Şermin KALAFAT

(Assoc. Prof. Dr., Visiting Scholar), University of Cambridge, Cambridge Digital Humanities / United Kingdom, sk2360@cam.ac.uk, 0000-0003-1000-0402

Triggering A Renaissance in Historical Textual Studies in Turcology: Text Encoding Initiative (TEI) and the Initial Standardised Ottoman Mathematics Text[*]

Türkoloji'deki Tarihi Metin Çalışmalarında Bir Rönesansı Tetiklemek: Metin Kodlama Girişimi (TEI) ve İlk Standartlaştırılmış Osmanlı Matematik Metni

Abstract: This article seeks to answer a critical question about the impact of digital transformation, which global technological and digital developments have made a necessity, on historical textual studies in Turcology. The article draws attention to the need for digital transformation in historical textual studies and the lack of literature on the use of TEI in Turcology. The research considers historical textual studies as an object of digital cultural heritage, digital humanities, and digital scholarship. It is evaluated through the Text Encoding Initiative (TEI) standard. The aim is to establish a symbiotic relationship between historical textual studies and the Text Encoding Initiative (TEI) and offer a new perspective on the latest digital environment for studying Turkish manuscripts and the future of this research. It also explains how manuscripts (especially in the Ottoman field) should be encoded at TEI to enable interdisciplinary cooperation in digital editions. A case study of an 18th-century Ottoman mathematical text was encoded in TEI, and the process is explained in detail for those working in Turcology and those who want to do text editions with this method. The article's results are shared with the reader by presenting a Github link. These results constitute a new source for the international literature on digital encoding of historical texts and set standards for digital text editions to be used in historical text studies in Turcology. In this way, it aims to trigger a renaissance promoting the use of TEI in historical textual studies and to take the first step towards its eventual transformation into a renewed field of research.

Keywords: Text Encoding Initiative (TEI), Digital historical text editions, Mathematics, Ottoman Turkish, Digital Ottoman Studies, Digital Turcology and Cultural Heritage

Öz: Bu makale, küresel teknolojik ve dijital gelişmelerin bir zorunluluk haline getirdiği dijital dönüşümün Türkoloji'dekitarihî metin çalışmalarına etkisine dair eleştirel bir soruya cevap aramaktadır. Makale, tarihî metin çalışmalarında dijital dönüşüme duyulan ihtiyaca ve Türkoloji'de TEI kullanımına ilişkin literatür eksikliğine dikkat çekmektedir. Araştırma, tarihsel metin çalışmalarını dijital kültürel miras (digital cultural heritage), dijital beşerî bilimler (digital humanities) ve dijital akademik araştırmaların (digital scholarships) bir nesnesi olarak ele almakta ve Metin Kodlama Girişimi (TEI) standardı üzerinden değerlendirmektedir. Amaç, tarihsel metin çalışmaları ile Metin Kodlama Girişimi (TEI) arasında simbiyotik bir ilişki kurmak ve Türkçe el yazmalarını incelemek için en son dijital ortam ve bu araştırmanın geleceği hakkında yeni bir bakış açısı sunmaktır. Ayrıca el yazmalarının (özellikle Osmanlı sahasında) dijital neşirlerde disiplinlerarası iş birliğinin sağlanması için TEI'da nasıl kodlanması gerektiğine dair bir bilgi akışı sunmaktadır. Bu amaçla bir 18. yüzyıl Osmanlı matematik metni TEI'da kodlanmış ve Türkoloji sahasında çalışanlar ve bu yöntemle metin neşri yapmak isteyenler için süreç ayrıntılı olarak açıklanmıştır. Makalenin sonuçları, Github linki sunularak okuyucuyla paylaşılmıştır. Bu sonuçlar, tarihi metinlerin dijital kodlanması konusunda uluslararası literatüre yeni bir kaynak oluşturmakta ve Türkoloji'deki tarihi metin çalışmalarında kullanılacak dijital

metin neşirleri için standartlar belirlemektedir. Böylelikle, tarihi metin çalışmalarında TEI kullanımını teşvik eden bir rönesansı tetiklemeyi ve nihai olarak yenilenmiş bir araştırma alanına dönüşmesi için ilk adımı atmayı amaçlamaktadır.

Anahtar Kelimeler: Metin Kodlama Girişimi (TEI), Dijital tarihsel metin neşri, Matematik, Osmanlı Türkçesi, Dijital Osmanlı Çalışmaları, Dijital Türkoloji ve Kültürel Miras

Introduction

The study of historical texts in Turcology is currently at an important crossroads in the context of ongoing digital transformation. How will this transformation, which is being driven by shifting supply-demand dynamics and the ensuing new academic balances, affect the future of this field of study's methodologies? If the response to this inquiry is predicated on the erroneous notion that progress in artificial intelligence, particularly in tools facilitating the comprehension and translation of historical texts, has rendered this type of research obsolete, or that the digital revolution in Turcology will be executed exclusively by a select few individuals and solely overseen by computer and software engineers, then, it is imperative to initiate a fresh discussion. A more profound examination reveals that the issue must be addressed within digital transformation, digitalisation, cultural heritage, the humanities, and the continuity of science. This perspective plays an essential role in initiating a critical and sensitive dialogue concerning the publication and analysis methods of historical text studies in Turcology, which will shape the coming period.

I. Current Status of Historical Textual Studies in Turcology

Historical textual studies constitute a broad research area in Turcology. Those working in different disciplines, language, literature, history, art, Islamic literature, history of science, information-document management, sociology, psychology, librarianship, architecture, medicine, pharmacy, mathematics, astronomy, chemistry and so on, complete the pieces of a mosaic by analysing textual materials belonging to the field of Turkishness. These studies focus on the Ottoman field, where manuscripts constitute a qualified heritage, and the Ottoman Empire stands out for its political and intellectual breadth. In addition, Eastern Turkish texts offer a research area at least as crucial as Ottoman manuscripts. This field contains valuable textual material covering different types of writing and a wide range of thematic content. Thus, Turcology has not only been concerned with studying Ottoman-oriented manuscripts. Manuscript textual research in both Turkish areas is comprehensively addressed in different fields of specialisation.

I.1. Digital Transformation in Historical Textual Studies

In the last few years, historical textual studies in Turcology have moved towards a new landscape within the digitalisation and digital transformation framework. However, this has not yet attracted much attention. If we add digital cultural heritage, digital humanities, and the concept of digital scholarship as the umbrella field of these two fields, historical textual studies will gain a place in this triangle. Although these three digital fields seem very close to each other regarding their content, similarity of their aims, case studies based on their methods, and project examples, they are all research areas of different communities.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines digital heritage as a combination of natural and cultural heritage. According to UNESCO, digital heritage is "computer-based materials of enduring value originating from different communities, industries, sectors and regions that should be preserved for future generations".[1] On the other hand, cultural heritage is defined as "sites, objects and intangible assets of cultural,

---

[1] "Concept of digital heritage", *UNESCO*, 2018, accessed: 5 March 2022, https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-heritage.

historical, aesthetic, archaeological, scientific, ethnological or anthropological value for society and the individual"[2] Consequently, digital cultural heritage facilitates the digitisation, preservation, processing, management, and accessibility of tangible or intangible cultural items. While digital humanities, formerly known as humanities computing, was initially shaped by fields such as linguistics and coding,[3] over time, it has spread to many fields such as museology, art history, history, new media, and archaeology and formed an identity with digital cultural heritage studies.[4] In more specific terms, digital humanities refers to using digital tools, methods, and technologies to study humanities disciplines such as literature, history, philosophy, linguistics, and cultural studies. This encompasses applying computational techniques to analyse and interpret cultural artefacts, texts, and data and creating digital resources for scholarly research, teaching, and public engagement in the humanities. It can be argued that almost every work done in this field, directly or indirectly, has contributed to digital cultural heritage.

Digital scholarship, which can be considered an umbrella term for these two fields, can be broadly defined as one that embraces the methods and functions of these fields with broader disciplines and represents digital pedagogy[5]. It refers to using digital technologies and methodologies to research various academic fields, disseminate findings, collaborate with others, and engage with audiences. It encompasses the humanities disciplines and the social sciences, natural sciences, engineering, and other academic fields.

Moreover, new approaches to support the development of research methods, enhancing digital efficiency, preserving resources, new documentation activities, and intellectual property are research directions that characterise these three fields holistically. In light of these definitions, it can be posited that Turcology is, in fact, more closely aligned with digital scholarship.

When manuscript texts are selected as the subject of study, it becomes evident that these domains are significantly interconnected. When a manuscript is digitised as a cultural heritage item requiring preservation, it serves as both a digital cultural heritage artefact and a subject of analysis in digital humanities and digital scholarship. The conversion of manuscripts into databases, the broadening of their applications, and the development of new techniques by specialists working on these materials are standard practices in research domains in collaboration with the field of digital humanities and digital scholarship.

I.1.2. Challenges and Solutions in Applying TEI to Turkish Manuscripts

There is a notable gap in the Turcological literature regarding procedures for digitising manuscripts and digitally transforming approaches used in historical textual studies. In Turcology, as in other humanities fields, historical textual studies face the challenge of incorporating new digital tools and methodology while maintaining the depth of traditional academic methods. Furthermore, advancements in artificial intelligence and the growing

---

[2] "Concept of digital heritage", *UNESCO*, 2018.

[3] Susan Hockey, "The history of humanities computing," *in A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, and John Unsworth (Oxford: Blackwell Publishing, 2004), accessed 11 April 2022, https://doi.org/10.1002/9780470999875.ch1; Patrik Svensson, "Humanities computing as digital humanities," (Digital Humanities Quarterly, 2009), accessed: 11 April 2022, http://www.digitalhumanities.org/dhq/vol/3/3/000065/000065.html.

