FIRAT ÜNİVERSİTESİ
SOSYAL BİLİMLER DERGİSİ
Journal of Social Sciences
p-ISSN:1300-9702 e-ISSN: 2149-3243

Osmanlıca-Türkçe Uçtan Uca Aktarım Sistemi-Osmanlica.com Projesinin Türk Dili ve Edebiyatı Alanına Katkıları

The Contributions of the End-to-End Ottoman Turkish to Turkish Conversion System-Osmanlica.com Project to the Turkish Language an Literature

Sare UZAN[1] ve İshak DÖLEK[2]

[1]Doktora Öğrencisi, İnönü Üniversitesi, Fen Edebiyat Fakültesi, Türk Dili ve Edebiyatı Bölümü, Malatya, sareuzan@gmail.com, orcid.org/0000-0003-1847-643X
[2]Dr. Öğr. Üyesi, İstanbul Medeniyet Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, ishak.dolek@medeniyet.edu.tr, orcid.org/0000-0002-5823-0103

Araştırma Makalesi/Research Article

Makale Bilgisi
Geliş/Received: 19.10.2024
Kabul/Accepted: 24.06.2025

DOI:
10.18069/firatsbed.1570398

Anahtar Kelimeler
Osmanlica.com, Osmanlıca OCR, edebiyat, alfabe ve dil çevirisi.

ÖZ

Türk dili ve edebiyatı alanında okutulan Osmanlı Türkçesi metinlerinin dijital ortamdan faydalanarak öğrenme sürecine ve metinlerin analizinin daha kolay yapılmasına nasıl katkı sağlayacağını belirlemek çalışmanın amacını oluşturmaktadır. Bu amaç doğrultusunda Osmanlı Türkçesiyle kaleme alınmış matbu (basılı) ve el yazması dokümanları Türkiye Türkçesine aktarmak için derin öğrenme modelleriyle metne dönüştüren web tabanlı Osmanlıca-Türkçe Uçtan Uca Aktarım Sistemi projesi kapsamında Osmanlica.com web sitesi oluşturulmuştur. Site; (i) Osmanlıca OCR/HTR (ii) Osmanlıca-Türkçe Alfabe Çevirisi, (iii) Osmanlıca-Türkçe Dil Çevirisi'nden oluşan aşamalar ile Osmanlıcadan Türkçeye aktarımı sağlamaktadır. Osmanlı Türkçesiyle yazılmış eserler web sayfasına; 1. Osmanlıca doküman (Resim) ile Türkçe (Latin alfabe) metinlerin satır satır eşleştirilmesi, 2. Eşleştirilmiş olan Resim-Latince metin satırların Osmanlıcaya çevrilmesi, 3. Resim-Osmanlıca metnin harf harf, kelime kelime kontrol edilmesi, 4. Osmanlıca metin ile Türkçe metnin kelime kelime eşleştirilmesi şeklinde veri kümelerinin hazırlanması adımlarından geçerek sisteme yüklenmektedir. Bu sayede el yazmaları ve nadir eserlerin korunması, bu eserlere kolaylıkla erişilmesi, kültür dünyamıza sahip çıkılması gibi hususlarda, çalışmanın önemi izah edilerek geçmişin birikimiyle geleceği inşa etme sürecine katkı sağlanması amaçlanmaktadır. Ayrıca Osmanlica.com sitesinin kullanıma sunduğu aruz vezni ve Türkçe sıklık analizine dair çalışmaların içeriğine değinerek Türk dili ve edebiyatı alanına katkıları ise örnekler üzerinden açıklanmıştır.

Keywords
Osmanlica.com, Ottoman OCR, literature, alphabet and language translation.

