

Osmanlı Çalışmalarında Dijital Edisyon ve Metin Kodlama: *Text Encoding Initiative-TEI*

FATMA ALADAĞ

Universität Leipzig, Doktora Öğrencisi.

(fatmaaladag1@gmail.com), ORCID: 0000-0003-3455-6181.

“ ” Aladağ, Fatma. “Osmanlı Çalışmalarında Dijital Edisyon ve Metin Kodlama: *Text Encoding Initiative-TEI*.” *Zemin*, s. 6 (2023): 190-211.

DOI: <https://doi.org/10.5281/zenodo.10435921>.

Metin Kodlama Girişimi (*Text Encoding Initiative-TEI*),¹ metin odağındaki dijital çalışmalar için standartlar oluşturan, özellikle tarihî belgelerin bilgisayar tarafından işlenmesi ve analiz edilmesini kolaylaştıran bir kodlama sistemi olarak ortaya çıkmıştır. Tarih, edebiyat, dilbilim gibi disiplinlerin araştırma kaynağı olan belgeler, arşivler ve yazma kitapların düzenlenmesi, kodlanması ve dijital olarak yayınlanması için TEI uluslararası bir standart sunan önemli bir araçtır. Osmanlı çalışmalarında, TEI'nin kullanımı, imparatorluğun *büyük verisini* oluşturan devasa arşiv külliyatının dijital ortamda yeniden oluşturulması ve analiz edilmesi için gündeme alınması gereken bir yaklaşımdır. Böyle bir teknoloji, Osmanlı Türkçesi metinlerin hem daha erişilebilir olmasına hem de araştırmacıların metinleri daha derinlemesine analiz etmesine olanak tanımaktadır.

TEI'nin sağladığı standartlaştırma, farklı kaynaklardan gelen metinlerin birbiriyle uyumlu hâle getirilmesine ve böylece daha geniş bir veri tabanının oluşturulmasına da imkân tanır. Nitekim TEI metinlerin yapısını ve içeriğini tanımlayarak, farklı araştırmacılar ve projeler arasında uyumluluğu desteklemektedir. Böyle bir yaklaşım, sayısı her geçen gün artan dijital beşerî bilimler kaynaklarını ve ilmî çıktılarını birbiriyle konuşabilir ve ilişkilendirilebilir hâle getirmektedir.²

Bu yazıda, dijital beşerî bilimlerde önemli bir alan olarak metin kodlama, *Text Encoding Initiative* (TEI) yönteminin Osmanlı çalışmalarındaki önemi uygulamalı örneklerle incelenmektedir. Osmanlı İmparatorluğu'nun geniş coğrafyası, yüzyıllar süren varlığı, çok dilli ve çok kültürlü yapısı, tarihî belgelerin ve metinlerin de tema, format ve hacim açısından zenginliğini sağlamıştır. TEI'nin sahip olduğu kapsamlı kodlama mimarisi, bu kaynakların dijital ortamda yayımlanması ve analiz edilmesi için bir yöntem sunar.

Çalışmada öncelikle TEI'nin kısaca tarihsel gelişimi ve kapsamından bahsedilecek ve ardından Osmanlı çalışmalarındaki potansiyeli örnek metin olarak seçilen *Küçük Mecmua* dergisinin ilk sayısı üzerinden uygulamalı olarak anlatılacaktır.

1 TEI, <https://tei-c.org/> (erişim 11 Kasım 2023). Bu çalışmada “TEI” şeklindeki kısaltma kullanılacaktır.

2 Fatma Aladağ, “Dijital Beşerî Bilimler ve Türkiye Araştırmaları: Bir Literatür Değerlendirmesi,” *Türkiye Araştırmaları Literatür Dergisi*, s. 18 (2020): 773-796, <https://dergipark.org.tr/tr/pub/talid/issue/58073/800406>.

Ayrıca TEI ile dijitalleştirilen metinler için Voyant Tools aracı kullanılarak metin analizi örneği sunulacaktır. Böylece TEI'nin Osmanlı tarihi üzerine yapılan araştırmalara nasıl katkıda bulunabileceği, teorik ve uygulamalı yönleriyle ele alınmış olacak ve TEI'nin sunduğu imkânlar somut örnekler üzerinden gösterilecektir. Bu yönüyle makale, TEI'ı Osmanlı çalışmalarının gündemine dâhil etmek için bir farkındalık oluşturmayı hedeflemektedir.

TEI'nin Gelişimi ve Kapsamı

1980'lerin başında, dijital teknolojilerin hızla gelişimine paralel olarak, metinlerin dijital ortamda ve standart bir formatta kodlanma ihtiyacı ortaya çıkmıştır.³ Metinlerle çalışan disiplinlerin özellikle de *Thesaurus Linguae Graecae*, *Trésor de la Langue Française* ve *Dictionary of Old English* gibi projeler tarafından ortaya çıkan kaynak metinlerin hesaplamalı olarak erişilebilir ve izlenebilir olmasına giderek ilgi artmıştır. Hızla değişen depolama teknikleri ve bunun sonucunda açık erişim ve bağımsız formatlara odaklanarak dijital kaynaklara uzun vadeli erişim sağlamaya yönelik tartışmalar akademik çevrelerde çoğalmıştır.⁴ Bu süreçte dilbilim, edebiyat, tarih, kütüphane ve mühendislik gibi farklı alanlardan uzman ve araştırmacılar, metinlerin dijitalleştirilmesi ve analizi için ortak bir dil ve yöntem geliştirmek adına yıllarca bir araya gelerek birlikte çalışmıştır. 1987 yılında Oxford, Michigan ve Virginia Üniversitelerinin iş birliğiyle resmi bir yapıya dönüşen TEI, bir inisiyatif olarak dijital metinler için kodlama standardı geliştirmek ve bu sayede dijital beşerî bilimlerin alanındaki çalışmalar arasında da iş birliği ve uyumluluğu artırmayı hedeflemiştir.

Metinleri kodlamak, bu metinlerin dijital ortamda işlenmesi, saklanması ve analiz edilmesi için zorunlu bir süreçtir. Microsoft Word veya Adobe PDF gibi programların sunduğu metin işleme ve görsel sunum özelliklerinin ötesine geçerek kodlama, araştırmacıların metinleri daha kolay yönetmelerine, bilgisayarların desteğini alarak hesaplamalı ve kompleks analizler yapmalarına ve görselleştirmelerine olanak tanır. Böylece metinler sistematik bir şekilde

3 Nancy M. Ide and C. M. Sperberg-McQueen, "The TEI: History, Goals, and Future," *Text Encoding Initiative Background and Context*, ed. Nancy Ide and Jean Véronis (Springer, 1995), 5-15.

4 Lou Burnard, "The Evolution of the Text Encoding Initiative: From Research Project to Research Infrastructure," *Journal of Text Encoding Initiative*, no. 5 (2013): 1-12, <https://journals.openedition.org/jtei/811> (erişim 25 Kasım 2023).

incelenabilir. Örneğin, tarihî belgeler veya edebî eserlerdeki farklı metin türleri veya versiyonlar, kodlama sayesinde belirli temalar, kavramlar, dil veya yapısal özellikler açısından karşılaştırılabilir ve analiz edilebilir. Ayrıca metinler içerisindeki kişi, kurum, yapı, olay ve daha birçok temadaki bilgiler büyük metin külliyatları içerisinde zaman, mekân veya ilişki ağı perspektifinden keşfedilebilir. Ayrıca, kodlama, metinlerin dijital olarak arşivlenmesi, erişilmesi ve korunmasında da önemli bir rol oynamaktadır.