[4] For an additional source in Turkish, see, Fatma Aladağ, "Dijital beşerî bilimler ve Türkiye araştırmaları: Bir literatür değerlendirmesi." *Türkiye Araştırmaları Literatür Dergisi* 18, no. 36 (2020): 773–796, accessed: June 2021, https://dergipark.org.tr/tr/pub/talid/issue/58073.

[5] Johanna Drucker, "Humanistic theory and digital scholarship," *Debates İn The Digital Humanities, Part II. Theorising The Digital Humanities*, (2012), accessed: 7 December 2022, http://dhdebates.gc.cuny.edu/debates/text/34; James P., Purdy, ve Joyce R. Walker. "Valuing digital scholarship: Exploring the changing realities of digital work", *Profession, Modern Language Association*, (2010), 177–195, accessed: 11 April 2022, https://www.jstor.org/stable/41419875.

influence of natural language processing applications indicate that Turcology should prioritise interdisciplinary approaches and adapt its methodologies. Vertan and Reimers and Soualah and Hassoun argue that it is crucial to develop a digitisation-compatible framework in order to incorporate manuscripts into the digital communication network[6].

To address this scenario, it is crucial to become familiar with the concept of digital text editions to digitise manuscripts and join the digital transformation network with updated methods. Digital text editions are scholarly editions explicitly created for online use in digital humanities. A scholarly edition refers to a reliable and consistent interpretation of a text that can recognise differences found in manuscripts and offer contextual information through footnotes or comments. A digital edition allows you to compare different document versions, change the appearance of annotations or corrections, and access additional supplementary content. Some digital versions may include both audio and video elements. Digital versions may use text analysis or text mining for comparison and other activities. The use of digital text editing is essential for the preservation, study and distribution of historical manuscripts. By digitising historical manuscripts, this approach not only preserves these valuable cultural artefacts but also increases their accessibility to a broader audience and enables multidisciplinary research.[7] Using digital text editions in Turcology can significantly transform how historical texts are studied and understood. According to Manoff, every digitised text has the potential to introduce novel concepts, analyses, theories, and tools that shape how we generate, receive, and engage with knowledge in the context of advancing digital publication and academic research.[8]

The utilisation and propagation of the Text Encoding Initiative (TEI) standard, which is directly associated with digital text editions, is imperative for the examination of Turkish manuscripts. As the most efficient digital humanities tool for creating a digital text edition, TEI has already proven its widespread and reliable use.[9] At the outset, it might be argued that the application of TEI in historical textual studies in Turcology could yield quick and advantageous results and adapt traditional approaches to modern technologies, as TEI can be easily used as a long-standing tool for historical text editions and semantic analyses. Contrary to expectations, this process is more complex than it seems. This is because the TEI coding procedure is primarily designed for texts in Western languages. For Turkish, especially when dealing with historical writings, it is crucial to manage the two steps effectively: disseminating the basic requirements to prepare for encoding and retesting the encoding for each new text to ensure standardisation. The latter is a lengthy process that can only be realised after a large number of Turkish manuscripts have been encoded. The former, on the other hand, needs to be put in place as a matter of priority. To elaborate on this point a little further, in traditional methods text editions differ in each specialisation and this hinders the interdisciplinary use of a text. For example, a manuscript text in specialised fields such as history, history of science and Turkish language is transcribed differently. If a text from one field is to be used in another, it has to be completely retranscribed. Researchers recognise that this technique is quite laborious and requires a considerable amount of time. Similarly, even in Turkish language and literature there

---

[6] Cristina Vertan and Stefanie Reimers, "A TEI-Based Application For Editing Manuscript Descriptions" *Journal of the Text Encoding Initiative*, 2 (2012), 1, accessed: 11 April 2022, https://doi.org/10.4000/jtei.392 , https://journals.openedition.org/jtei/392; Mohammed Ourabah Soualah and Mohamed Hassoun, "A TEI P5 Manuscript Description Adaptation For Cataloguing Digitised Arabic Manuscripts", *Journal of the Text Encoding Initiative*, 2, no. 2 (February 2012), 2, accessed, 11 April 2022, https://doi.org/10.4000/jtei.398, https://journals.openedition.org/jtei/398

[7] Patrick Sahle, "What is a Scholarly Digital Edition?" *In Digital Scholarly Editing*: *Theories and Practices*, (Cambridge, 2017), 19-25, accessed: March 2022, https://books.openedition.org/obp/3397.

[8] Marlene Manoff, "The Materiality of Digital Collections: Theoretical and Historical Perspectives", *Libraries and the Academy,* 6, no. 3 (2006), 311, accessed: January 2022, http://hdl.handle.net/1721.1/35689.

[9] One indisputable proof of this is the awarding of the Antonio Zampolli Prize to TEI in 2017, which recognizes individuals and communities that have made notable contributions to the field of digital humanities (see https://tei-c.org/).

are different transcriptions. Although this may seem a simple problem, few researchers have fully addressed the issues of textual corrections, comparative analysis of copies, genealogical tracing, the presence or absence of marginal notes, edition criticism and interlingual translation that make up an edition and accompany transcription. There is also a notable lack of consistency in typography, the symbols used for words that cannot be read due to the difficulty in deciphering the old script, the display of footnotes, and how verses and hadiths should be incorporated into the text.[10]

Although the limitations of classical approaches have been disregarded and the importance of interdisciplinarity has been ignored, the use of TEI in the digitisation of manuscripts emphasizes the necessity of multidisciplinary usability in digital text editions. Preserving and displaying cultural treasures and disseminating open data in databases are two of the many components that make up digital text editions.[11] Standardising and communicating the fundamentals of transcription, typography, and printing is critical to guarantee the compatibility of any encoded text. Following this method will ensure high-quality data and make Turkish manuscripts more usable. Nevertheless, TEI should not be solely regarded as the process of converting a transcribed text into a certain format. Transcription, similar to traditional working methods, is a first step that must be undertaken in order to facilitate analysis. TEI enables the execution of a wide range of analysis, including those that may be performed using traditional methods, as well as additional analyses.

Ultimately, whether it is done digitally or not, a historical textual study primarily focuses on the processes of production and manipulation of a cultural artefact. Therefore, the importance and continuity of textual studies in Turcology will continue in the digital age, as it has to date. However, given the ongoing digital transformation, this scenario also serves as a cautionary reminder that we need to bring a constructive mindset to our field of study regarding the methods and approaches by which historical text publication and related research will continue in the digitalisation process. Therefore, the main aim of the research is to outline standard features and standardisation measures that should be taken into account to ensure interdisciplinary cooperation when encoding manuscripts in Turcology using TEI. Furthermore, an Ottoman mathematical manuscript is presented as a case study as an example of how TEI can be applied in the context of manuscripts. This encoding experience serves a pedagogical purpose to show how historical textual studies in Turcology, which is still very unfamiliar with TEI, can be integrated with TEI (Text Encoding Initiative) and pave the way for its effective use in graduate programs.

II. Text Encoding Initiative (TEI)

In parallel with the increasing use of computers and digital tools, text encoding is a process that started in the second half of the 19th century. Two text encoding concepts can be mentioned in digital infrastructures. Their distinction is essential. The first is a coding standard called ASCII (American Standard Code for Information Interchange), created in 1960 for electronic storage and processing of texts and started the text coding process. It represents text characters with 7-bit binary numbers, consists of 128 characters and is thus supported by many

---

[10] See also, Şermin Kalafat, "Türkoloji araştırmaları kapsamında yazma eser çalışmalarına yönelik bir inceleme", *The Second International Congress of Turcology*, "600 Years of Polish-Turkish Relations: Research on Turcology Yesterday, Today, and Tomorrow," Warsaw, Poland, June 25, 2014; Berat Açıl, Sadık Yazar, Kadir Turgut, Özgür Kavak, *Tenkitli Neşir Kılavuzu (Osmanlı Türkçesi Metinleri İçin)*, (İstanbul: İSAM, 2022).

[11] Murray Pittock, "The Scottish Heritage Partnership Immersive Experiences: Policy Report", (University of Glasgow,   2018:   5), https://www.gla.ac.uk/schools/humanities/research/informationstudiesresearch/researchprojects/scottishnationalheritage/#/policyreport:immersiveexperiences; Stuart D. Lee, *Digital Imaging: A Practical Handbook*, (London: Facet Publishing, 2001), 160; Tolga Çakmak, "Kültürel bellek kurumlarında dijitalleştirme: Ankara'daki 17 kurumun uygulamalarının analizi", *Ankara Araştırmaları Dergisi* 6, no. 2 (2018), 168.

computers and software.[12] However, this standard is suitable for English and cannot describe characters from other languages. Therefore, in the 1980s and 1990s, especially in the humanities, specialised coding standards were developed for different languages and character sets. These form the second structure and differ from ASCII in their purpose. Indeed, calls for reusability, interchange, system and software independence, portability and collaboration in the humanities were answered with the emergence of the Standard Generalised Markup Language (SGML), which became an ISO standard in 1986.[13] SGML is a standard generalised markup language for structuring, tagging and formatting text-based documents. It uses a set of tags to describe documents' content, structure and relationships.[14] Charles Goldfarb developed it for IBM's "Document Composition Facility Generalised Markup Language".[15] Following this encoding standard in 1986, TEI (Text Encoding Initiative) attracted attention as a new standard in 1987.