ABSTRACT

The aim of this study is to indicate how Ottoman Turkish texts taught in the field of Turkish Language and Literature will contribute to the learning process and enable easier analysis of texts by using the digital environment. For this purpose, the Osmanlica.com website was created within the scope of the web-based "End-to-End Ottoman-Turkish Conversion System" project, which converts printed and handwritten documents written in Ottoman Turkish into texts using deep learning models to convert them to Turkish. The site provides transfer from the Ottoman to the Turkish in stages: (i) Ottoman OCR/HTR (ii) Ottoman-Turkish transliteration, and (iii) Ottoman-Turkish translation. Works written in Ottoman Turkish can be added to the web page as follows; 1. Matching Ottoman document (Image) with Turkish (Latin alphabet) text line by line, 2. Translating the matched Image-Latin text lines into Ottoman, 3. Checking the Image-Ottoman text letter by letter and word by word, 4. Matching the Ottoman text with the Turkish text word by word and preparing the data sets are uploaded to the system. In this way, it aims to explain the work's importance in terms of protecting manuscripts and rare works, easily accessing these works, protecting our cultural world, and contributing to the process of building the future with the accumulation of the past. In addition, it aims to explain the content of the studies on Aruz meter and Turkish frequency analysis offered by the Osmanlica.com site and their contributions to the field of Turkish language and literature through examples.

Atıf/Citation: Uzan, S. ve Dölek, İ. (2025). Osmanlıca-Türkçe Uçtan Uca Aktarım Sistemi-Osmanlica.com Projesinin Türk Dili ve Edebiyatı Alanına Katkıları. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 35, 3, 907-924.
Sorumlu yazar/Corresponding author: Sare UZAN, sareuzan@gmail.com

1. Giriş

Osmanlı Türkçesi, Osmanlı Devleti sınırları içinde yaşayan kişilerin, Arap harfleriyle yazdığı konuşma ve yazı dilidir. Arapça ve Farsça kelimelerin geniş yer tuttuğu bu yazı dili, Eski Anadolu Türkçesi diye adlandırılan ilk dönemi de içine alarak 13. yüzyıldan 20. yüzyılın başlarına kadar varlığını koruması sebebiyle Türkiye Türkçesinden bağımsız bir yazı dili olarak düşünülemez (Özkan, 2007: 483). Türk milleti sahip olduğu geniş coğrafya üzerinde çeşitli diller ile etkileşime girmiş ve bu dillerden birtakım özellikleri kendi diline ya doğrudan ya da Türkçeleştirerek dâhil etmiştir. Bundan dolayı Osmanlı Türkçesi iki ayrı seviyede gelişim göstermiştir: İlk evre kimi noktalarda konuşma dilinden farklı olsa da geniş bir okur-yazar kitlesi tarafından kolayca okunup anlaşılan bir yazı dili hüviyetine sahiptir. Çünkü daha ziyade faydayı ön plana alan ve yazı dilini bir araç olarak kullanan bir özellik göstermektedir. Pek çok yazar, eserlerinin kolay anlaşılmasını ve geniş kitlelere hitap etmesini önceliğinden dolayı bu seviyede bir dil kullanılmasını yeğlemişlerdir. İkinci seviye ise üst dil de diyebileceğimiz daha zengin bir kelime kadrosu ve yabancı söz kalıplarının varlığı ile İslam kültürünün de dâhil olduğu yüksek zümreye hitap eden bir seviyedir. Bu yazı diliyle yalnızca edebî eserler değil, içtimai anlamda toplumun her alanına dair eserler verilmiştir (Tulum, 2011: 3-4). Bu durum Türk dilinin ve doğal olarak Osmanlı Türkçesinin zaman içerisinde başka dillerden de aldığı kelimeler sayesinde zenginleşerek ilerlediğini göstermektedir. Bu hususta geniş bir coğrafyaya hükmeden Osmanlı Türkçesi ile kaleme alınan eserler üzerinde çalışılması, dil ve kültür adına önem arz ettiği gibi geleceğin inşa edilmesine kaynaklık edeceği aşikârdır. Bu dilin Arapça ve Farsça gibi dillerden etkilenmesi ve özellikle İslam medeniyetine ait kelime ve terkipleri bünyesinde muhafaza etmesi, beraberinde mezkûr dillerden alınan kelimelerin bazılarının Türkçeleşmesine ve yeni anlamlar kazanmasına olanak sağlamıştır. “Kökü mazide olan âtî”ler olabilmenin ilk önemli adımı, yüzyıllar süren bu medeniyetin sahip olduğu eserleri araştırmak ve incelemektir.