TEI topluluğu ilk yıllarında temel kodlama kılavuzları ve öneriler sunarak başlamış, zaman içinde bu kılavuzları genişletmiş ve güncellemiştir.⁵ Bu süreç aynı zamanda akademik çevrelerin geri bildirimleri ve teknolojik gelişmeler doğrultusunda da şekillenmiştir. TEI'nin ortaya çıkışının temel nedeni, mühendislerin ve bilgisayar programcılarının kullanımına yönelik geleneksel kodlama dillerinin, özellikle beşerî bilimlerdeki metinlerin semantik ve karmaşık yapısını yeterince temsil edememesi olarak gösterilmektedir. Nitekim, beşerî bilimcilerin tarihsel ve edebî metinlere özgü ortaya çıkan kod dili ihtiyacı bu alandaki araştırmacıların bir araya gelerek TEI gibi bir kod sisteminin oluşmasını sağlamıştır. Özellikle beşerî bilimlerin araştırma metinleri farklı versiyonlar ve dillerin oluşturduğu çeşitliliğin yanı sıra görsel öğeler, referanslar ve derkenar gibi detayların da olduğu bölümleri içermektedir. Dolayısıyla geleneksel mühendislik alanındaki kodlama dilleri, bu tür zengin metin yapılarını yeterli bir kapsamda kodlanması ve dijital olarak ifade edilmesinde tercih edilmemiştir.

TEI, genel bir işaretleme dili olan SGML üzerine kurulmuş ve zamanla metinlerin yapısal ve semantik özelliklerini tanımlamak için XML (*eXtensible Markup Language*) tabanlı bir sisteme geçilerek bir dizi kılavuz geliştirilmiştir. Böylece internet ve web teknolojileri ile uyumlu olan XML formatı sayesinde kodlanmış metinlerin çevrimiçi paylaşımı ve işlenmesi için kolaylık oluşmuştur. Ayrıca XML, üst (*meta*) veri yani *metinler hakkında bilgi* kodlamasını da desteklemekte ve metinlerin yazarı, yayın bilgileri, arşiv konumu gibi ek bilgilerin kaydedilmesine de olanak tanımaktadır. Daha da önemlisi XML, araştırmacıların kendi kod etiketlerini tanımlamasına izin verir. Bu da belirli bir araştırma veya proje için özel gereksinimleri karşılayacak biçimde metinlerin kodlanmasını sağlar ki bu şekilde özelleştirilebilir bir etiketleme yapısı ve esnek dil işleme kapasitesi

⁵ C.M. Sperberg-McQueen and Lou Burnard, *Guidelines for Electronic Text Encoding and Interchange* (Chicago and Oxford, 1994).

Osmanlı Türkçesi gibi özel ve karmaşık dillerin gereksinimlerini karşılama oldukça önemlidir.

1990'lar ve 2000'ler boyunca TEI, gelişen teknolojiyle uyumlu olarak ve farklı dillerin de bu uygulamayı kullanmasıyla birlikte geliştirilmiş ve güncellenmiştir. Bu süreçte, TEI kodlama kılavuzları sürekli olarak güncellenmiş ve beşerî bilimcilerin ihtiyaçlarına göre genişletilmiştir. Böylece farklı dillerdeki ve çeşitli biçimlerdeki metinleri kapsayacak şekilde geniş bir standart hâline gelen TEI artık metin analizi, arşivleme ve dijital edisyon gibi çeşitli alanlarda kullanılan evrensel bir standarttır. Bu yönüyle TEI ortaya çıkan ihtiyaçlara ve teknolojiye göre sürekli geliştirilen ve güncellenen yapısı sayesinde metinlerin dijitalleştirilmesinde yeni yaklaşımların ve metodolojilerin ortaya çıkmasına da öncülük etmektedir.

TEI Yönergeleri

TEI kodlama yönergeleri, uluslararası akademisyenler, kütüphaneciler, dilbilimciler ve diğer uzmanlar tarafından geliştirilmiştir. Bu yönergeler esnek bir yapıya sahiptir ve farklı diller, metin türleri ve projeler için genişletilebilir veya özelleştirilebilir bir yapıdadır. Yönergeler belirli temalarda nasıl kodlama yapılacağı ve hangi kod etiketlerinin kullanılacağına dair araştırmacılara bilgi sağlamaktadır. Bu temalar a) metin yapısı, b) metin içeriği, c) üst veri, d) özel karakter ve semboller, e) görseller ve multimedya, f) notlar, açıklamalar ve referanslar, g) dilbilimsel özellikler gibi belirli alanlarda kategorilere ayrılmaktadır. Metin yapısı; başlıklar, paragraflar, listeler, tablolar gibi metindeki yapısal öğeleri tanımlamak için etiketler sunar. Metin içeriği; anlatı, diyalog, şiir gibi farklı metin türleri için etiketleme seçeneklerini kapsar. Üst veri ise yazar, yayın tarihi, haklar, yayıncı gibi bilgileri etiketleme içeriğine odaklanır.

TEI yönergelerinin kullanımında, metinlerin XML tabanlı kodlanması ve standardizasyonu için DTD (Document Type Definition) dosyası hazırlamak esastır. DTD, XML belgelerinin kod yapısını tanımlayan ve bu yapının belirli kurallara uygun olmasını sağlayan bir dildir. Bir diğer ifadeyle, DTD bir araştırma kaynağı için hazırlanan XML'de hangi etiketlerin bulunacağını, amaçları ve bu etiketlerin nasıl bir hiyerarşi içinde olması gerektiğini tanımlayan dosyadır. DTD ile standartlaştırma, veri bütünlüğü ve kod dosyasının geliştirme süreçleri sağlanmış olur ve bu da farklı sistemler ve uygulamalar arasında da veri alışverişini kolaylaştırır.

Bu yönergeler yazma eserleri, edebî ve tarihî belgeleri orijinallerine en yakın şekilde kodlayarak dijital temsilini sağlamaktadır. Örneğin bir yazmada yer alan başlıkları, kenar notlarını, paragrafları ve hatta sayfa düzenini (iki sütunlu sayfa yapısı gibi) kodlamak için etiketler sunar. <head> etiketi başlıklar için, <note> kenar notları için ve <div> metni bölümlere ayırmak için kullanılabilir. Dil <lang>, terimler <term>, özel karakter ve semboller <char> ve <glyph>, metne sonradan eklenen veya çıkarılan kısımlar için <add> ve , metnin yazarı, yazıldığı tarih, kaynak bilgileri gibi üst veri için <fileDesc> ve bu bölüm içinde yer alan <titleStmnt>, <publicationStmnt> ve <sourceDesc> etiketleri kullanılmaktadır. Bunların yanı sıra başlık ve altbaşlıklar <title>, <subtitle>, paragraflar <p>, alıntılar ve içerikler <quote>, <cit>, liste ve öğeleri <list>, <item>, referans ve linkler <ref>, <link>, biçimlendirme ve stiller <hi>, <rend> gibi çeşitli yapısal etiketler bulunmaktadır.

Metnin içeriğiyle ilgili semantik (anlamsal) etiketler de bulunmaktadır. TEI daha önce de vurgulandığı üzere metin içindeki yapılar, kişiler, olaylar, yerler, tarihler ve lokasyonlar için özel etiketlerle metni tanımlamaya olanak tanır. Kişilerin isimleri, metin içinde <persName> etiketi kullanılabilir. Örneğin, bir Osmanlı tarihi kaynağında, Sultan II. Mehmed'in adı <persName>II. Mehmed</persName> şeklinde kodlanabilir.