The Text Encoding Initiative (TEI) is a text encoding standard that arguably spans various research areas. TEI (Text Encoding Initiative), which can be translated into Turkish as *Metin Kodlama Girişimi*, is an initiative launched in 1987 by a community of researchers in the United States of America to develop, maintain and disseminate hardware and software independent methods for encoding machine-readable text sources in the humanities. The community includes universities, librarians, and academics. It is supported by the Association for Computers and Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing. The TEI community published its first guidelines (Text Encoding Initiative) in 1994. Since the beginning of the 2000s, the TEI community, which has become more widespread internationally, has sought to create an independent resource by using the TEI's text encoding system as a kind of digital archiving system and, at the same time, making the sharing of these texts as fast and easy as possible. Thus, the guidelines in the TEI's guidelines have been redefined in XML (Extensible Markup Language), the most appropriate text working form for the TEI's guidelines, and have become a digital archiving resource.[16] The new formation of TEI via XML is significant. To explain this briefly for those unfamiliar with software issues, XML is a subset of SGML. As a communication standard designed by W3C (World Wide Web Consortium), it was created to be readable by man and machine. It offers reliable, accessible, comprehensive possibilities for creating and sharing digital data sets, archives, and mega data sources. You can define your language in XML and create a system for your work by building hierarchical relationships between the components of your language. In this respect, it is one of the most helpful exchange languages that can provide information circulation between applications with different grammatical structures. This process can be better seen in the picture below.

---

[12] ASCII, see accessed: 3 March 2022, https://www.ascii-code.com/.

[13] ISO 8879 "SGML Standard 1986," *ISO*, accessed: 3 March 2022, https://www.iso.org/standard/16387.html; Charles. F. Goldfarb, The SGML handbook (Oxford: Clarendon, 1990); Edward Vanhoutte, "An Introduction to the TEI and the TEI Consortium", *Literary and Linguistic Computing*,19, no. 1 (2004), 10, accessed: 11 April 2022, https://www.edwardvanhoutte.org/pub/2004/teillc.htm.

[14] SGML Standard 1986.

[15] Charles, F Goldfarb, 32.
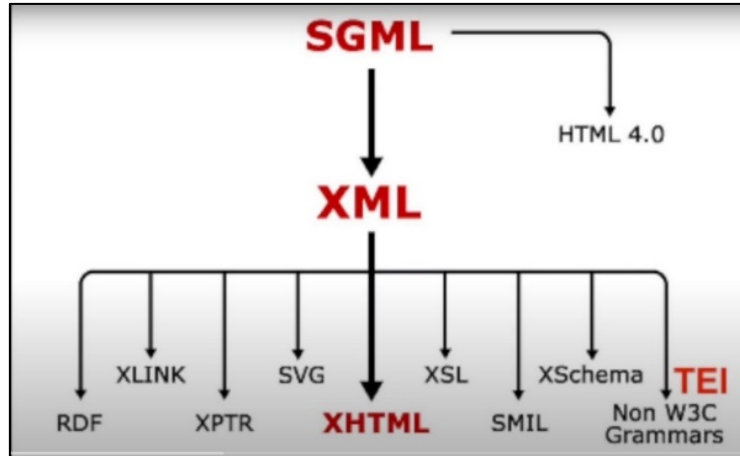
[16] Edward Vanhoutte, 12.

Figure 1. TEI

The TEI guidelines include best practice standards of text markup and guidelines representing a theoretical approach. They were developed to address the digital representation and analysis of complex textual resources. Therefore, it can be considered as a method and coding philosophy.[17]

The theoretical framework of TEI consists of XML-based coding systems that encompass components and attributes describing the structure and content of texts as well as their grammatical, historical, cultural and literary features. TEI facilitates comparative analyses, the creation of grammatical and functional definitions, etymological analyses, the elaboration of rhyme and redaction studies in poetry, the creation of dictionaries and indexes, syntax analyses and the standardisation of structural and semantic aspects of language. This encoding philosophy brings together structural, semantic and contextual factors, enabling TEI to serve as a tool for the linguistic study of digital texts. Using TEI to encode historical manuscripts enhances research by providing a structured framework that facilitates data processing, annotation and cross-comparisons. Researchers can thus gain deep insights into linguistic patterns, historical contexts and cultural nuances, fostering more efficient and innovative research approaches. Furthermore, TEI XML's independence from any programming environment reduces the risk of data loss and enables seamless text analysis across various systems. Supported by an expanding academic research community, TEI fosters interdisciplinary collaboration through the integration of textual reproduction and analysis.[18]

---

[17] C. Michael Sperberg-McQueen, "Textual criticism and the Text Encoding Initiative." *Paper presented at the annual convention of the Modern Language Association*, (San Diego, California: December 1994), accessed: 11 April 2022, https://tei-c.org/Vault/XX/mla94.html; Allen Renear, "Text Encoding," *in A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, and John Unsworth (Oxford: Blackwell Publishing, 2004), accessed: 11 April 2022, https://companions.digitalhumanities.org/DH/?chapter=content/9781405103213_chapter_17.html; Sarah L. Pfannenschmidt and Tanya E. Clement, "Evaluating digital scholarship: suggestions and strategies for the text encoding initiative," *Journal of the Text Encoding Initiative,* 7 (November 2014), 7; accessed: 21 March 2022, https://journals.openedition.org/jtei/949, https://doi.org/10.4000/jtei.949; James Cummings, "The Text Encoding Initiative and the study of literature" in Ray Siemens & Susan Schreibman (Eds.), *A Companion to digital literary studies* (Oxford: Blackwell Publishing, 2007) accessed: 7 March 2022, https://doi.org/10.1002/9781405177504.ch25.

[18] Lou Burnard, *What is the Text Encoding Initiative? How to add intelligent markup to digital resources* (Marseille: OpenEdition Press, 2014), 12, accessed: 7 March 2022, https://books.openedition.org/oep/426, https://doi.org/10.4000/books.oep.426,

---

II.1. TEI's system and code layout

The coding tags for encoding text generated via TEI-XML are based on elements that activate these tags, attributes, and script subtags that enable using attributes according to the material. Although these structures have commonalities with XML and other web languages such as HTML, CSS, Java, Xquery, Xquery Json, and Typescript, the limited combinations of other languages can be extended since XML has its unique tag design. At this point, although TEI is created on XML, other web software languages are needed in the publishing process. While TEI is not independent of these, it should be noted that the community has created templates and an advanced manual to facilitate the use of these structures together.[19]

It is necessary to explain the systemic functioning of TEI by relating it to the term DTD (Document Type Definition). A document type definition (DTD) is a set of rules and guidelines defining how a document will be processed per TEI standards, i.e., the document's structure, content, and formatting. The verifiability, reproducibility and shareability of the digital structure of the document depend on these. DTD defines the elements that may be needed in any text, from essential to sophisticated tags. Tags for images, facsimiles, drawings, and calculations are to be added to the text. In addition, depending on the field of study, it may be possible to create a tag that is not previously defined in TEI or used for different purposes. TEI leaves the researcher free in this regard. Thus, each new creation will contribute to the factual processes of TEI, so much so that these additions or innovations in different coding formats have created a literature network on TEI.[20]

Through DTD, the tags used and the hierarchy created with these tags vary according to the type of text at hand (such as manuscript, poem, story, novel) and the purpose of processing the text (e.g., grammar, terminology, semantic constructs, concept relationship networks, rhyme schemes, etc.). In TEI, tags can be used for all kinds of text analysis operations usually performed, such as focusing on morpho-syntactic structures in the text, creating a rhyme scheme, defining terms, coding meaning changes and defining the thematic structure of the concept area based on the text, conducting functional studies, adding dictionary, index, etymological definitions, interpreting the relationship networks in stories and novels. It is essential to know which tag does what and to use it for what purpose. This is measured by your mastery of web design languages, especially XML, in the background for each desired operation. We have learnt from long hours of work on this subject. However, knowing that even those who have been using TEI internationally for many years are not familiar with all the labels in the TEI manual gives you the confidence to create your limitations in your original research method.

Regarding textual studies in Turcology, as poetry and prose language can differ, specialising in their genres also creates originality. The first experience of a researcher who has worked on a genre in poetry for a long time in the face of a prose text may cause hesitation initially, as it will show differences in method and approach. This is also the case for TEI. Coding a poem versus coding a prose text, creating a text that requires text repair, or each new perspective chosen for text analysis will create differences in the application of labels, attributes, and other web languages, even among those coding the exact text. It is all a matter of the researcher's ability and how deep he or she can go in the research.

---

[19] The majority of the content discussed in this section is extensively elucidated on TEI's official website. This section provides a concise overview of the information for researchers who are unfamiliar with the text encoding in Turcology, with the intention of making the knowledge more applicable and useful."Text Encoding Intiative," Text Encoding Intiative: TEI, accessed: June 2021, (See also https://tei-c.org/).

[20] Cristina Vertan and Stefanie Reimers, "A TEI-based application for editing manuscript descriptions," Journal of the Text Encoding Initiative 2 (2012),5, accessed: 11 April 2022, https://doi.org/10.4000/jtei.392 , https://journals.openedition.org/jtei/392.

III. Material and Method

The case study, which is the method of the article, can be considered a learning technique in which the researcher is confronted with a specific problem or situation. The case study facilitates the investigation of a genuine issue within a defined context using various data sources. It is a research approach used to examine and understand a particular situation, event or phenomenon in detail.[21] The primary material of the case study is a mathematics text from the late Ottoman period. The selection of a specific Ottoman mathematical text in this article aims to draw attention to neglected texts in this field and, since these texts contain mathematical operations, to create an important model to demonstrate the encoding of these operations in Turcology texts to be encoded in TEI. The research is analysed within the framework of this case study. The methodology of this approach requires the interweaving of TEI principles and their application to the analysis of digital historical texts with the principles of Turcology. Thus, it also allows for exploring the limits of TEI's capabilities while encoding Ottoman mathematical text. The encoding process is classically based on a manuscript prototype for simplicity and clarity. Accordingly, the encoding result again takes the form of transcription, intralingual translation, dictionary, and facsimile edition.