Yurt içi ve yurt dışındaki devlet arşivlerinde, kütüphane ve özel koleksiyonlarda Osmanlı Türkçesi ile kaleme alınmış farklı türden eserler bulunmaktadır. Yüzyıllar boyunca ortaya konan eserlerin tümünün insan eliyle yazıya geçirilmesi, incelenmesi mümkün görünmemektedir. Dolayısıyla bu eserleri Türkiye Türkçesine aktararak tarihine, kültürüne ve diline sahip çıkmak isteyenler için yapay zekâ aracılığıyla geliştirilen yöntemler sayesinde kuşaktan kuşağa aktarılacak dijital bir kütüphanenin varlığına ihtiyaç duyulduğu bir gerçektir. Halil İnalcık bu hususa binaen “Bana Osmanlı arşivlerini verin, size bir kültür imparatorluğu kurayım.”[1] diyerek kültür tarihimize sahip çıkmanın ehemmiyeti üzerinde durmaktadır. Özellikle Arap harfli Osmanlı Türkçesi ile kaleme alınan eserlerin zengin kültür birikimini gelecek kuşaklara aktarmak için dil, edebiyat ve tarih alanında el yazmaların ve nadir eserlerin okunulması, araştırılması ve üzerine çalışmaların yapılması, kültür tarihimizin idrak edilmesi ve benimsenmesi noktasında elzem bir yer teşkil etmektedir. Bu amaçla farklı alanlarda Osmanlı Türkçesi ile kaleme alınan eserlere ulaşabilme ve bu eserlerin anlaşılabilmesi için Dijital Beşeri Bilimler alanında çalışmalar önem kazanmış ve bu çalışmalar için yöntem ve araçlar geliştirilmiştir. Özellikle tarihî ve edebî metinlerin bilgisayar ortamında işlenmesi ve analiz edilmesi noktasında pratik bir şekilde verim alabilmek için kodlama ve tarama işlemleri yapılmaktadır. Bu sayede farklı disiplinlerin temel kaynak niteliği taşıyan Osmanlı Türkçesi ile kaleme alınan yazıların, arşiv belgelerinin ve yazma eserlerin dijital ortamda incelenip yeniden teşekkül ve analiz edilmesi sayesinde bu kaynak eserlere erişilmesi ve eserler üzerinde derin bir analiz ağının sağlanması olanağı doğmuştur.

Doç. Dr. Atakan Kurt ve Dr. Öğr. Üyesi İshak Dölek, Osmanlı Türkçesi metinlerini dijitalleştirme noktasında güç ve karmaşık problemlere yapay zekâ aracılığıyla üç adımdan oluşan bir çözüm sunmaktadır. Yapay zekânın “görüntü işleme”, “doğal dil işleme” ve “derin öğrenme teknikleri”ni Osmanlı Türkçesine ve Türkiye Türkçesine uygulanmasıyla geliştirilen araçlar, Osmanlica.com sistemi üzerinde uygulanarak Osmanlı Türkçesiyle kaleme alınan dokümanlar, kolay bir şekilde Türkiye Türkçesine aktarılabilmektedir. Geliştirilen site üzerinden eklenen Arap harfli bir metin dosyası otomatik olarak Türkiye Türkçesine çevrilmektedir. Böylelikle Osmanlı Türkçesi öğrenimi daha kolay hâle geldiği gibi metin analizleri de gerçekleşebilmektedir. Site, Osmanlı Türkçesi gramer kuralları ve sözcük dağarcığı noktasında da aktif ve etkili bir şekilde öğrenim gerçekleştirilecek bir platform sunmaktadır. Sistem üzerine aktarılan eserler üzerinde koleksiyonların oluşturulmasıyla dönemsel anlamda metinlerin dil ve tema noktasında incelemelerin yapılması için ciddi bir veri ağı oluşmaktadır. Çalışmamız metinlerin dijital ortamda indekslenmesi sayesinde hem akademisyenlerin

[1] https://www5.tbmm.gov.tr// [Erişim Tarihi: 03.09.2024]

hem öğrencilerin istifade edeceği bir platformun varlığı ile interaktif bir eğitim sürecinin gerçekleşmesinin mümkün olacağının belirtme gayesi taşımaktadır. Osmanlica.com sitesi aracılığıyla aruz çözümleme ve Türkçe sıklık analizinin de yapılabileceği örnek şekiller üzerinden izah edilmiştir ve böylece Türk dili ve edebiyatı alanı için fayda sağlayacak bir platformun varlığı tanıtılmıştır.