Önemli tarihî veya edebî olaylar <event> etiketi ile tanımlanabilir. Örneğin, İstanbul'un fethi, <event type="fetih">İstanbul'un Fethi</event> şeklinde kodlanabilir.

Coğrafi yerler ve lokasyonlar <placeName> ve <location> etiketleri ile temsil edilmektedir. Örneğin, Ayasofya, <placeName>Ayasofya</placeName> etiketiyle kodlanabilir.

Tarihler, <date> etiketi kullanılarak kodlanabilir ve böylece hem belirli tarihler hem de dönemler metinde ayrıştırılabilir. Örneğin, İstanbul'un fethinin tarihi, <date>1453</date> olarak kodlanabilir. Ayrıca Hicri ve Miladi olarak da tarihler <date when="1453" when-hijri="857"/> şeklinde dosyada versiyonlanabilir.

TEI, ayrıca kişiler, yerler ve olaylar arasındaki ilişkiler için de kod etiketleri sunmaktadır. Böylece metinlerdeki karmaşık ilişkisel yapıları göstermek mümkün olabilir. Örneğin, bir kişinin belirli bir yerle veya olayla ilişkisi <relation> etiketi ile kodlanabilir.

Bu etiketleme sistemi, tarihî ve edebî metinlerdeki karmaşık yapıların ve ilişkilerin dijital ortamda doğru ve orijinal yapısına uygun bir şekilde temsil

edilmesini sağlamaktadır. Bu sayede, metinlerdeki kişilerin, olayların ve yerlerin bağlamları daha açık bir şekilde anlaşılabilir ve araştırmacılar için bilgisayar destekli tematik analiz imkânı doğar. Nitekim böyle hesaplamalı bir yaklaşımla büyük hacimli metinleri analiz etmek için bu metinleri kodlama diline sahip bir formata dönüştürülmesi gerekmektedir. Araştırmacıların bu tür detaylı kodlamayla, tarihî metinlerin derinlemesine incelenmesi, farklı versiyonlarında ya da metnin içerisindeki diyaloglar, kişi, yapı, organizasyon, olay veya yer odağında karşılaştırmalı bir okuma ve görselleştirme yapmasını sağlamaktadır. Ayrıca edebî bir eserin farklı basımlarının içeriği, sayfa yapısı, kaynağın kâğıt ve baskı türü veya metinsöylemi eş zamanlı karşılaştırılabilir, dizinleri hazırlanabilir. Ya da metinde geçen kişi, olay, yer veya kurumların ilişki ağı, haritası veya veri tabanları oluşturulabilir.

Bu yönergelerin uygulanmasıyla ilgili çeşitli rehberler, kılavuzlar ve çevrimiçi eğitim materyalleri üretilmiştir. Bu kaynaklar, TEI'nin kullanımıyla ilgili örnekler ve adım adım uygulamanın nasıl yapılacağına dair bilgiler içermektedir. *TEI Consortium*⁶ web sitesi resmi dokümantasyonlar sunmaktadır. Ayrıca *TEI by Example*⁷ bu yönergeleri örneklerle anlatan ve açıklayan bir web sitesidir.

TEI dokümanlarını görüntülemek ve yayınlamak için de çevrimiçi araçlar geliştirilmiştir. *TEI Viewer*⁸ ve *TEI Publisher*⁹ beşerî bilimcilerin XML dosyalarını ve içeriklerini tematik olarak yayınlama imkânı sunmaktadır. *Vincent van Gogh: The Letters*¹⁰ TEI ile tarihî metinlerin dijital edisyonlarının hazırlanması üzerine güzel bir örnektir. Proje kapsamında Van Gogh'un kardeşi Theo'ya, sanatçı arkadaşları Paul Gauguin ve Emile Bernard'a ve daha pek çok kişiye yazdığı mektupları TEI ile tematik kodlanarak dijital ortamda yayımlanmıştır. Projenin web sitesinde tarihî metinler “kişiler” (1782 adet) ve “mektuplar” (928 adet) olarak iki farklı temada görüntülenmektedir (*Görsel 1*).

6 TEI, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html> (erişim 11 Eylül 2023).

7 TEI by Example, <https://teibyexample.org/exist/> (erişim 13 Eylül 2023).

8 TEI Viewer, <https://alexandermendes.github.io/tei-viewer/> (erişim 13 Eylül 2023).

9 TEI Publisher, <https://teipublisher.com/> (erişim 13 Eylül 2023).

10 Vincent van Gogh: The Letters, <https://teipublisher.com/exist/apps/vangogh/index.html> (erişim 13 Eylül 2023).

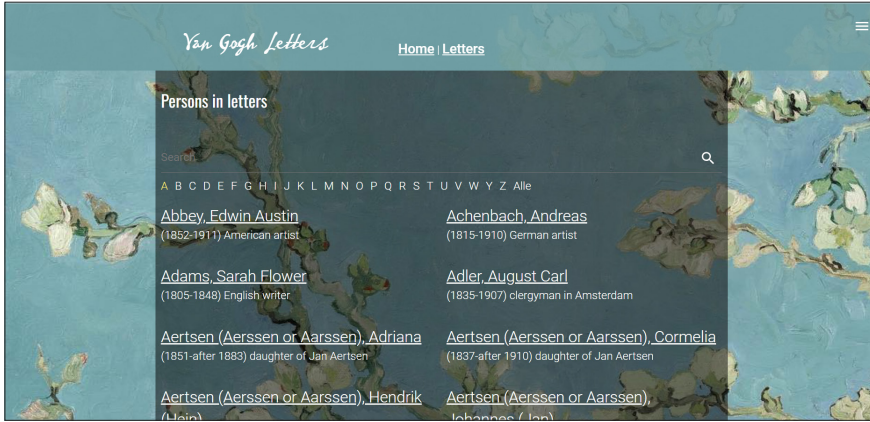


1: Vincent van Gogh: The Letters'in TEI Publisher ile hazırlanmış dijital edisyon ana sayfası.

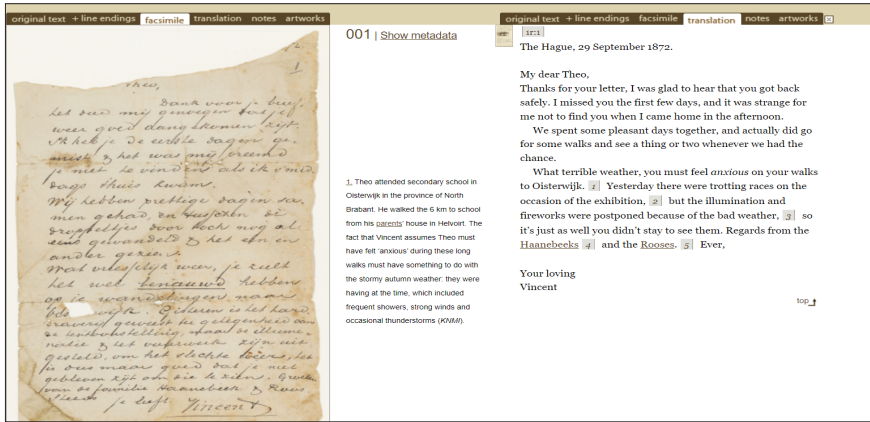
Dijital edisyon web sitesinde TEI ile kodlanmış metinlerin üst verisi, orijinal Hollanda dilindeki versiyonu, farklı dillere göre transkripsiyonları, editoryal açıklamalar, notlar ve görsellerle birlikte eş zamanlı okunabilir şekilde erişime açılmıştır (Görsel 2). Ayrıca metinlerde geçen kişilerin indeksi hazırlanmıştır ve bu sayede araştırmacılar bağlı veri (*linked data*) sistemiyle bu kişilerle ilgili bibliyografik harici veri tabanlarına ulaşabilmektedirler (Görsel 3). Projenin orijinal bir başka web sitesinde ise mektuplar TEI yanı sıra orijinal belge formatlarıyla eş zamanlı görüntülenebilir şekilde yayınlanmıştır (Görsel 4). Edisyonda satır aralıkları orijinal belgedeki satır sayısına sadık kalarak kodlanmıştır. Ayrıca cümle sonlarındaki referans numaralarına tıklanarak editoryal ek açıklamalar görüntülenebilmektedir.