In the article, the step-by-step coding process is described in detail. This is to provide straightforward guidance to interested researchers and graduate students. Each stage has been carefully considered, starting from a meticulous transcription of the original text at the beginning to a translation that overcomes language barriers. A facsimile is also included so readers can visually connect with the historical text. Since mathematical texts are rarely analysed, a glossary of terms is provided to help understand technical nuances. In addition, the necessity of encoding notation and formulae with the text is clarified, as text-specific problems require modern solutions. The case study, which embodies the research methodology, was prepared in the context of a mathematics text written by Yenişehirli Mehmed Esad Efendi in the 18th-century Ottoman period. The subject of the text is the introduction and application of a mathematical tool called Sine. Yenişehirli Mehmed Esad Efendi wrote various treatises on the use of mathematical tools..[22] The selection of the text was based on its brevity (2 drafts), which would allow it to be presented in the article.

IV. Preparation phase

IV.1. Transcription of the Manuscript

The transcription of texts is of great importance both in traditional studies and in the digitisation of texts. In order to increase the applicability of results to other disciplines, it is crucial to perform a comprehensive transcription of texts. There are various methodologies for text transcription in digital applications, the three most commonly used being:

a) Transcription using traditional method: This is a highly reliable strategy for researchers working with texts. Regardless of the technique, all transcriptions should be subjected to cross-validation according to the specified protocol. This study used the described text transcription method to conduct the research.

b) Performing text transcription with optical character recognition (OCR): This technique can be implemented using two separate methods. Its primary function is to facilitate the conversion of text from typewritten works, especially theses from the 1980s or 1990s, into

---

[21] See also, Pamela Baxter ve Susan Jack, "Qualitative case study methodology: Study design and implementation for novice researchers," *The Qualitative Report*, 13, (2008), 547, accessed: 27 July 2022, https://nsuworks.nova.edu/tqr/vol13/iss4/2/, https://doi.org/10.46743/2160-3715/2008.1573.

[22] Since Atilla Polat and Halime Mücella Çavuşoglu have written a detailed article on the author and his works, this information will not be given here Atilla Polat ve Halime Mücella Demirhan Çavuşoğlu, "Mehmed Said Efendi'nin Misâha Risâlesi". Osmanlı Bilimi Araştırmaları 21, 2 (July 2020): 249-70, accessed: 23 March 2021, http://dx.doi.org/10.26650/oba.650926, https://iupress.istanbul.edu.tr/en/journal/oba/home.

word processors. This significantly benefits the user by eliminating the need to transcribe text from scratch. However, it is essential to be aware of the limitations of this process. The physical condition of the text used in optical character recognition (OCR) is critical. Various problems such as smudges, dark images, smearing and paper quality can affect resolution and prevent the text from being transferred. Furthermore, in the case of Ottoman manuscripts, where reading accuracy can vary from 80% to 88%, it is paramount to analyse the text in depth to obtain accurate statistics, given the altered appearance of the transcribed characters. Therefore, this approach seems more challenging than the first strategy, especially for those with a deep understanding of text comprehension approaches and significant expertise in this field.[23]

c) Use of crowd-based transcription system: The use of switching systems requires users to participate in the management of significant amounts of data and to adopt a community-oriented approach. This method encourages individuals to meticulously analyse a digital copy of the original text and accurately transcribe each character or word. The transcription process, therefore, depends on people's cognitive capacities in terms of visual perception and verbal comprehension. Researchers use keying systems to optimise the process of searching, indexing and accessing collections. The participation of volunteers facilitates the data entry process and encourages information sharing. However, this technique also requires specific precautionary considerations. The data provided by readers should be subject to rigorous and thorough scrutiny. This approach can be likened to the process whereby master's students are given a substantial corpus of information to read. Then, this corpus is assessed by their supervisors. This can be conceptualised as an immersive experience involving over a hundred students and guaranteeing public participation. Furthermore, choosing this methodology may raise ethical concerns in the comprehensive analysis of written material. Data entry staff may be unable to thoroughly scrutinise the content, resulting in only the first version being available. However, a competent researcher may be able to correct and utilise this data. Therefore, the qualifications, identity and other relevant information of the people involved in the reading process should be considered when adopting such a system. It is also important to include appropriate annotations to guarantee adherence to intellectual property rights.[24]

After deciding on the text transcription method, the font choice for transcription can be started. Today, *Munevver*, *Oktay New Transcription* (ONT), *Timesefras*, and *Timestrans* are used in classical text studies in Turcology. These typefaces contain deficiencies in providing some transcription marks. For example, since there is no lisp s in Münevver, the researcher has to use an underscore for each s in Word; in ONT, the sign for the back palatal n is given in this way (ñ). The marking that gives the actual value of the sound in the IPA is the hooked version (ŋ). Although this is a point that researchers sometimes do not pay attention to, it is now accepted in Turcological research that these two signs characterise different sounds. Using the hooked form of the sound in the IPA system has become widespread. However, despite this, it is a fact that different disciplines do not pay attention to this. Insisting on such a usage would be a critical error and a significant limitation for the interdisciplinary use of texts, as it would lead to an incorrect representation of the sound. Unfortunately, a researcher who wants to publish such a text must redo the transcription of this sound. This would require an additional time-consuming process for the researcher, and careful comparison with the main text would be necessary to keep this process under control.

---

[23] See also, Dölek, İshak, ve Atakan Kurt. "Derin Sinir ağlarıyla Osmanlıca Optik Karakter tanıma". Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi 38, 4 (April 2023), 2579-94, https://doi.org/10.17341/gazimmfd.1062596, https://dergipark.org.tr/en/pub/gazimmfd/issue/75069.

[24] Victoria Anne Van Hyning, ve Mason A. Jones. "Data's Destinations: Three Case Studies in Crowdsourced Transcription Data Management and Dissemination". Startwords, 2 (01 October 2021) accessed: 9 March 2022;Crowsourching,https://www.qub.ac.uk/researchcentres/CentreforDataDigitisationandAnalysis/ServicesandOu tputs/ToolkitsandGuides/Crowdsourcing/; Tim Blackman, What is Crowd-Based Translation: A beginner's Guide, April 25, 2018, https://www.brombergtranslations.com/crowdsourcing-translation-crowd-based-translation/ accessed: 9 March 2022.

Osmanlı Mirası Araştırmaları Dergisi / Journal of Ottoman Legacy Studies
Cilt 11, Sayı 30, Temmuz 2024 / Volume 11, Issue 30, July 2024

270

Concerning the font chosen for transcription, there is always the concern that the text is displayed correctly. If a particular typeface is not installed on your computer or your text is not a PDF file, the transcription marks may be distorted, and the text may become unreadable. This is also the case when creating TEI-XML files. In digital publishing, Unicode-compliant Gentium and Cora typefaces should be preferred so that transcription marks are not corrupted in XML and other web languages. However, according to the data obtained from our experience, a general problem has been identified while encoding the LISPs. It is known that the transcription sign of the lisp s, which appears as "underlined s" in the specified typefaces, may appear as "underlined t" in a distorted form in the code line. Therefore, in order to overcome the problem, the attribute assignment can be provided and encoded with a numeric value in Unicode. This will ensure that the transcription copied from the web view or XML file is also in usable format in Word.

An important point to be mentioned in transcription is the closed e sound. The representation of this sound is still controversial in historical textual studies. In classical Western Turkish texts, it is usually written as /i/ because the sound has been moved to a variant sound position different from its original sound position in Eastern Turkish, has lost its distinctive feature of meaning, and does not contain an additional feature that has the value of a sign or symbol with which it establishes an indicative relationship. This is not to deny the existence of the sound but to ensure the consistency of the marking system with the transcription. Of course, some studies favour the (ė) marking. Since this dual use is justifiable, it cannot be considered erroneous in classical studies and is preferred in the Turkish language and literature. However, in digital editions, the situation requires a little more caution. The algorithm is a system based on cycles. Since digital edition studies are not only an edition but will also be used in the production of future language technologies and language modelling, different values of the same word may cause cyclic errors in the software programs to be produced, making coding difficult. Although these problems can be solved with an additional implementation method, the data used always has the potential to generate errors in the analysis results. Therefore, it is necessary to say that the use of /i/ is mandatory in Western Turkish texts do not pose a problem in future technological adaptations due to its lack of meaning discrimination, lack of a marker and its status as a variant sound, as previously mentioned.

The use of closed (ė) is compulsory and widespread in the texts of Eastern Turkish since the sound has a sign, and its meaning distinguishing feature continues. The differences in alphabets used in Eastern Turkish and the fact that sounds can be inferred from signs and meanings in Arabic-letter texts make this transcription sign compulsory in digital editions. In the final analysis, the (ė) sign should be preferred for closed e in digital editions in Eastern Turkish. Thus, the spelling of closed e in Eastern and Western Turkish texts can be differentiated by coding. In my opinion, this is a parsing that researchers will find safe as an endeavour that makes the historical linguistic trace more meaningful in future studies and is believed to gain value in the long run. For this reason, it is brought to the attention of researchers as an essential and obligatory suggestion to ensure the standardisation of the coded texts in this way. For this reason, it is brought to the attention of researchers as an essential and obligatory suggestion to ensure the standard in the coded texts in this way.[25]

Another issue that is overlooked in transcribed texts and often creates problems in terms of interdisciplinarity is that the letters p and ç at the end of Turkish words and suffixes are written as b and c in Arabic letters, so they are read as alub (آلوب), aġac (آغاج). However, the rule that there are no letters b, c, d, or g at the end of words and suffixes in Turkish should also come into play here, and such words should be read as "olup, agaç".

