can improve diplomatic studies. The strong wish of authors is to put it into discussion within the community of scholars, in order to open new perspectives for diplomatics in the context of digital humanities.

Жарко Вујошевић – Павле Николић

Перспективе дигиталних издања дипломатичког материјала на примеру докумената краља Стефана Уроша III Дечанског (1321–1331)

Резиме

Примена компјутерских метода у хуманистичким наукама одавно није новост. У последњих двадесет и више година значајни резултати постигнути су и на плану обраде и презентације дипломатичке грађе, при чему нарочити искорак представљају дигитална издања средњовековних докумената. Као и „класичне“ (штампане) едиције, она доносе критичку интерпретацију изворног текста, али то „чине“ посредством рачунарског програма, који је носилац не само садржаја, већ и функција (као што су претрага, систематизација, визуализација и др). Осим што повећавају доступност дипломатичког материјала, дигиталне едиције тако могу унапредити и сам истраживачки поступак.

Дигитално издање подразумева маркирање (анотацију) текста документа, као и одговарајућих метаподатака, чиме се релевантне информације генеришу у виду компјутерски читљивих кодова. Међу рачунарским маркационим језицима и схемама за кодирање текстуалне грађе као стандардни у већини хуманистичких дисциплина намећу се XML (Extensible Markup Language) и TEI (Text Encoding Initiative), потискујући старије (SGML, HTML) и специјализоване (CEI) системе. Широка примена TEI схеме отвара могућности за упоредну анализу различитих текстуалних жанрова, али њен веома општи карактер захтева додатна решења за дефинисање недостајућих дипломатичких концепата (нпр. ауктор/дестинатар, традиција документа и др) и формула (нпр. аренга, санкција, потпис и др). С тим у вези предложена је допуна TEI номенклатуре терминима вишејезичког лексикона *Vocabulaire international de la diplomatique* (VID), етаблираног у свету дипломатике и у међувремену доступног у дигиталном облику. Теоријски и практично образложени метод XML-TEI-VID кодирања дипломатичког материјала у раду је приказан на примеру корпуса повеља и писама краља Стефана Уроша III (Дечанског), обрађеног у оквиру пројекта Diplomatarium Serbicum Digitale, а који се састоји из 22 повеље и писма, сачуваних у 27 примерака разноврсне типологије и традиције.

Посредством неколико одабраних повеља представљени су изглед и садржај XML „позадине“ (backend) са кодираним метаподацима и одговарајућим деловима текста појединачних докумената, као и „интерфејс“

(frontend) дигиталног издања читавог корпуса са функцијама приказа и претраге маркираних елемената као што су личности, географски појмови, дипломатичка структура документа, формуле и др. Истраживачки потенцијал овако приређене дигиталне едиције додатно је тестиран применом компјутерских алата XPath и XQuery, уз чију помоћ анализа текста може бити знатно ефикаснија и комплекснија него уз примену традиционалних метода. Као пример истакнуте су специфичне фразе у српском дипломатичком материјалу (као што су добро позната „честник грчким странам“ краљевских повеља Стефана Душана, и „први цар“ у неколико повеља издатих након 1346), чија би анализа у одговарајућем контексту допринела бољем разумевању порекла и околности настанка појединих докумената. Оба поменута навигациона језика, XPath и XQuery, омогућују претрагу и упоредну анализу кодираних елемената (лична имена, титуле, места и установе, дипломатичке формуле или специфичне фразе), док се уз помоћ овог другог може извести и претрага кључне речи у контексту (Key Word in Context – KWIC). Њихова употреба приказана је кроз примере као што су инде-ксирање помена река у првој дечанској хрисовуљи, издвајање дипломатичког састојка нарације (narratio) из свих докумената Стефана Уроша III и преглед употребе речи „родитељ“ у њеном непосредном лексичком окружењу кроз читав корпус.

Предложени методи израде и коришћења дигиталне едиције у истраживачке сврхе овде су изнете на дискусију широј заједници проучавалаца не само српске, већ документарне грађе уопште. Осим што могу бити модел за једно свеобухватно издање српских средњовековних повеља и писама, они се препоручују као допринос разматрању циљева и могућности дигиталне дипломатике у веома разуђеном и све динамичнијем спектру дигиталне хуманистике.

Кључне речи: Дипломатичка грађа, дигитално издање, повеље и писма Стефана Дечанског, XML, Text Encoding Initiative, Vocabulaire International de la Diplomatique, дигиталне истраживачке методе.

Чланак примљен: 30. септембра 2025.
Чланак прихваћен: 12. новембра 2025.