2: Vincent van Gogh: The Letters'dan örnek bir dijital edisyon görünümü.



3: Vincent van Gogh: The Letters'daki kişi isimlerinin dizini.



4: Vincent van Gogh: The Letters projesinin alternatif web sitesinde dijital edisyon görünümü.

Mektupları tematik olarak ve farklı versiyonlarıyla birlikte keşfetmeye ve görselleştirmeye imkân sağlayan dijital edisyon mimarisinin arka planı TEI XML kod formatıdır (Görsel 5). Bu kod dosyasında başlık <title>, yayıncı <publisher>, editör <editor>, yayın yeri <pubPlace> gibi birçok üst verilerin yanı sıra kişi ismi <rs type="pers">, yer <vg:placeLet>, tarih <date> gibi farklı etiketler kullanılarak tarihî mektuplar tematik olarak dijitalleştirilmiştir (Görsel 5).

```

1 <!-- Last Changed Date: 2018-07-18 12:32:41 +0200 (Wed, 18 Jul 2018) | SVN Revision: 2206 --><!--Licensed under
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)--><?xml-model href="https://xslschema
.huygens.knaw.nl/vgodd.rng" type="application/xml" schematypens="http://relaxng.org/ns/structure/1.0?>
2 <?TEI xmlns="http://www.tei-c.org/ns/1.0" xmlns:vg="http://www.vangoghletters.org/ns/"
3 <teiHeader xmlns:xsl="http://www.w3.org/2001/XMLSchema-instance">
4 <fileDesc>
5 <titleStm>
6 <title>To Theo van Gogh. The Hague, Sunday, 29 September 1872.</title>
7 <editor>Leo Jansen</editor>
8 <editor>Hans Luijckx</editor>
9 <editor>Nienke Bakker</editor>
10 </titleStm>
11 <publicationStm>
12 <publisher>
13 <name>Huygens Instituut voor Nederlandse Geschiedenis (KNAW)</name>
14 <email>info@huygens.knaw.nl</email>
15 </publisher>
16 <publisher>Van Gogh Museum</publisher>
17 <pubPlace>Amsterdam</pubPlace>
18 <date type="first" when="2009" date="2009">2009</date>
19 <availability status="restricted">
20 <license target="http://creativecommons.org/licenses/by-nc-sa/4.0/ https://creativecommons.org/licenses/by-nc-sa/4.0
/legcode">
21 <p>Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) </p>
22 </license>
23 </availability>
24 <ptr target="http://vangoghletters.org/orig/let001/">
25 </publicationStm>
26 <sourceDesc>
27 <sourceDesc>
28 <vg:letIdentifier>

```

5: Vincent van Gogh: The Letters'in TEI XML kod mimarisinden bir görünüm

Projenin iki ayrı web sitesi çevrimiçi ve açık erişim olarak araştırmacıların hizmetine sunulmuştur. Bu yaklaşım sayesinde arşiv belgeleri hem dijital edisyon imkânlarına dâhil edilmiş hem de tematik olarak analize hazır hâle getirilmiştir.

Bu proje örneğinde görülen dijital edisyon mimarisi çok daha gelişmiş tematik görselleştirmelerle Osmanlı Türkçesi kaynakları yayınlamak ve içeriğini analiz etmek için önemli bir potansiyele sahiptir. Nitekim Osmanlı çalışmalarının zengin kaynaklarında yer alan kişi, yer, mekân, yapı, olay, tarih, harita, gravür ve daha birçok temadaki içerik bilgisi, kaynakların dönemler içerisinde mekânsal ve kurumsal dolaşımı ve etkileşimi TEI ile erişilebilir, ilişkisel ve semantik olarak analiz edilebilir ve görselleştirilebilir.

Osmanlı Çalışmaları ve TEI

Osmanlı belgeleri farklı dillerde (Osmanlı Türkçesi, Arapça, Farsça vb.) yazılmış ve süslemeler, karmaşık yazı stilleri ve sayfa düzenleri içermektedir. TEI, bu belgelerin metin yapısını ve görsel özelliklerini dijital ortamda orijinaline yakın bir şekilde temsil etmek için gerekli olan detaylı etiketleme sistemini sunma kapasitesine sahiptir.

Osmanlı Türkçesi matbu ve yazma kitaplar ile belgelerin dijitalleştirilmesi hususunda son yıllarda önemli çalışmalar yapılmaktadır.¹¹ Başlangıçta, belgelerin

¹¹ Osmanlı Türkçesi metinlerin dijitalleştirilmesi çalışmalarıyla ilgili kapsamlı bir literatür için bk. Fatma Aladağ ve Elif Derin Can, "Programlamadan Yapay Zekâya Osmanlı Türkçesi için Metin Tanıma ve Otomatik Transkripsiyon," *Dijital Beşerî Bilimler ve Osmanlı Çalışmaları*, haz.

basit bir şekilde taranıp sayısallaştırılmasıyla sınırlı olan süreç, artık içerik araması, otomatik transkripsiyon ve analiz için OCR (Optical Character Recognition) ve HTR (Handwritten Text Recognition) gibi teknolojilerle gelişmiştir. OCR, metin tanıma sürecinin öncüsü olarak, karakterlerin şekillerini analiz eden bir tanıma işlemidir ve sonrasında geliştirilen tekniklerle daha etkili hâle gelmiştir. HTR ise özellikle el yazısı metinlerin tanınması için geliştirilmiştir. Osmanlı Türkçesi metin tanıma ve otomatik transkripsiyon için çeşitli arama motorları, sözlükler, platformlar ve araçlar bulunmaktadır. *LexiQamus*,¹² *Wikilala*,¹³ *Müteferriqa*¹⁴ ve *osmanlica.com*¹⁵ gibi ticari girişimlerinin yanı sıra IRCICA,¹⁶ AKİS¹⁷ ve Digital Ottoman Corpora¹⁸ gibi akademik projeler yürütülmektedir. 2023 yılında Digital Ottoman Corpora ekibi *Transkribus*¹⁹ yapay zekâ platformunu kullanarak ilk kez matbu Osmanlı Türkçesi otomatik transkripsiyon modeli geliştirmiş, açık erişim ve ücretsiz olarak yayına açmıştır.²⁰ *Transkribus* platformu TEI standardında çıktı formatı sağlaması ve hem otomatik transkripsiyon hem de

Yunus Uğur (İstanbul: Vakıfbank, 2023), 45-77; Süphan Kırmızıaltın and David Wrisley, “Automated Transcription of Non-Latin Script Periodicals: A Case Study in the Ottoman Turkish Print Archive,” *Digital Humanities Quarterly* 16, no. 2 (2022), <https://www.digitalhumanities.org/dhq/vol/16/2/000577/000577.html> (erişim 04 Kasım 2023); Esmâ F. Bilgin Taşdemir, “Printed Ottoman Text Recognition Using Synthetic Data and Data Augmentation,” *International Journal on Document Analysis and Recognition*, no. 26, (2023): 273-287; Y. Sait Can, M. Erdem Kabadayı, “Line Segmentation of Individual Demographic Data from Arabic Handwritten Population Registers of Ottoman Empire,” *Document Analysis and Recognition – ICDAR 2021 Workshops*, edited by Elisa H. Barney Smith and Umapada Pal (Cham: Springer International Publishing, 2021), 312-21. https://doi.org/10.1007/978-3-030-86198-8_22; Dijital Osmanlı çalışmalarında yürütülen proje ve yayımlar için bkz. Digital Ottoman Studies Platformu, <https://www.digitalottomanstudies.com/> (erişim 14 Kasım 2023).