---

[25] Over time, some differences will inevitably emerge in the coding methods of the Eastern and Western fields. Nevertheless, this is a subject of study in itself. The closed "e" sound was taken into consideration because it is a distinctive element in the texts of the two research fields and because it is controversial in historical texts.

In many manuscript texts, it is frequently witnessed that the same word is written with a dot on the top line but not on the bottom line. This is because Arabic letters should be written without raising the hand along the line, i.e., the dots should be added later. After this process, the transcription script is ready for encoding.

IV.1.2. Editor Choice

Editors are powerful tools designed to manipulate text, regardless of the programming languages involved. The choice of an appropriate editor for modelling data is a critical decision. Different editors have varying capabilities for developing programming languages, with some offering extensive features while others may be more limited. For TEI, Oxygen is typically the editor of choice. However, it is possible to use other tools for coding, such as Notepad and Dreamweaver. In recent times, Oxygen, eXide, and Visual Studio Code (VSC) have gained considerable popularity. VSC offers several advantages over other editors. Its compatibility with Java and Python libraries makes it an excellent choice for coding large projects.

On the other hand, developers commonly use Oxygen, an XML editor, to create these types of structures. Although Oxygen requires payment, VSC is available for free. However, TEI has created eXide as an editor. The article provides illustrative examples of the use of editors such as Oxygen and VSC for working with TEI-XML, and it also explains the usage of eXide when necessary.

IV.1.3. TEI Schema Selection

Choosing a TEI schema that adheres to TEI standards once the editor has been selected is paramount. TEI has developed a plethora of schemas to enhance the usability of diverse types of text materials. The appropriate schema can be selected based on the structure and purpose of the content to generate relevant code blocks. This article showcases a comprehensive architecture that seamlessly integrates manuscripts and mathematical calculations. The tags in the code blocks will vary according to the selected schema. Figure 2 below illustrates the All schema in its simplest form.

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_allPlus.rng" type="application/xml" 
3  <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_allPlus.rng" type="application/xml"
4      schematypens="http://purl.oclc.org/dsdl/schematron"?>
5  <TEI xmlns="http://www.tei-c.org/ns/1.0">
6    <teiHeader>
```

Figure 2. Link to Code Block Schematic [All] (Oxygen)

IV.2. Coding

IV.2.1. Metadata Input

The term' metadata' is used in the data context to describe or offer information about other data. Metadata is distinct in that it provides a set of qualities, properties or descriptors of a particular dataset or piece of information, and it is used to categorise data, search it and retrieve and understand it.

Data management, integration, cataloguing and information sharing frequently employ metadata. The data may encompass several details, including the dataset's name, description, creation date, source, format, size, classification and other relevant information. The data referred to as "metadata" is located within a TEI document's <teiHeader> element. The <teiHeader> section contains metadata about the document, including the author, title, publication date, language, source information, descriptions, and other descriptive and administrative elements associated with the text. It encompasses the bibliographic details of the text material awaiting coding. This framework encodes both the effort exerted and the

researcher/programmer responsible for its digital representation, as well as the publication location of the digital data. It is important to note that the metadata portion is not visible in web publications, namely in HTML and XHTML file conversions. From an ethical standpoint, this establishes a binding agreement between the author and the publisher to store this digital data. Nevertheless, it is essential to acknowledge that TEI possesses a versatile framework. Thus, it is feasible to personalise the presentation of metadata. It is possible to modify or personalise configuration files in the chosen file format for online display, incorporating metadata throughout the publication process. This can be achieved by adjusting the code or incorporating CSS properties.

```
5  ▽ <TEI xmlns="http://www.tei-c.org/ns/1.0">
6  ▽   <teiHeader>
7  ▽     <fileDesc>
8  ▽       <titleStmt>
9              <title>Title</title>
10           </titleStmt>
11 ▽       <publicationStmt>
12            <p>Publication Information</p>
13           </publicationStmt>
14 ▽       <sourceDesc>
15            <p>Information about the source</p>
16           </sourceDesc>
17         </fileDesc>
18       </teiHeader>
19 ▽     <text>
20 ▽       <body>
21            <p>Some text here.</p>
22           </body>
23         </text>
24       </TEI>
```

Figure 3. Metadata Defined Code Block (Oxygen)

Figure 4 showcases an expanded version of the code block in Figure 3, incorporating extra details. The tags are used in conjunction with the tag, and an inline CSS property is applied to ensure that the title is displayed in red. Information about the author is provided in the <author> tag, with the author's name specified using the <persName> tags, and the date of the author's death added using the <date> tag. The "when" attribute helps the date tag function. The publicationStmt tag includes the code for the publisher, which is indicated through the addName and settlement tags. The sourceDesc tag is used to describe the source text. The text is a manuscript; information is conveyed through the "Manuscript Description" (msDesc) tag below this designation.

The xml:id attribute within this tag serves as a unique identifier for the msDesc tag. The information enclosed within quotation marks is subject to the researcher's discretion. The term "MS1" is used as a mnemonic device to represent "manuscript 1." This reference is used for identification in other sections of the text. Using concise and numeric qualifiers improves memory retention and streamlines the coding process. The xml:lang attribute is used to specify the language of the content within the <msDesc> element, which is set to "osmT". Language codes are essential for establishing a standardised way to indicate the specific language in which a document is authored. This will streamline the process of reading code. After examining the sub-tags within the msIdentifier tag, the settlement tag stores the name of the specific geographical location where the relevant manuscript is located. In the past, this element was included in the <publicationStmt> tag. Tags should be used appropriately based on the specific need at hand. The <repository> tag indicates the location where the document is stored. The following <idno> tag indicates the unique identifier or shelf number assigned to the work. The identifying type is indicated by the "type" element and is defined as "shelfmark." The <msContents> tag indicates the relevant information present in the text. This is similar to the informative data found in library records. The relevant information is contained within the <summary> tags. The history element represents the manuscript's archival data. The <origin> tag is paired with a <when> property to establish the date and time of the manuscript precisely.

```xml
<TEI xmlns="http://www.tei-c.org/ns/1.0">
<teiHeader>
  <fileDesc>
    <titleStmt>
        <title><hi rend="bold" style="color:red;">Risāle-i Sinüs li-Misāḥati'l-bu'd</hi></title>
            <author>
                <persName> Müft'iżāde-i Yenişehrî Mehmed Saîd Efendi </persName>
                <date when="1767"> 1767 (18.yy) </date>
            </author>
        </titleStmt>
    <publicationStmt>
        <publisher>
            <addName>  Doç.Dr. Şermin KALAFAT </addName>
            <settlement> istanbul Medeniyet Üniversitesi / Türk Dili ve Edebiyatı Bölümü / Yeni Türk Dili</settlement
        </publisher>
    <date when="2023-04-10"> 10 Nisan 2023 </date>
    </publicationStmt>
    <sourceDesc>
        <msDesc xml:id="MS1" xml:lang="osmT" >
            <msIdentifier>
                <settlement> İstanbul </settlement>
                <repository> Topkapı Sarayı Müzesi</repository>
                <idno type= "shelfmark">Hazine nr. 609/1 (8v-10r) </idno>
            </msIdentifier>
            <msContents>
                <summary>"Uzaktaki yerlerin mesahasını ölçmekte kullanılan sinüs adlı bir matematik
                    aletinin yapımı ve kullanımı hakkında bir risaledir."</summary>
            </msContents>
            <history>
                <origin when="1154"> 1741 (18.yy)</origin>
            </history>
        </msDesc>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

Figure 4. Metadata Defined Code Block (Oxygen)

The section where the text is entered is crucial for the design of the text. Furthermore, this section allows for including various elements, such as text, calculations, pictures, maps, and more, based on personal preference. This section is the text portion coded between <body> </body> tags. By observing Figure 2, it becomes evident that there is solely a 'paragraph' tag present. Div tags are commonly used to partition content to shape a webpage's appearance based on your text or interpretation. Figure 5 presents an expanded version, for instance. The text layout is formed by the tags between the <body> tags, as well as the <div></div> tags. The <head></head> and <p></p> tags are located between them. These tags are fundamental HTML codes and are also defined in XML. You can include the <figure> tag, the <graphic> tag, and the "url" attribute. The tags used here are for the parts that form the comment line, where the author or researcher records notes and follows up on the steps. These are not efficient when the text is displayed. The <pb/> tag marks the start of a page. The attribute n represents the numerical value assigned to pages, chapters, sequences, etc.

```xml
<text>
    <body>
        <pb n="1"/>
        <!-- first page starts here -->
        <div>
            <head><!-- heading of first section  --></head>
            <p>some text here</p>
            <figure>
                <graphic url= "http://www.tei-c.org/logos/TEI-glow.png"/>
            </figure>
        </div>
        <div n= "2">
            <head> <!-- heading of second section --></head>
            <p> some text here </p>
        </div>
    </body>
</text>
```

Figure 5. Extended Version of TEI's Body Section (Oxygen)

IV.2.2.1. Coding of the Transcription Section

Given the subject matter of this study, namely mathematics, we intend to demonstrate the various approaches employed by different editors in this context.

Figure 6 demonstrates the utilisation of the "type" element positioned directly within the <div> tag. This attribute specifies the section type as "section". The attribute "n" represents the numeric value assigned to a section. The "head" tag indicates the title of this section. Additionally, the "n" attribute is utilised in this context to specify the numerical value of the title. The "rend" element in the header determines the visual style of the header, resembling a CSS effect, with properties like bold font and colour. The "style" property provides additional information regarding the style or visual presentation of the title. The section is marked as "Part I: Transcription". Next, a new <div> element is opened. The <div> tag creates subsections within the main section. The value of this tag can be adjusted based on personal preference. The value of this section is determined by the attribute "n". The <pb> tag is used to select the final page number. Within this context, the "n" attribute signifies the page number's numerical value. In this situation, the value is clearly defined as n= "8v.".