2. Araştırmanın Konusu ve Önemi

Osmanlı Türkçesi ile kaleme alınan eserler aklâm-ı sitte; sülüs, rika, nesih, tevki, muhakkak, reyhani ve talik şekillerinden oluşmaktadır. Edebî eserlerde ve vesikalarda talik; devletin önemli belgelerinin yazımında rika; mali kayıtların yazımında ise siyakat gibi yazı şekilleri kullanılmıştır (Çetin, 1989: 276, Derman, 1997: 430). İslam dininde resim ve heykel gibi sanatların yasaklanması ile Osmanlı sanatkârları, yeteneklerini hüsn-i hatta ve tezhipte göstermişlerdir. Yazı çeşitlerinde özgün stili beraberinde getiren bu durum, Osmanlı Türkçesi ile kaleme alınan eserlerin incelenmesinde bazı müşkül durumları beraberinde getirmiştir. Bu durum ayrı bir başlık altında ayrıca tetkik edilecektir. Yapılan projede eserlerin bilgisayar ortamında incelenmiş olması, aslında bilgisayar algoritmasında Osmanlı Türkçesinin hem doğru okunması hem de dönemin yazı dilinde beliren farkların ortaya konması adına projenin önemini vurgulamaktadır. Çünkü projede kullanılan eserler, yazı çeşitliliğine ve her yazı çeşidiyle alakalı farklı belgelerin olmasına dikkat edilerek belirlenmiştir.

"Osmanlıca-Türkçe Uçtan Uca Aktarım Sistemi" adlı projede, Osmanlı Türkçesi ile kaleme alınan eserlerin Türkiye Türkçesi karşılıkları, bilgisayar ortamında bütünden parçaya doğru bir inceleme yolu ile verilerin oluşturulması ve veriler üzerinde analiz yapılarak yazı dilinin dijital ortamda doğru bir şekilde aktarılması, çalışmanın konusunu oluşturmaktadır. Eserlerin tafsilatlı bir düzlemde ele alınması, erişim kolaylığı sağlanması ve geniş kitlelere ulaştırılmasının hedeflenmesi, özellikle teknoloji çağı diye nitelendirilen bu yüzyılda projenin dijital alanda gerçekleşiyor olması oldukça kıymetlidir. Ayrıca Osmanlı Türkçesine ilgisi olanların kolaylıkla erişebileceği bir sitenin -proje vesilesiyle- olması dil şuurunun oluşmasını/gelişmesini gündeme getirmektedir.

Belgelerin bilgisayar ortamında herhangi bir değişikliğe ve bozulmaya maruz kalmadan kullanılabilir olması, bir imparatorluğun kültürel mirasını disiplinler arası çalışmalar vesilesiyle dijital ortama aktarılma ihtiyacını doğurmuştur. Beşeri bilimler arasında teknolojik araçların sağladığı imkânlardan faydalanma ve yeni şeyler üretme noktasında özellikle edebiyat ve tarih biliminin sunduğu olanaklardan istifade etmek adına yapılan bu çalışma, yazı stillerinin çeşitli olması ve kimi karmaşık sayfa düzenlerinin varlığını içeren metinlerin varak yapısı ve görsel nitelikleri, dijital ortamda orijinal hâline en yakın biçimde aktarılarak -Arap harfli bir metin dosyasının Latin harfe doğru bir şekilde- çevrilmesi amaçlanmıştır. Ayrıca proje kapsamında geliştirilen *Osmanlica.com* sitesi üzerinde erişime sunulan ve Türk dili ve edebiyatı alanına yeni bir bakış açısı sağlanılacağı düşünülen "Bâkî" isimli aruz vezni çözümleme aracı ve "Ergin" isimli yazılım ekseninde geliştirilen Türkçe sıklık analizinin yapılacak çalışmalara kaynaklık edebileceğini belirtmek de çalışmanın amacını teşkil etmektedir.