¹² LexiQamus, <https://www.lexiqamus.com/tr> (erişim 04 Kasım 2023).

¹³ Wikilala, <https://www.wikilala.com/> (erişim 04 Kasım 2023).

¹⁴ Muteferriqa, <https://mutferrriqa.com/> (erişim 04 Kasım 2023).

¹⁵ osmanlica.com, <https://www.osmanlica.com/> (erişim 04 Kasım 2023).

¹⁶ IRCICA, <https://library.ircica.org/Projects/OttomanOCR> (erişim 04 Kasım 2023).

¹⁷ AKİS, <https://dhlab.sabanciuniv.edu/tr/akis-0> (erişim 07 Kasım 2023).

¹⁸ Digital Ottoman Corpora, Süphan Kırmızıaltın, Elif Derin Can ve Fatma Aladağ, <https://www.digitalottomancorpora.org/> (erişim 04 Kasım 2023).

¹⁹ Transkribus, <https://readcoop.eu/transkribus/> (erişim 04 Kasım 2023).

²⁰ Matbu Osmanlı Türkçesi Yapay Zekâ Otomatik Transkripsiyon Modeli, <https://readcoop.eu/model/ottoman-turkish-print/> (erişim 04 Kasım 2023).

dijital edisyon için Osmanlı Türkçesi belgelerde önemli bir potansiyele sahiptir. Ayrıca Zooniverse²¹ (kitle-kaynak), noadgoat²² (veritabanı yönetim sistemi) gibi birçok platformla entegre çalışma özelliğine sahip olarak dijitalleştirilen metinlerin ileri düzey çalışmalarını da sağlamaktadır.

TEI yöntemiyle Osmanlı Türkçesi belgelerin dijitalleştirilmesi için –her ne kadar sayısı az olsa da– kıymetli çalışmalar bulunmaktadır. Örneğin *Corpus Musicae Ottomanicae*-CMO²³ projesi Osmanlı müziğinin ritmik ilkelerini göz önünde bulundurarak çoğunlukla çeşitli alfabelerle yazılmış şiirlerin bir dijital edisyonunu oluşturmak için TEI yaklaşımını benimsemiştir. Till Grallert, *Kânûn-ı Esâsî*'nin Arapça, İngilizce ve Osmanlıca dillerinde TEI kodlamasını gerçekleştirmiştir.²⁴ Gisela Procházka-Eisl ve Hülya Çelik tarafından Viyana Üniversitesi'nde yürütülen MECMUA Online projesi on altıncı yüzyıl âlimi ve şairi Nev'î'nin *Netâyicü'l-Fünûn* adlı ansiklopedisi başta olmak üzere çeşitli eserlerinin yazmalarını dijitalleştirmiştir.²⁵ Tamás Kiss, Osmanlıca belgelerde de kullanımı mümkün olan TEI XML görüntüleyici ve arama motoru yazılımı geliştirmiştir.²⁶ Avusturya Bilimler Akademisi Habsburg ve Balkan Çalışmaları Enstitüsü tarafından yürütülen QhoD projesi, on beş ve on altıncı yüzyılların başında iki imparatorluk arasındaki diplomatik ilişkilerin başlangıcından Birinci Dünya Savaşı'nın sonuna kadar Habsburg-Osmanlı diplomasisi hakkındaki kaynakları düzenlemek için başlatılan dijital bir altyapıdır.²⁷

Görüldüğü üzere Osmanlı Türkçesi metinlerin TEI gibi geçmişi uzun ve köklü bir dijitalleşme teknolojisiyle entegrasyonu oldukça az sayıdaki proje ve kişisel çalışmalarla sınırlıdır. Oysa tarihî dillerin incelenmesi için özel olarak

21 Zooniverse'de Osmanlı Türkçesi belgelere dair yürütülen bir proje için bk. OTurC Projesi <https://www.zooniverse.org/projects/skirmizi/ottoman-turkish-crowdsourcing> (erişim 04 Kasım 2023).

22 Nodegoat, <https://nodegoat.net/> (erişim 04 Kasım 2023).

23 CMO Projesi, <https://www.uni-muenster.de/CMOEdition/en/> (erişim 04.11.2023); Sven Gronemeyer, "Representing the Sung Poetry of Ottoman Art Music in a Critical Digital Edition in TEI XML," *MSGSÜ Sosyal Bilimler Dergisi*, s. 27 (2023): 61-90.

24 Till Grallert, <https://github.com/tillgrallert/kanun-i-esasi> (erişim 04 Kasım 2023).

25 Mecmua Online Projesi, <https://mecmua.acdh.oew.ac.at/home.html> (erişim 04 Kasım 2023).

26 Tamás Kiss, https://dh-ottoman.univie.ac.at/fileadmin/user_upload/k_dh_ottoman/Rumi-1-1.pdf (erişim 8 Kasım 2023).

27 QhoD Projesi, <https://qhod.net/>; <https://zenodo.org/records/10182359> (erişim 30 Kasım 2023).

tasarlanan TEI, Osmanlı Türkçesinde yazılmış edebî, hukuki, sanatsal metinler, yazmaları, arşivler, haritalar, gravürler ve diğer tüm belgeler için de kodlama kapasitesine sahiptir. Bu devasa Osmanlı *büyük verisi* TEI ile kodlandığında kültürel ve tarihsel bağlam içinde derinlemesine bir araştırma ortamı oluşacak ve böyle bir dijitalleşme süreci farklı disiplinlerden gelen araştırmacıların Osmanlı metinleri üzerinde iş birliği yapmasını da kolaylaştıracaktır. Özellikle tarihçiler, dilbilimciler, edebiyatçılar ve hatta sanat tarihçileri, TEI'nin sunduğu ortak format sayesinde bir araya gelebilir. Bu aynı zamanda Osmanlı çalışmalarında yeni dijital beşerî bilimler metodolojilerinin geliştirilmesine de zemin hazırlayacaktır. Nitekim dijital metin analizi, veri madenciliği ve görselleştirme gibi hesaplamalı teknikler, Osmanlı tarihine önemli perspektifler kazandırma potansiyeline sahiptir. Evliya Çelebi'nin *Seyahatnamesi*'nin İstanbul kısmı üzerine yapılan bir çalışmada Osmanlı Türkçesi metinler NLP ve çeşitli metin madenciliği teknikleriyle analiz edilmiş ve seyahatnamenin TEI ile dijital edisyon örneği sunulmuştur.²⁸

İlmî katkıların yanı sıra TEI ile dijitalleştirilen Osmanlı dönemine ait yazma eserler ve belgeler, fiziksel olarak da koruma altına alınmış olur. Böyle bir yaklaşımla dijitalleştirilen kütüphaneler ve arşivler, araştırmacıların yanı sıra kamunun da Osmanlı yazılı kültürüne ve tarihine çok daha kolay erişimini sağlayacaktır.