The <figure> element is commonly utilised to represent an image or a geometric shape. The <graphic> tag is used to indicate the file reference. On the other hand, the "url" element is utilised to specify the image file. The file name of the facsimile stored on your computer is surrounded by double quotation marks, like "MS1_1v-1r.JPG". It is worth mentioning that facsimiles can be utilised in PNG and JPG formats. It is necessary to consider the specific image in question in order to provide an accurate answer. The second <head> tag is used for subheadings specifically for page titles. The indicator in the header shows whether the sheet is the front or back side of sheet 8v (or 8b). The <quote> tag is commonly used within this tag to incorporate quotations, critical additional comments, verses and hadiths, or introductory sections. In this coding example, the prayer section at the beginning is emphasised. Instead, you can use independent <div> tags to display introductory sections or similar comments rather than relying on this tag. This reflects the researcher's personal preference. In order to identify the expression within this tag, the "type" element needs to specify a particular type of expression once again. Here, the type is specified as "prayer". As explained previously, paragraphs within <div> tags are numbered using the "n" attribute. The "xml:lang" element within the <p> tag is used to indicate the language of the paragraph. These qualifiers help differentiate between various languages in multilingual text. As an example, the text could be written in Ottoman Turkish. In certain sections of the internal chapters, it may be necessary to utilise Arabic, Persian, or other languages exclusively. Here, "Ar" stands for the Arabic language. Comment lines have been included in the text to enhance clarity and comprehension of the codes, making it easier for other researchers to navigate.

```
<text>
    <body>
        <div type="chapter" n="1">
            <head n="1" rend="bold" style="color:black;"> I. Bölüm - Transkripsiyon </head>
        <div n="1v"> <!--1. varak-->
            <pb n="8v"/> <!--1.varak ön yüz-->
                <figure>
                    <graphic url="MS1_1v-1r.JPG"/>
                </figure>
            <head n="1v" rend="bold" style="color:black;"> [8v] </head>
            <quote type="dua">
            <p n="1" xml:lang="Ar">
                <seg n="1"> (1) Elḥamdulillāhü'l-bāḳī elleẕī lā yezāl lā cennül mezḫāl ilā ḥāl ve  </seg>
                <seg n="2"> (2) es-ṣalātu ve's-selām 'alā nebiyyehü'l-meb'ūs bi-eṣref el-ḥışāl ve 'alā ālihi</seg>
                <seg n="3" xml:id="p1Satır3"> (3) ve eṣḥābihi ve ehl-i beynehu ve āmennehu ecma'īn.</seg>
            </p>
            </quote>
```

Figure 6. Beginning of body part (Oxygen)

The <seg> element includes pieces of text within <p> tags, as shown in Figure 7. The <seg> element refers to a specific section or fragment of text. It is imperative to understand the importance of using tags such as <l> for line encoding. However, these will organise the lines in a vertical order. The goal is to create blocks of text in paragraph structure, but we need a tag that allows parallel sentence organisation. Consequently, the <seg> tag serves a practical purpose. The attribute "n" once again specifies the line numbers within the tag.

```
<p n="2" xml:lang="OsmT">
    <seg n="3" xml:id="p1Satır3_cont">
        <hi rend="bold" style="color:red;"> ve ba'd </hi> sinüs dimekle</seg>
    <seg n="4"> (4)ma'rūf olan āleti resm eylemek murād olınsa gerek pirinç ve gerek </seg>
    <seg n="5"> (5) ağaç her ḳanġısından olursa rāst gönye şeklinde dü pāy </seg>
    <seg n="6"> (6) olup bir pāyı bir pāyından dirāz ola. Ol dirāz olanı ḳāme</seg>
    <seg n="7">(7) farż idüp meşelā <hi rend="bold"> و</hi> yāḫūd <hi rend="bold"
        >ﻭ</hi> yāḫūd <hi rend="bold"> ﻼﺑ</hi>bunlardan</seg>
    <seg n="8"> (8) her birisine ale's-seviyye ḳısmet idüp per-gār müteġayyir olmadan
        yine ḳısmet-i</seg>
```

Figure 7. Body (Oxygen)

Figures 8 and 9 show a solution to a common problem in text printing. Often, sentences can be complex or lengthy, causing words to span lines or sentences to extend into separate sections. These sentences can be challenging to analyse when determining specific numerical values of index divisions, especially in the traditional survey arts. This can lead to challenges in formatting the text and ensuring that sentences flow smoothly. To overcome this problem, we applied the following strategy: When a line is extended, the first part of the sentence is recognised by assigning it an "id" number. According to Figure 10, line 22 continues on page 3. In order to maintain compatibility with the "id" in line 22 and to ensure that it continues smoothly, a new identifier is assigned to the continued section. Firstly, "v1p2Line22" is used, followed by the abbreviation "cont" to indicate the continued section. This method is used consistently to link sentences.

```
<p n="1" xml:lang="OsmT">
    <seg n="1"> (1) ḍarb u ḳısmet ile ḥāṣıl olur ve ṭarīḳ-i ṣānī nisbet ile </seg>
    <seg n="2"> (2) olur. Ḍarb u ḳısmet ile olan ṭarīḳ budur ki meşelā "Bir şey'üŋ</seg>
    <seg n="3" xml:id="v1p1Satır3"> (3) bu'duŋ bilinmek murād olınsa?"</seg>
</p>
<p n="2" xml:lang="OsmT">
    <seg n="3" xml:id="v1p1Satır3_cont"> <add>Cevāb:</add> Sinüsi bir düz maḥallde vaż'</seg>
    <seg n="4"> (4) idüp ve 'iżādenüŋ ḫālī olan başını ḳāme üzerine rast</seg>
    <seg n="5"> (5) vaż' idüp ve sinüsi ol bu'dı bilinmek murād olınan şey'i</seg>
```

Figure 8. Encoding of combining continuous sentences (Oxygen)

```
<p n="3" xml:lang="OsmT" >
  <seg n="22" xml:id="v1p2Satır22_cont"> Meşelā: "Bir cānibinden bir cānibe varınca ne miḳdār</seg>
  <seg n="23" xml:id="v1p3Satır23"> (23) ẕirā'dur ẕirā'lanması mümkin olmaya bir nehr-i kebīrüŋ ḳarşu?"</seg>
    </p>
<p n="4" xml:lang="OsmT">
  <seg n="24"> (24) <add>Cevāb:</add> Cānibinde bir ḥacer olsa, ve cānibinden bir sinüsi bir maḥalle vaż'</seg>
  <seg n="25"> (25) idüp hedefeteynden ol ḥaceri görüp sinüsi</seg>
  <seg n="26" xml:id="v1p3Satır26"> (26) vaż' olınan maḥalle bir 'alāmet vaż' eyledük. Ba'dehū</seg>
</p>
```

Figure 9. Encoding of combining continuous sentences (VSC)

Figure 10 shows the sections without any text or needing to be restored with text. Utilising the <add> tag to incorporate words, sentences, symbols, or expressions is highly beneficial. At this stage, the researcher must note the sections marked with the <add> tag and ensure that they are coded in the appropriate locations within the text transfer/interlanguage translation section. Ensuring consistency and accuracy of meaning is crucial.

```
<p n="4" xml:lang="OsmT">
   <seg n="24"> (24) <add>Meŝelā:</add>"Ve eger ol buʿdeynüŋ miḳdārı murād olınan şeyʾi sinüsüŋ</seg>
   <seg n="25"> (25) ikinci vażʿ olınan maḥallden ne miḳdār?" diyü suʾāl olınursa</seg>
   <seg n="26"> (26) yine keʾl-evvel ḍarb u ḳısmet ve nisbet ile olur.
      <add>Cevāb-ı evvel:</add> Ḍarb u ḳısmet</seg>
```

Figure 10. Encoding of the text repair section (Oxygen)

IV.2.2.2.Coding Translation Section (Transfer/Intralingual Translation)

This segment uses the same lines of code for content transfer. Thus, we can replicate the framework to ensure seamless inclusion of translation segments. However, it is essential to consider a few key points. When performing classical analyses, word-for-word translation becomes difficult due to conjunctions such as "ki, çün, dahı", complex verse structure, lyrical fragments and rearranged sentence structure. In transcriptions, line numbers are typically shown as "(1), (2)..." in translation sections, but they are usually shown as whole numbers enclosed in square brackets []. For example, the notation [1-5] indicates the task of translating sentences numbered 1-5. This has been carefully considered in the coding process. Applying this standard in digital versions is essential for creating indexes and glossaries.