3. Osmanlı Türkçesi Metinlerin Dijitalleşmesinde İzlenen Yöntemler

Osmanlı Türkçesiyle kaleme alınan metinlerin dijitale aktarılmasında Osmanlıca OCR, Osmanlıca-Türkçe alfabe çevirisi, Osmanlıca-Türkçe dil çevirisi olmak üzere üç adımdan oluşan ve her bir adım için de geliştirilen üç ayrı araç kullanılmaktadır (Dölek, vd., 2022: 2). Çalışmamızda bu adımların gerçekleşmesi için Osmanlı Türkçesiyle ilgili dokümanların (.png, .jpg, .pdf vb.) bilgisayar ortamında düzenlenebilir metin (editable text) hâline gelmesi ve çalışabilmesi için Osmanlı Türkçesi dokümanı yazım şekliyle (nesih, rika, talik, sülüs vb.) alakalı veri setlerin toplanması ve hazırlanması gerekmektedir. Yüzyılları kapsayan bir süreçte kullanılan Osmanlı Türkçesinin karmaşık dil yapısı ve farklı yazı türleri, veri setlerinin oluşturulmasında önemli zorluklar yaratmaktadır. Özellikle el yazması metinlerin okunabilirliği ve yapay zekâ tarafından doğru bir şekilde analiz

edilebilmesi için yüksek kalitede görüntü verilerine ihtiyaç duyulmaktadır. Bu sayede, metinlerin içeriği daha doğru ve eksiksiz bir şekilde çıkarılabilmektedir. Veri seti oluşturulurken metinler ister matbu ister el yazması şeklinde yazılmış olsun asıl gaye veri çeşitliliğinin geniş olmasıdır. Veri çeşitliliği ise rika, divani, siyakat vb. yazı stillerinden ziyade hangi yazı stili üzerinde çalışılıyorsa hazırlanacak veri setlerinin de o yazı türünden farklı metin türlerinin içermesine dikkat edilmektedir. Bu sayede veri setlerindeki başarı yüzdesinin arttırılması hedeflenmektedir. Örneğin, matbu metin özelinde hikâye, roman, şiir vb. edebî eserlerle birlikte tarih, astronomi, resmî yazışmalar ve gazeteler gibi farklı kelime ve cümle yapılarını barındıran belgelerin veri çeşitliliğine dikkat edilmektedir.

Veri setleri hazırlanırken ayrıca transkripsiyon alfabesini kullanıp kullanmamak ya da sınırlı ölçüde kullanmak gibi seçeneklerden biri tercih edilmektedir. Bu hususta önemli olan hazırlanan verilerin standart bir yapıda olmasıdır. Kısacası veri setleri hazırlanırken metinlerin toplanması, tasnif edilmesi ve dijitalleştirilip işlenmesi gibi aşamalar dikkatli bir şekilde yapılmaktadır. Osmanlı Türkçesiyle yazılmış belgeler için gerekli veri seti hazırlandıktan sonra karakter/kelime tanıma işlemleri yapılmıştır.