Burada altı çizilmesi gereken önemli bir husus ise Osmanlı metinlerinin çeşitliliğine dair standart bir etiketleme yaklaşımının geliştirilme ihtiyacıdır. Özellikle edebiyat, tarih ve dilbilim alanındaki araştırmacıların dijital-doğan –yani geleneksel kâğıt tabanlı kayıtların aksine, çeşitli teknolojiler sayesinde dijital biçime dönüştürülen ve saklanan– Osmanlı Türkçesi metinler için ortak bir transkripsiyon modeli ve dijital edisyonların etiketleri ve hiyerarşisi yönünde alınacak kararlara öncülük etmesi beklenmektedir. Bu noktada araştırmacılar ve projeler arasında Osmanlı Türkçesi özelinde konsensüs oluşturma çabaları, ortak standartlar ve yöntemler üzerinde anlaşma sağlanması çalıştaylar, seminerler ve projeler aracılığıyla gerçekleştirilebilir.

²⁸ Fatma Aladağ, “Osmanlı Çalışmalarında GPT'nin Potansiyeli: Evliya Çelebi Seyahatnamesinin NLP ve Metin Madenciliği ile Uygulamalı Analizi ve TEI Yöntemiyle Dijital Edisyonu,” *I. Evliya Çelebi Sempozyumu*, Ekim 2023 (preprint versiyon), https://www.researchgate.net/publication/375925805_Osmanli_Calismalarinda_GPT'nin_Potansiyeli_Evliya_Celebi_Seyahatnamesinin_NLP_ve_Metin_Madenciligi_ile_Uygulamali_Analizi_ve_TEI_Yontemiyle_Dijital_Edisyonu (erişim 30 Kasım 2023).

Osmanlı Türkçesi İçin TEI Uygulaması: *Küçük Mecmua*

Ziya Gökalp tarafından 1922-1933 yılları arasında yayımlanan *Küçük Mecmua*; fikir, edebiyat ve siyaset temalarında haftalık yayımlanan matbu Osmanlı Türkçesi bir dergidir.²⁹ *Küçük Mecmua*'da Ziya Gökalp'ın sosyoloji, felsefe, Türk medeniyeti tarihi, lisan, folklor ve gündelik meseleler üzerine yazdığı makalelerin yanı sıra masallar, manzumeler ve piyesler de bulunmaktadır. *Küçük Mecmua*, dönemin edebî akımlarını ve eğilimlerini gözlemleme fırsatı sunan önemli bir dergidir. Özellikle Türkiye'de edebiyat ve sanat alanında modernleşme sürecine rol oynamış ve yeni nesil yazar ve sanatçıların yetişmesine öncülük etmiştir.

Ziya Gökalp, *Küçük Mecmua*'nın farklı temadaki makalelerini ayrı kitaplar veya özetler hâlinde yayımlamıştır.³⁰ Ayrıca eser Şahin Filiz tarafından Latin harflere aktarılmıştır.³¹

Böyle bir edebî eserin dijital edisyonunun hazırlanması ve metinlerde yer alan kişi, yer, olay, etnik grup, tarih gibi tematik etiketlerle dizinli olarak çevrimiçi yayımlanması TEI ile mümkündür. Bir TEI uygulaması olarak bu çalışma kapsamında, Digital Ottoman Corpora (DOC)³² ekibi tarafından *Transkribus* yapay zekâ otomatik transkripsiyon platformunda dijitalleştirilen ve ilk kez dijital edisyona hazırlanan *Küçük Mecmua* dergisinin ilk sayısı kaynak olarak kullanılmıştır.

Derginin matbu Osmanlı Türkçesi'nin ilk sayısı öncelikle bahsi geçen *Transkribus* projesi kapsamında otomatik olarak transkribe edilerek dijitalleştirilmiştir. Daha sonra Latin harflerine dönüştürülmüş bu metinler içerisinde geçen her bir kişi, etnik grup, tarih ve yer ismi için *Transkribus* platformunda etiketleme (*tagging*) işlemi yapılmış ve TEI formatında çıktı olarak kaydedilmiştir. TEI, metinlerin satır bazında orijinal belgeleriyle transkripsiyonlarının eş zamanlı takip edilebilmesi için de etiketlemelere sahiptir (*Görsel 6*). Bu çalışmada faksimile kullanılmamış, transkribe edilmiş metin içeriğinin kodlanmış kısımları üzerinden dijital edisyon örneği hazırlanmıştır. 16 sayfadan oluşan dergi metni-

29 Alim Kahraman, "Küçük Mecmua," *TDVİA*, XXVI, 528-529, <https://islamansiklopedisi.org.tr/kucuk-mecmua> (erişim 28 Kasım 2023).

30 Ziya Gökalp, *Türk Medeniyeti Tarihi* (İstanbul: Matbaa-i Âmire, 1341 [1925]); Ziya Gökalp, *Türk Töresi* (İstanbul: Matbaa-i Âmire, 1339 [1923]).

31 Şahin Filiz, *Küçük Mecmua I-III* (Antalya: Yeniden Anadolu ve Rumeli Müdafaa-i Hukuk Yayınları, 2009-2010).

32 Digital Ottoman Corpora, Süphan Kırmızıaltın, Fatma Aladağ, Elif Derin Can, <https://www.digitalottomancorpora.org/> (erişim 11 Kasım 2023).

nin transkribe edilmiş versiyonu TEI XML formatında üst veri de dâhil olmak üzere kodlanmıştır (Görsel 7).

The screenshot displays the Transkribus platform interface. On the left, there is a 'Server Overview Layout Metadata Tools' panel with options like 'Document...', 'Find', 'Document Manager', 'User Manager', 'Versions', 'Jobs', 'Recent documents...', and 'User activity'. Below this is a 'Collections' section for 'KucukMecmua (236021, Owner)' and a 'Documents Model Data' table with columns for ID, Title, Pa., Uploader, and Uploaded. The table lists 15 documents, all with 'ederinca...' as the title and 'Mon Oct...' as the upload date. On the right, a document image is shown with a list of metadata items:

- 1-1 Sâhib-i İmtiyâz ve Müdiri-i Mesûl
- 1-2 Ziyâ Gökâlp
- 1-3 Yıl - 1 Sayı - 7
- 1-4 Nüshası 10 guruşdur
- 1-5 İdârehânesi
- 1-6 Diyarbekir
- 1-7 Hükümet Dâiresinde

6: Küçük Mecmua dergisinin Digital Ottoman Corpora tarafından Transkribus platformunda hazırlanan dijital edisyonu.