In addition, sections requiring repair are presented using the same <add> tag. This time, however, translation lines are also provided. This guarantees consistency when switching between scripts.

```
</div>
   <div n= "2r">
      <pb n="9r"/> <!--1.varak arka yüz-->
      <head n="2r" rend="bold" style="color:black;"> [9r] </head>
      <p n="1" xml:lang="ModernT">
         <seg n="1-3"> [1-3] çarpma ve bölme ile bulunur ve ikinci yol da orantı ile olur. Çarpma ve bölme ile olan yöntem şudur:
            Problem: "Bir şeyin uzaklığı bilinmek istenirse?"
         </seg>
      </p>
      <p n="2" xml:lang="ModernT">
         <seg n="3-22"> [3-22] <add>Cevap:</add>Sinüs aletini düz bir yerde ortaya çıkarıp cetvelin boş hali olan başını dikme üze
         </seg>
```

Figure 11. Coding of the in-language translation section. (Oxygen)

```
</div>
   <div type="chapter" n="2">
      <head n="2" rend="bold" style="color:black;">II. Bölüm - Metin Çevirisi</head>
      <div  n="1v"> <!--1. varak-->
         <pb n="8v"/> <!-- 1. varak ön yüz-->
         <head n="1v" rend="bold" style="color:black;" > [8v] </head>
         <quote type="dua">
         <p n="1" xml:lang="T">
            <seg n="1-3"> [1-3] Sonsuz olan Allah'a şükürler olsun ki, O'nun dışında her şey yok olup gi
         </p>
         </quote>
         <p n="2" xml:lang="modernT">
            <seg n="3-24">
            [3-24]  ve sonra, sinüs demekle bilinen aleti resmetmek istense gerek pirinç ve gerek ağa
            </seg>
         </p>
```

Figure 12. Coding of the in-language translation section. (VSC)

Osmanlı Mirası Araştırmaları Dergisi / Journal of Ottoman Legacy Studies
Cilt 11, Sayı 30, Temmuz 2024 / Volume 11, Issue 30, July 2024

277

The study focuses on mathematical content and includes current mathematical problem formulation translations. The mathematical problems were encoded using two different encoding types: LaTeX (Figures 13 and 14) and TEI-MathML (Figures 15 and 16), with examples for each encoding style given separately. The LaTeX code was converted to XHTML and seamlessly integrated with the MathJax JavaScript library to generate the web view.

```
<p n="2" xml:lang="ModernT">
  <seg n="7-14"> [7-14] <add>İkinci cevap:</add> Orantı ile çözümlendiği takdirce de cetvel yaklaşık 63 çap geçmişti. Bu iki işareti
  <!-- İkinci yöntem, orantıya göre, cetvel 63 çap geçmişse bunu 30zira ve 63 derece ile oranlarız.Yine 30*63=1890 olur ve bunu 20'ye
  <formula notation="TeX">
    <![CDATA[
    \begin{formula}
    $$\frac{\text{30 zira}}{\text{20 derece}}=\frac{\text{uzaklık}}{\text{sinüsün ikinci konulan yere kadar olan uzaklığı}}$$
    $$\text{uzaklık} = \frac{\text{sinüsün ikinci konulan yere kadar olan uzaklığı} \cdot (2 + 0.5)}{2} + \frac{\text{63 çap}}{2}$$
    $$\text{uzaklık} = \frac{\text{sinüsün ikinci konulan yere kadar olan uzaklığı} \cdot 2.5}{2} + \text{31.5 çap}$$
        \end{formula}]]>
  </formula>
</p>
```

Figure 13. Representing mathematical operations with LaTeX (Oxygen)

```
<p xml:lang="ModernT" lang="ModernT"><span>[23-26] <span class="add">Cevap:</span> Bir tarafından sinüsü bir
    bir işaret koyduk. Sonra</span>
    <div>
        <span class="mathjax">\begin{equation}\text{Uzaklık (zira)} = \dfrac{20^\circ\times 30}{60^\circ} =
            10 \text{ zira} \end{equation}</span>
    </div>
    <div>
        <span class="mathjax">\begin{equation}\text{Uzaklık} = \text{90 zira}
            = 9 \times \text{10 hesap} = \text{9 hesaplamalık} \end{equation}</span>
    </div>
```

Figure 14. Adaptation of LaTeX-JavaScript using MathJax (VSC)

The symbol "×" in Figure 15 represents mathematical multiplication. Using symbols such as plus (+) and minus (-) is also possible. Figure 16 shows the implementation of a MathML character called a zero-width space, denoted by "⁢". This distinguishes between the variables "x" and "y". This method improves the clarity and readability of the expression "xy". It is essential to understand that when mathematical code generated using MathML in XML is translated into XHTML for web representation, it does not require any changes or additional support from MathJax. In my opinion, it is more flexible than the LaTeX formulation.

```
<formula xmlns:ns="http://www.w3.org/1998/Math/MathML">
    <ns:math>
        <ns:mrow>
            <ns:mfrac linethickness="0">
                <ns:mrow>
                    <ns:mrow>
                        <ns:mn>30</ns:mn>
                        <ns:mo>zira</ns:mo>
                    </ns:mrow>
                    <ns:mo>&#xD7;</ns:mo>
                    <ns:mrow>
                        <ns:mn>63</ns:mn>
                        <ns:mo>derece</ns:mo>
                    </ns:mrow>
```

Figure15. Encoding maths operations with MathML. (Oxygen)

```
<ns:mrow>
    <ns:mtext>sinüsün ikinci konulan yere kadar olan uzaklığı</ns:mtext>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mo>&#8290;</ns:mo>
    <ns:mn>2</ns:mn>
    <ns:mo>+</ns:mo>
    <ns:mn>0.5</ns:mn>
</ns:mrow>
```

Figure 16. Encoding maths operations with MathML., HTML (VSC)

IV.2.2.3. Coding the Dictionary Section

This study has included a glossary section to enhance comprehension of the text using a conventional method. Having index and glossary sections is crucial for any textual analysis. These sections offer great time-saving benefits and make it more convenient to use the text for future research projects. Thus, it is crucial to establish standardisation for these sections in digital text editions.

The TEI provides a range of techniques for developing dictionaries. Even in the field of lexicography, various codes can be utilised. In this case, only fundamental coding has been utilised to ensure clear understanding. Figure 18 demonstrates the implementation of a newly created section using the <div> tag, as previously illustrated. This section is titled "Glossary". Later on, the <list> tag became the preferred choice. This approach aims to organise the glossary by listing the terms alphabetically. Afterwards, the "type" attribute is added with the value "terms" to specify the particular type of list that will be created from this tag. This indicates that the list will be composed of "term" elements. The <item> tag is the container for identifying items within a list or array. Open a subtag of this tag to specify language codes using the 'xml' property. The first words of each element are styled in bold using the <hi> element with the "rend" property set to "bold". Furthermore, <add> tags can be utilised to specify the gender or plurality of the expression. In this context, the <add> element offers additional information that improves the description, similar to how it is used in text restoration. Following this, the expressions are explained using the "description" element. Specifying the language code within the <desc> element by using the "xml:lang" attribute is necessary. In addition, using a comprehensive "id." simplifies the process of centralising language codes, eliminating the requirement to enter a language code for each annotation separately. By adopting a non-traditional approach to entering language codes for Turkish terms, the inclusion of codes for both terms and meanings in the coding has expanded the potential usefulness of the text for future digital studies. Implementing standardised practices will enhance the versatility of the text across various fields.

```
<div type="chapter" n="3">
    <head n="3" rend="bold" style="color:black;">III. Bölüm - Terim Sözlüğü</head>

    <list type="terms">
     <item>
        <term xml:lang="Ar"><hi rend="bold">aʿdād </hi>(Ar. <add>çoğ.</add>)<hi
            rend="bold">:</hi>
        </term>
        <desc xml:lang="T"> sayılar</desc>
     </item>
  <item>
```

Figure 17. Encoding the Glossary section (Oxygen)

Section 2.2.4 provides a more detailed description of the coding of facsimiles. As mentioned in Section 3.2.2.2.1, when coding facsimiles at TEI, it is important to determine their position (at the beginning, in the middle, or interspersed within the text) and their display format. In Figure 18, the facsimile was meant to be positioned right below the first page of the folio, which is why the "page break" tag was used. Interestingly, this tag can be used independently without requiring another "page break" tag to close it. The facs attribute "facsimile" specifies the file description where our image was stored. The "n" property specifies the folio page corresponding to "9a". Section 2.2.1 introduces a div element, followed by a figure element, as mentioned earlier. The graphic child element contains the image with the "url" attribute. The div tag includes additional details indicated by the "type" attribute, which is set to "facsimile," indicating that the embedded image is a facsimile. The value "n" determines the order of the image. The facsimile enclosed in the <div> tags will appear as a separate fragment.

```
<pb facs="MS1_1v.jpg" n="9a"/>
```

Figure 18. Facsimile encoding (Oxygen).

```
<div n="1v"> <!--1. varak-->
    <pb n="8v"/> <!--1.varak ön yüz-->
        <figure>
            <graphic url="MS1_1v-1r.JPG"/>
        </figure>
```

Figure 19. Facsimile encoding (Oxygen).

Converting images from XML to HTML simplifies the process of specifying CSS properties like size, alignment, and more, making it effortless to insert them. Here are three coding examples in HTML/CSS showcasing this process, with the facsimile section at the end. In the first two examples (Figures 20 and 21), the facs section is added using the <img src> element commonly found in HTML. In the other example, a different tag is used. Figure 22 displays the facsimile that will appear after each page.

```
<div>
    <section class="chapter" id="index.xml-body.1_div.4"></section>
    <header> <h1><span class="headingNumber"> 4. </span><span style="color:□black;;font-weight:bold">Tıpkıbasım</span></h1></header>
    <img src="MS1_1v-1r.JPG">
    <img src="MS1_2v-2r.JPG">
</section>
</div>
```

Figure 20. Display the facsimile in HTML at the end tab. (VSC)

```
<section class="chapter" id="index.xml-body.1_div.4"></section>
    <header> <h1><span class="headingNumber"> 4. </span><span style="color:□black;;font-weight:bold">IV. Bölüm - Tıpkıbasım</span></h1></header>
    <figure class="figure"><img src="MS1_1v-1r.JPG" alt="" class="graphic"/></figure>
    <figure class="figure"><img src="MS1_2v-2r.JPG" alt="" class="graphic"/></figure>
</section>
</div>
```

Figure 21. Display the facsimile in HTML at the end tab. (VSC)

```
<div class="teidiv1" id="index.xml-body.1_div.1_div.3">
    <figure class="figure"><img src="MS1_2v-2r.JPG" alt="" class="graphic"/></figure>
    <div class="pagebreak" id="index.xml-pb-d31e265">[Page 2r]</div>
```

Figure 22. Display the facsimile in HTML on the start tab. (VSC)

Conclusion

The specific results of this study can be categorised into two main groups. The first group pertains to the most important criteria enabling interdisciplinary cooperation in the development of digital text editions. These issues are discussed in detail in the sections of the article that address the preparation of the coding procedure for the text, transcription, and text editing. We anticipate widespread acceptance and adoption of these criteria, as well as the recognition of their importance in utilising TEI applications for historical text research in Turcology.

The second concerns the visual representation of the coding technique, which is described in detail in the paper. This might, of course, be displayed in a more visually appealing manner using TEI Publisher or other software. However, in this article, the HTML picture is utilized and a connection is provided to the readers through their GitHub site. The website is accessible at the URL: https://serminography.github.io/. The XML files and JPEG photographs can be accessible at the URL: https://github.com/serminography/serminography.github.io, with a GitHub account.

To summarize, the significant influence of TEI in the field of digital text studies indicates that the use of digital transformation is now necessary in Turcology, particularly in the analysis of historical texts. The TEI's ability to enable a wide range of research activities, including as lexicography, terminology, morpho-syntactic and morpho-semantic investigations, and active indexing procedures, showcases its potential to go beyond conventional methods in this sector. Consequently, this will result in the development of further open resources that are readily available, secure, and enable data-driven interpretations. In addition, the digitisation project seeks to include not only the period of the Ottoman Empire but also all written works that demonstrate the historical evolution of the Turkish language. Therefore, conducting additional research on the subject of Turcology with TEI will result in substantial progress in the analysis of historical texts.

Suggestions

To successfully adjust to this transition and avoid falling behind, the most straightforward method is to actively promote the acquisition of knowledge and expertise in software and AI technologies among graduate students. Turcology can leverage the younger generation's growing flexibility and inclination towards digital instruments. This approach will allow young researchers from different fields studying historical texts in Turkish to improve their skills by actively using TEI and discover novel methods of interpreting historical texts through this standardised framework. This will enhance our capacity to articulate novel insights regarding the revival of a historical text through digital transformation. Consequently, there will be a reciprocal exchange of information and engagement between the researcher and the field. This would enable a thorough digital modernisation of historical textual studies, a well-established area of research in Turcology, and provide chances for the progress of digital research.

Bibliography

Açıl, Berat, Yazar, Sadık, Kadir Turgut, Özgür Kavak. *Tenkitli Neşir Kılavuzu (Osmanlı Türkçesi Metinleri İçin)*, İstanbul: İSAM, 2022.

Aladağ, Fatma, "Dijital beşerî bilimler ve Türkiye araştırmaları: Bir literatür değerlendirmesi", *Türkiye Araştırmaları Literatür Dergisi*, 18, no. 36 (2020): 773–796, accessed: June 2021, https://dergipark.org.tr/tr/pub/talid/issue/58073.

ASCII, American Standard Code, https://www.ascii-code.com/, accessed: 3 March 2022.

Baxter, Pamela ve Susan Jack. "Qualitative case study methodology: Study design and implementation for novice researchers," *The Qualitative Report*, 13, (2008), 547,

accessed: 27 July 2022, https://nsuworks.nova.edu/tqr/vol13/iss4/2/, https://doi.org/10.46743/2160-3715/2008.1573.

Blackman, Tim. What is Crowd-Based Translation: A beginner's Guide, April 25, 2018, https://www.brombergtranslations.com/crowdsourcing-translation-crowd-based-translation/ accessed: 9 March 2022.

Burnard, Lou. What is the Text Encoding Initiative? How to add intelligent markup to digital resources (Marseille: OpenEdition Press, 2014), 12, accessed: 7 March 2022, https://books.openedition.org/oep/426, https://doi.org/10.4000/books.oep.426.

"Concept of digital heritage", UNESCO 2018, accessed: 5 March 2022, https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-heritage.

Cummings, James. "The Text Encoding Initiative and the study of literature" in Ray Siemens & Susan Schreibman (Eds.), A Companion to digital literary studies Oxford: Blackwell Publishing, 2007, accessed: 7 March 2022, https://doi.org/10.1002/9781405177504.ch25.

Çakmak, Tolga. "Kültürel bellek kurumlarında dijitalleştirme: Ankara'daki 17 kurumun uygulamalarının analizi", *Ankara Araştırmaları Dergisi 6*, no. 2 (2018): 167–182. https://doi.org/10.5505/jas.2018.30502.

Dölek, İshak ve Atakan Kurt. "Derin Sinir ağlarıyla Osmanlıca Optik Karakter tanıma", *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi* 38, 4 (April 2023): 2579-94, https://doi.org/10.17341/gazimmfd.1062596,https://dergipark.org.tr/en/pub/gazimmfd/issue/75069.

Drucker, Johanna. "Humanistic theory and digital scholarship," *Debates İn The Digital Humanities*, *Part II. Theorising The Digital Humanities*, (2012), accessed: 7 December 2022, http://dhdebates.gc.cuny.edu/debates/text/34.

Goldfarb, Charles, F. *The SGML handbook*, Oxford: Clarendon, 1990.

Hockey, Susan."The history of humanities computing," *in A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, and John Unsworth (Oxford: Blackwell Publishing, 2004), accessed 11 April 2022, https://doi.org/10.1002/9780470999875.ch1

ISO 8879. "SGML Standard 1986," ISO, accessed: 3 March 2022, https://www.iso.org/standard/16387.html.

Kalafat, Şermin. "Türkoloji araştırmaları kapsamında yazma eser çalışmalarına yönelik bir inceleme", The Second International Congress of Turcology, "600 Years of Polish-Turkish Relations: Research on Turcology Yesterday, Today, and Tomorrow," Warsaw, Poland, June 25, 2014.

Lee, Stuart D., *Digital Imaging: A Practical Handbook*, London: Facet Publishing, 2001.

Manoff, Marlene. "The Materiality of Digital Collections: Theoretical and Historical Perspectives", *Libraries and the Academy*, 6, no. 3 (2006), 311, accessed: January 2022, http://hdl.handle.net/1721.1/35689.

Pfannenschmidt, Sarah L. and Tanya E. Clement. "Evaluating digital scholarship: suggestions and strategies for the text encoding initiative," Journal of the Text Encoding Initiative, 7 (November 2014), 7; accessed: 21 March 2022, https://journals.openedition.org/jtei/949, https://doi.org/10.4000/jtei.949.

Pittock, Murray. "The Scottish Heritage Partnership Immersive Experiences: Policy Report", (University of Glasgow, 2018: 5),

https://www.gla.ac.uk/schools/humanities/research/informationstudiesresearch/researchpr
ojects/scottishnationalheritage/#/policyreport:immersiveexperiences.

Polat, Atilla ve Halime Mücella Demirhan Çavuşoğlu. "Mehmed Said Efendi'nin Misâha
Risâlesi". Osmanlı Bilimi Araştırmaları 21, 2 (July 2020): 249-70, accessed: 23 March
2021,   http://dx.doi.org/10.26650/oba.650926,
https://iupress.istanbul.edu.tr/en/journal/oba/home.

Purdy, James P., ve Joyce R. Walker. "Valuing digital scholarship: Exploring the changing
realities of digital work", *Profession, Modern Language Association*, (2010), 177–195,
accessed: 11 April 2022, https://www.jstor.org/stable/41419875.

Renear, A., "Text Encoding," in A Companion to Digital Humanities, ed. Susan Schreibman,
Ray Siemens, and John Unsworth (Oxford: Blackwell Publishing, 2004) accessed: 11
April   2022,
https://companions.digitalhumanities.org/DH/?chapter=content/9781405103213_chapter_
17.html

Sahle, Patrick. "What is a Scholarly Digital Edition?" *In Digital Scholarly Editing: Theories
and   Practices*,   (Cambridge,   2017),   19-25,   accessed: March   2022,
https://books.openedition.org/obp/3397.

Soualah, Mohammed Ourabah and Mohamed Hassoun. "A TEI P5 Manuscript Description
Adaptation For Cataloguing Digitised Arabic Manuscripts", *Journal of the Text Encoding
Initiative*,   2,   no.   2   (February   2012),   2,   accessed,   11   April   2022,
https://doi.org/10.4000/jtei.398, https://journals.openedition.org/jtei/398

Sperberg-McQueen, C. Michael. "Textual criticism and the Text Encoding Initiative." Paper
presented at the annual convention of the Modern Language Association, (San Diego,
California: December   1994),   accessed: 11   April   2022,   https://tei-
c.org/Vault/XX/mla94.html

Svensson, Patrik. "Humanities computing as digital humanities," (Digital Humanities Quarterly,
2009),   accessed: 11   April   2022,
http://www.digitalhumanities.org/dhq/vol/3/3/000065/000065.html.

"Text Encoding Initiative." Text Encoding Initiative: TEI, accessed: 11 June 2021, https://tei-
c.org/.

Van Hyning, Victoria Anne, ve Mason A. Jones. "Data's Destinations: Three Case Studies in
Crowdsourced Transcription Data Management and Dissemination", *Startwords*, 2 (01
October   2021)   accessed: 9   March   2022,   Crowsourching,
https://www.qub.ac.uk/researchcentres/CentreforDataDigitisationandAnalysis/Servicesan
dOutputs/ToolkitsandGuides/Crowdsourcing/.

Vanhoutte, Edward. "An Introduction to the TEI and the TEI Consortium", Literary and
Linguistic Computing,19, no. 1 (2004), 10, accessed: 11 April 2022,
https://www.edwardvanhoutte.org/pub/2004/teillc.htm.

Vertan, Cristina and Stefanie Reimers. "A TEI-based application for editing manuscript
descriptions," Journal of the Text Encoding Initiative 2 (2012), accessed: 11 April 2022,
https://doi.org/10.4000/jtei.392, https://journals.openedition.org/jtei/392.