The screenshot shows the TEI XML code for the document. The code is displayed in a text editor with a menu bar (File, Edit, Navigate, Buffers, Application, XQuery, XML, Help, Login) and a toolbar (New, New XQuery, Open, Save, Close, Eval, Run). The code includes the following elements:

```

1 xquery version "3.1";
2 <?xml version="1.0" encoding="UTF-8"?><TEI xmlns="http://www.tei-c.org/ns/1.0">
3   <teiHeader>
4     <fileDesc>
5       <titleStm><title type="main">KucukMecmua_issue07_17Temmuz338-17July1922</title></titleStm>
6       <seriesStm><title>KucukMecmua</title></seriesStm>
7     <sourceDesc>
8       <bibl><title type="main">KucukMecmua_issue07_17Temmuz338-17July1922</title><idno type="Transkribus">1598172</idno><idno type="external">NA</idno></bibl>
9     </sourceDesc>
10  </fileDesc>
11  </teiHeader>
12  <text>
13    <body>
14      <div>
15        <pb facs="#facs_1" n="1" xml:id="img_0001"/>
16        <ab facs="#facs_1_r1" type="text">
17          <b facs="#facs_1_r11" n="N001"/>Sâhib-i İmtiyâz ve Müdiri-i Mesûl
18          <b facs="#facs_1_r113" n="N002"/><rs type="person">Ziyâ Gökâlp</rs>
19          <b facs="#facs_1_r112" n="N003"/>Yıl - 1 Sayı - 7
20          <b facs="#facs_1_r114" n="N004"/>Nüshası 10 guruşdur
21          <b facs="#facs_1_r115" n="N005"/>İdârehânesi
22          <b facs="#facs_1_r119" n="N006"/><rs type="place">Diyarbekir</rs>
23          <b facs="#facs_1_r1110" n="N007"/>Hükümet Dâiresinde
24          <b facs="#facs_1_r1112" n="N009"/>Vilâyet Matba'ası
25          <b facs="#facs_1_r1117" n="N009"/>Telgrafiyen:< rs type="place">Diyarbekir</rs>
26          <b facs="#facs_1_r1119" n="N010"/>Küçük Mecmua'a
27          <b facs="#facs_1_r116" n="N011"/>Abone Şeraiti
28          <b facs="#facs_1_r1111" n="N012"/>Seneliği: 500
29          <b facs="#facs_1_r1114" n="N013"/>Altı Aylık: 250
30          <b facs="#facs_1_r1116" n="N014"/>guruşdur

```

7: Küçük Mecmua dergisinin TEI XML kod mimarisinden bir görünüm.

Küçük Mecmua TEI kod mimarisini yakından incelemek XML dosyasındaki her bir etiketin hangi açıdan işe yaradığını anlamak açısından yardımcı olacaktır:

<teiHeader> → Başlık
 <fileDesc> → Metnin dosya açıklaması
 <title type="main">KucukMecmua_17Temmuz338</title> → Ana Başlık
 <title>KucukMecmua</title> → Bir metnin serisi ve koleksiyonu
 <sourceDesc> → Belgenin kaynağı
 <bibl><idno type="Transkribus">1598172</idno> → Bibliyografik bilgi
 ...
 <body> → Metnin asıl ve ana içeriği
 <div>
 <pb facs="#facs_1" n="1" xml:id="img_0001"/>
 → 'pb': sayfa başı, 'facs': faksimile, 'xml:id': belgenin kimlik numarası (ID)
 <lb facs="#facs_1_r111" n="N001"/>Sâhib-i İmtiyâz ve Müdür-i Mesûl
 → 'lb': satır başı, 'facs': faksimile, 'n': satır numarası

Metnin sayfa düzeni (*layout*) ve içeriği eş zamanlı olarak dijitalleştirilebilir. Böylece belgede derkenarlar, karalanan veya sonradan metne eklenen kısımlar, italik veya bold vurgular ve daha birçok detay ayrı kod etiketi ile orijinal belgede görüldüğü şekliyle dijital ortamda temsil edilebilir. Metinde içerikle ilgili etiketler (*tags*) kişiler, yer ismi, tarih ve etnik grup olarak ayrıca etiketlenmiştir:

<ethnic_group>Avrupalı</ethnic_group>larla → Metinde geçen etnik gruplar
 <rs type="place">İtalya</rs>'da → Metinde geçen yer isimleri
 <rs type="person">Karl Marks</rs> ile tarafdârları → Metinde geçen kişi isimleri
 <date>23 Zilka'de sene 340 Pazartesi 17 Temmuz sene 338</date> → Metinde geçen tarihler

Bu tematik etiketler her proje ve belgenin ihtiyacına göre artırılabilir ve çeşitlendirilebilir. Böylece kaynağın içerisinde önemli olduğu öngörülen her bir içerik mikro ve makro ölçekte erişilebilir. Böyle bir çalışmayı kullanıcı dostu bir arayüzle araştırmacıların erişimine açmak ve tematik etiketleri izlenebilir kılmak için bir web sitesinde dijital edisyonun yayınlanması mümkündür. Bu yaklaşım nispeten XML dosya formatlarının karmaşık kod görüntüsünü, dijital okuryazarlığı az olan araştırmacılar ve kamu için daha anlaşılır bir hâle getirmektedir.

XML dosyasının bir web sitesinde yayınlanması için HTML formatına dönüştürülmesi gerekmektedir. Bunun için de bazı ön işlemler ve teknik aşamalar gerçekleştirilmelidir. İlk olarak, TEI XML belgelerinin yapısal unsurların tanımlanması için gerekli olan XSLT formatına dönüştürülmesi daha sonra ise XML yapısındaki kodların hiyerarşisini doğru bir şekilde yorumlamak için gerekli olan bir DOM (Belge Nesne Modeli) ayrıştırıcısının oluşturulması gerekir. Bu işlemler, beşerî bilimciler için ek bir eğitim ve zaman gerektiren teknik bir yetenek ihtiyacı doğurmaktadır. Ancak son dönemde çığır açan bir teknoloji olarak araştırmacıların hayatında giren yapay zekâ platformu ChatGPT ile bu işlemler çok daha kolay ve kısa işlem adımıyla çözülebilir hâle gelmiştir. Nitekim TEI XML formatındaki *Küçük Mecmua* ChatGPT-4³³ Büyük Dil Modeli ile otomatik olarak HTML'e dönüştürülmüştür (Görsel 8). Böylece metin içerisinde etiketlenen temalar kullanıcılar tarafından bir liste şeklinde erişilebilir ve metin bağlamında takip edilebilir. Ayrıca metinde bahsedilen yer isimleri ve koordinat bilgileri yine GPT-4 ile otomatik olarak eşleştirilerek harita üzerinde gösterilebilir (Görsel 9). Her ne kadar bu çalışmanın sınırları kapsamında web sitesi basit ve temel düzeyde tasarlanmış olsa da bu teknoloji kullanılarak çok daha detaylı ve kapsamlı indeksler, kaynağın farklı versiyonları veya farklı dillerdeki yayınları eş zamanlı olarak yayınlanabilir ve görselleştirilebilir.

33 OpenAI, ChatGPT-4 <https://chat.openai.com/> (erişim 11 Kasım 2023).

Küçük Mecmua TEI

Küçük Mecmua: TEI ile Dijital Edisyon

Sâhib-i İmtiyâz ve Müdür-i Mesûl Ziyâ Gökâlp
 Yıl - 1
 Sayı - 7
 Nüshası 10 guruşdur
 İdârehânesi Diyarbekir Hükümet Dâiresinde Vilâyet Matba'ası
 Telgrafiyen: Diyarbekir
 Küçük Mecmû'a Abone Şerâti: Seneligi: 500 Altı
 Aylık: 250 guruşdur
 1922 - 338
 Haftada bir çıkar 'ilmi, edebî, siyâsî, iktisâdî mecmû'adır
 23 Zilkâ'de sene 340 Pazartesi
 17 Temmuz sene 338

Kişi Tarih Etnik Grup Yer Harita

Ziyâ Gökâlp
 Jar Jak Ruso
 Nâmîk Kemâl'in hem idâre
 Emrâh
 Mahmûd Es'ad
 Ahmed Midhat Efendiler
 Şemseddin Sâmî,
 Tevfik
 Fikret
 Karl Marks

© 2023 All Rights Reserved

8: *Küçük Mecmua*'nın TEI dosyasının HTML formatına dönüştürülmüş web sitesi görünümü.

Küçük Mecmua TEI

Küçük Mecmua: TEI ile Dijital Edisyon

Kişi Tarih Etnik Grup Yer Harita



© 2023 All Rights Reserved

9: *Küçük Mecmua*'daki yer isimlerinin haritalanması.

olarak etiketlenmesini de sağlamaktadır. Böyle bir çalışma ile kavramların duygu puanlarına göre metnin söylemi ve duygu analizi yapılabilir.

Bu tespitler bu makalenin sınırları kapsamında kısıtlı bir örnek olarak *Küçük Mecmua*'nın yalnızca ilk sayısı üzerinden nicel bir analiz sonucu ortaya çıkmıştır. Ancak uygulanan yöntemler sayesinde eserin tüm sayıları dâhil olacak şekilde çok daha detaylı nicel ve niteliksel bir metin analizi gerçekleştirilebilir. Ayrıca yazarın diğer eserleri ile kıyaslamaların yanı sıra aynı dönemdeki diğer edebi eserlerle de karşılaştırmalı analizler yapılabilir.

Sonuç

Bu çalışma, dijital beşerî bilimlerde uluslararası bir kodlama standardı olan TEI'nin Osmanlı Türkçesi özelinde tarih ve edebiyat disiplini içerisindeki potansiyellerini uygulamalı örnekler üzerinden incelemektedir. Bu kapsamda örneklem olarak çalışılan *Küçük Mecmua*'nın TEI kodlaması ve HTML tasarımı aracılığıyla, tarihî metinlerin nasıl dijital edisyonunun hazırlanacağı ve içeriklerinin analiz edileceği konusunda somut bir örnek sunulmuştur. Ayrıca, Voyant Tools ile gerçekleştirilen metin analizi, edebî eserlerin uzak okuma yaklaşımıyla tematik olarak incelenmesinde nasıl bir yol izlenebileceğine dair bir çerçevede sunmuştur. Böylece, tarihsel metinlerin evrensel bir bakış açısı ve standart ile erişilebilir ve anlaşılır olmasını sağlayan dijital araçların ve metodolojilerin, beşerî bilimlerde içerisindedir tarih ve edebiyat çalışmalarındaki kapasitesine dikkat çekilmesi hedeflenmiştir.

Kaynaklar

- Akis Projesi, <https://dhlab.sabanciuniv.edu/tr/akis-osmanlica-transkripsiyon-araci> (erişim 07 Ekim 2023).
- Aladağ, Fatma. "Dijital Beşerî Bilimler ve Türkiye Araştırmaları: Bir Literatür Değerlendirmesi." *Türkiye Araştırmaları Literatür Dergisi*, s. 18 (2020): 773-796.
- _____ ve Can, Elif Derin. "Programlamadan Yapay Zekâya Osmanlı Türkçesi için Metin Tanıma ve Otomatik Transkripsiyon." *Dijital Beşerî Bilimler ve Osmanlı Çalışmaları*, hazırlayan Yunus Uğur, 45-77. İstanbul: Vakıfbank, 2023.
- _____. "Osmanlı Çalışmalarında GPT'nin Potansiyeli: Evliya Çelebi Seyahatnamesinin NLP ve Metin Madenciliği ile Uygulamalı Analizi ve TEI Yöntemiyle Dijital Edisyonu." *I. Evliya Çelebi Sempozyumu*. İstanbul, Ekim 2023 (erişim 30 Kasım 2023).
- Bilgin Taşdemir, Esmâ F. "Printed Ottoman Text Recognition Using Synthetic Data and Data Augmentation." *International Journal on Document Analysis and Recognition*, no. 26 (2023): 273-287.
- Can, Y. Sait, M. Erdem Kabadayı. "Line Segmentation of Individual Demographic Data from Arabic Handwritten Population Registers of Ottoman Empire." *Document*

- Analysis and Recognition – ICDAR 2021 Workshops*, edited by Elisa H. Barney Smith and Umappeda Pal, 312–21. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-86198-8_22.
- CMO Projesi, <https://www.uni-muenster.de/CMOEdition/en/> (erişim 02 Kasım 2023).
- Digital Ottoman Corpora, <https://www.digitalottomancorpora.org/> (erişim 07 Kasım 2023).
- Digital Ottoman Studies, <https://www.digitalottomanstudies.com/> (erişim 07 Kasım 2023).
- Filiz, Şahin. *Küçük Mecmua I-III*. Antalya: Yeniden Anadolu ve Rumeli Müdafaa-i Hukuk Yayınları, 2009-2010.
- Grallert, Till. <https://github.com/tillgrallert/kanun-i-esasi> (erişim 04 Ekim 2023).
- Gronemeyer, Sven. "Representing the Sung Poetry of Ottoman Art Music in a Critical Digital Edition in TEI XML." *MSGSÜ Sosyal Bilimler Dergisi*, s. 27 (2023): 61-90.
- Gökalp, Ziya. *Türk Medeniyeti Tarihi*. İstanbul: Matbaa-i Âmire, 1341 [1925].
- _____. *Türk Töresi*. İstanbul: Matbaa-i Âmire, 1339 [1923].
- Ide, Nancy M. and C. M. Sperberg-McQueen, "The TEI: History, Goals, and Future." *Text Encoding Initiative Background and Context*, edited by Nancy Ide and Jean Véronis, 5-15. Springer, 1995.
- IRCICA, <https://library.ircica.org/Projects/OttomanOCR> (erişim 01 Kasım 2023).
- Kahraman, Alim. "Küçük Mecmua," *TDVİA*, c. 26, <https://islamansiklopedisi.org.tr/kucuk-mecmua> (erişim 28 Kasım 2023).
- Kırmızıaltın, Süphan and David Wrisley. "Automated Transcription of Non-Latin Script Periodicals: A Case Study in the Ottoman Turkish Print Archive." *Digital Humanities Quarterly* 16, no. 2 (2022), <https://www.digitalhumanities.org/dhq/vol/16/2/000577/000577.html> (erişim 04 Kasım 2023).
- LexiQamus, <https://www.lexiqamus.com/tr> (erişim 04 Kasım 2023).
- Matbu Osmanlı Türkçesi Yapay Zekâ Otomatik Transkripsiyon Modeli, <https://read-coop.eu/model/ottoman-turkish-print/> (erişim 04 Kasım 2023).
- Mecmua Online Projesi, <https://mecmua.acdh.oeaw.ac.at/home.html> (erişim 04 Kasım 2023).
- Muteferriqa, <https://mutferrriqa.com/> (erişim 04 Kasım 2023).
- Nodegoat, <https://nodegoat.net/> (erişim 04 Kasım 2023).
- OTurC Projesi, <https://www.zooniverse.org/projects/skirmizi/ottoman-turkish-crowdsourcing> (erişim 04 Kasım 2023).
- osmanlica.com, <https://www.osmanlica.com/> (erişim 04 Kasım 2023).
- QhOD Projesi, <https://qhod.net/>; <https://zenodo.org/records/10182359> (erişim 30 Kasım 2023).
- Sperberg-McQueen, C.M. ve L. Bumard. *Guidelines for Electronic Text Encoding and Interchange*. Chicago and Oxford, 1994.
- TEI, <https://tei-c.org/> (erişim 13 Eylül 2023).
- TEI by Example, <https://teibyexample.org/exist/> (erişim 13 Eylül 2023).
- TEI Publisher, <https://teipublisher.com/> (erişim 13 Eylül 2023).
- TEI Viewer, <https://alexandermendes.github.io/tei-viewer/> (erişim 13 Eylül 2023).
- Vincent van Gogh: The Letters, <https://teipublisher.com/exist/apps/vangogh/index.html> (erişim 13 Eylül 2023).
- Wikilala, <https://www.wikilala.com/> (erişim 04 Kasım 2023).