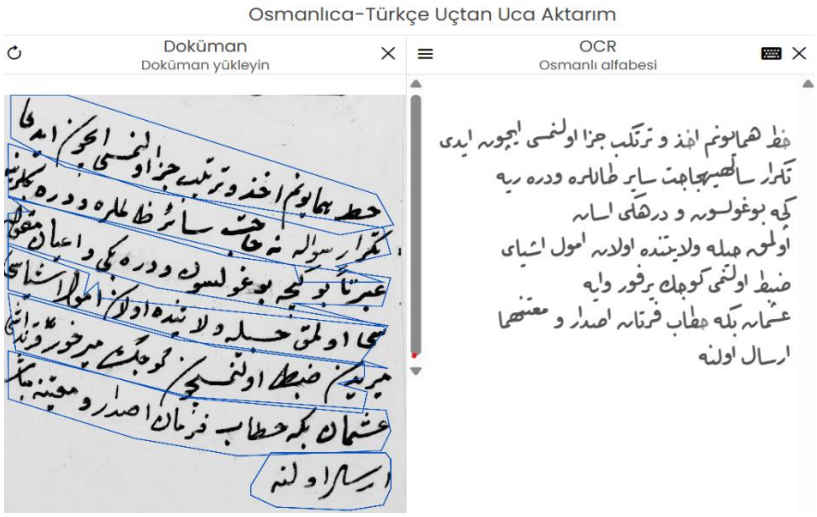
3.1. Osmanlıca OCR

Osmanlica.com projesinin alt modülü olan Osmanlıca OCR metinlerin görüntüsünün bilgisayar ortamına aktarılması için bir resim üzerinde saklanan görüntüdeki karakterlerin işleme ve makine öğrenme tekniklerini kullanarak metne (string) dönüştürme işlemi olan OCR sistemi kullanılmıştır. OCR, esasında karakterlerin tasnif edilme problemidir. El yazmaları ve nadir eserlerin görüntülerdeki metinleri düzenlenebilir metin hâline getirme işlemi olan OCR, "bazen bağlı-karakter-kataı (connected component) tanıma problemi veya kelime tanıma problemi olarak da ele alınmaktadır." (Dölek, vd., 2022: 34). OCR sisteminde işlemler temelde beş adımdan oluşmaktadır:

1. Görüntü ön işleme (image preprocessing): Bu aşamada bir görüntü dosya formunun ve boyutunun normal hâle gelmesi, gerekli olduğu takdirde metin formunun ve boyutunun ayarlanması, görüntünün filtrelenmesi, eşikleme (tresholding), siyah ve beyaz forma dönüştürme (binarization), gürültüyü azaltma/temizleme (noise reduction), görüntü iyileştirme (image enhancement) gibi ön aşamaları içerir. Buradaki asıl gaye görüntünün çözünürlüğünü en iyi hâle getirmektir.
2. Görüntü bölümleme (image segmentation): Bilgisayarın kullanılan görüntüdeki karakterleri tanıması için görüntünün tanınacak birimlere ayrılması gerekmektedir. Bu süreç, sayfa düzeni analizi, sayfa segmentasyonu, satır, kelime ve karakter seviyesinde bölümlenme gibi görüntü işleme teknikleriyle desteklenir. Bu sayede, metnin görsel temsili içindeki sınırları belirlenir; dil, alfabe, yazı tipi ve stil gibi metinsel özellikler tespit edilir. Bu ön işleme adımları, OCR sisteminin doğruluğunu ve verimliliğini artırmak için kritik öneme sahiptir.
3. Özellik çıkarma (feature extraction): Sınıflandırıcı bir modelle yapılan görüntüyü tanıma işlemi için bazı özelliklere ihtiyaç duyulabilir. Modelin kullanmasını sağlayan ve baştan belirlenen özelliklerle birlikte görüntüdeki genişlik ve yükseklik, yakınlaştırma ve uzaklaştırma gibi farklı parametreler ve geometrik özellikler kullanılabilir.
4. Tanıma (recognition, OCR): Görüntüde bilgisayar tarafından tanınması istenilen kısmın (eğri, karakter, alt-kelime, kelime) bir model aracılığıyla tanıma aşamasıdır.
5. Düzenleme ve hata düzeltme (post processing & error correction): Görüntünün tanınmasının ardından OCR çıktısındaki karakterlerin nomalizasyonu ile OCR'da tanınmayan karakter veya kelimelerin doğru tahmini için yapılan düzenleme ve hata giderme işlemi aşamasıdır (Dölek, vd., 2021: 400). OCR aracılığıyla bir metin dosyasının bilgisayara aktarılma işlemi aşağıda verilmiştir:

Osmanlıca matbu veya el yazmalarını dijital ortamda düzenlenebilir metin haline getirmek amacıyla geliştirilen Osmanlıca OCR sistemi, derin öğrenme tabanlı yaklaşımlar sayesinde önemli ilerlemeler kaydetmiştir. Denetimli öğrenme (supervised learning), etiketli veriler kullanarak modelin eğitildiği bir öğrenme tekniğidir. Daha açık bir ifadeyle, giriş (input) verilerine karşılık gelen doğru çıkışlar (output) (etiketler) önceden belirlenmiştir ve model, girdileri bu etiketlere uygun şekilde tahmin etmeyi öğrenir. Bu bağlamda, CRNN (Convolutional Recurrent Neural Network) mimarisi kullanılarak Osmanlıca OCR modeli eğitilmiş ve optimize edilmiştir. Modelin başarısını artırmak için öncelikle görüntü ön işleme aşamasında metinlerin boyutları normalize edilmiş, gürültü azaltma ve eşikleme teknikleri uygulanarak görsel kalitesi iyileştirilmiştir. Ardından, görüntü bölümleme süreçleri ile sayfa, satır ve karakter bazında segmentasyon gerçekleştirilmiş, metnin yapısal özellikleri belirlenmiştir. Osmanlıca OCR sistem, Osmanlıca metinlerin yüksek doğrulukla

okunmasını sağlayarak tarihî belgelerin dijitalleştirilmesine ve dilbilimsel çalışmaların genişletilmesine katkıda bulunmuştur:



Şekil 1. Osmanlıca OCR aracı

3.2. Osmanlıca-Türkçe Alfabe Çevirisi

Alfabe çeviri kısmı, Osmanlı Türkçesiyle yazılmış metinlerin transkripsiyon edilmesi şeklinde denilebileceği gibi Arap harflerin Latin harflere çevrilmesi olarak da yorumlanabilir. Bu durum sadece Türkçe eserler için söz konusu olmadığı gibi "kaynak alfabedeki bir metnin hedef alfabeye aktarımı olarak tanımlanabilir. Bu aktarımda amaç hedef dilde yazma veya hedef dilde okuma olabilir." (Dölek, 2022: 17). Dolayısıyla hedef dilde yazım noktasında yapılan aktarım "ortografik alfabe çeviri" olarak adlandırılırken okuma maksadıyla yapılan aktarım ise "fonetik alfabe çevirisi" şeklinde isimlendirilmektedir. Ortografik çeviri asıl metindeki harflerin, hedef dile aktarılmasını temsil ederken fonetik çeviri ise transkripsiyon şeklinde bilinmektedir. Alfabe çevirisi hedef ve kaynak alfabedeki seslerin karşılığının eşlenmesi olarak ele alınmaktadır. Ortografik alfabe çevirisi aşağıda verilen adımlar ile gerçekleştirilmektedir:

1. Ön İşlemler
a. Metin dönüştürme ve temizleme
b. Metin normalizasyonu
c. Cümle ve kelime ayrıştırma
d. Varlık ismi tanıma
2. Ortografik Alfabe Çevirisi
a. Osmanlıcada gövdeleme
b. Osmanlıca-Türkçe sözlüklü aktarım
c. Türkçe bitiştirme
3. Hata Düzeltme
a. Kelime Bölümleme
b. Yazım Düzeltme
c. Seslendirme
4. Son İşlemler
a. Kelime tahmini

b. İsim tamlaması ve birleşik kelimeler

Osmanlıca-Türkçe alfabe çevirisi için geliştirilen sistem, geleneksel Doğal Dil İşleme (NLP) yöntemlerine dayanmaktadır ve çeviri sürecini dört ana aşamada gerçekleştirir. İlk olarak, ön işleme aşamasında metin temizlenir, normalleştirilir ve kelimelere ayrıştırılarak varlık isimleri belirlenir. Ardından, ortografik alfabe çevirisi kapsamında Osmanlıca kelimeler gövdeleme yöntemiyle analiz edilir, sözlük bazlı aktarım gerçekleştirilir ve Türkçede uygun ekleme kurallarıyla yeniden oluşturulur. Hata düzeltme aşamasında, kelime bölümleme, yazım kontrolü ve seslendirme yöntemleri kullanılarak doğruluk artırılır. Son olarak, kelime tahmini algoritmaları ve isim tamlaması kuralları uygulanarak akıcı ve anlamlı bir metin elde edilir. Bu sistem, Osmanlıca metinlerin Türkçeye daha doğru ve okunabilir şekilde aktarılmasını sağlayarak dil tarihi ve edebiyat çalışmalarına katkıda bulunur.

Alfabe çeviri kısmında özellikle çalışmanın içeriği noktasında Arap harfleriyle yazılmış bir metnin Latin alfabesindeki karşılığını edinmek için izlenen yöntemler, metnin çeviri noktasında doğruluk oranını artırmak ve alfabe-kelime-cümle ve paragraf şeklinde birbirini takip eden parçadan bütüne doğru giden süreçte bilgisayarın algoritmasındaki kavrayışı hızlandırmak için geliştirilmektedir. Alfabe çevirisinde Osmanlı Türkçesiyle yazılmış metindeki harflerin karşılığı olarak Latin alfabesindeki karşılıklarının birebir çevrilmesi hedefi olsa da bilgisayarlı alfabe çevirisinin yapısı gereği karmaşık bir durumla da karşılaşılmaktadır. Çeviri esnasında sadece alfabelerin eşlenmesi koşulu olmamakla birlikte imla ve gramer yapılarına da dikkat edilerek eserin orijinal hâline uygun bir çeviri amaçlanmıştır:

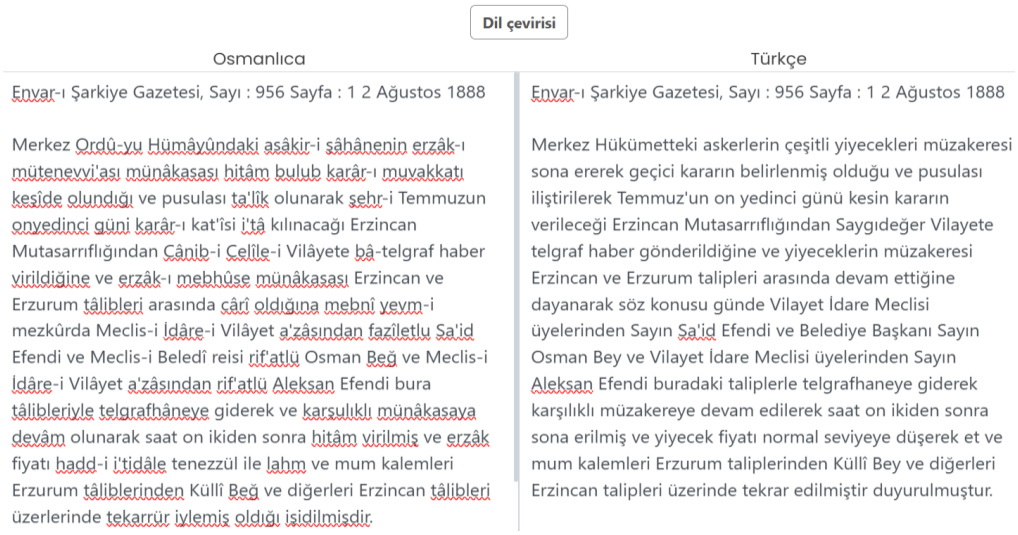


Şekil 2. Osmanlıca-Türkçe alfabe çevirisi aracı

3.3. Osmanlıca-Türkçe Dil Çevirisi

Osmanlıca-Türkçe dil çevirisi adımında Türk dili ve edebiyatında metnin Türkiye Türkçesine aktarıldıktan sonra manayı ifadelendirmek için kullanılan "dil içi çeviri" kısmı ifade edilmektedir. Bir metnin belirli kurallar çerçevesinde sadeleştirilmiş versiyonu olarak da tarif edebileceğimiz bu çeviri, bilgisayar üzerinden gerçekleştiğinden dolayı doğal dil işlemenin en mühim ve en zor problemini de beraberinde getirmektedir. Bir dilin başka bir dile çevirisi noktasında oluşan problemlerin zorluğunun aksine yazı dili farklı olup aynı dili temsil eden yani Osmanlı Türkçesinden Türkiye Türkçesine aktarımın daha basit yöntemlerle gerçekleştiği görülmüştür. Dil içi çevirisi yapılırken kelime grupları tabanlı şeklinde birebir çeviri yapılacağı gibi derin öğrenme yöntemleriyle cümle tabanlı makine çevirisi yaklaşımlarıyla da çeviri yapılabilir. Bunun için de "Çeviri öncesi (pre processing), sırası ve sonrasında (post processing) yüzeysel veya tam yapı bilimsel ve/veya söz dizimsel çözümleme, üretim, belirsizlik giderme (shallow/full morphological & grammatical parsing,

generation, disambiguation), kelime anlamı belirsizliği giderme (Word sense disambiguation), varlık ismi tanıma (named entity recognition) vb. işlemlerin yapılması gerekmektedir." (Dölek, 2022: 37). Şekil: 3'te Envâr-ı Şarkiye Gazetesi'nin 956. Sayısının ilk sayfalarında yer alan metin, çeviri aracı kullanılarak dil çevirisi verilmiştir:



Şekil 3. Dil çeviri aracı

4. Osmanlı Türkçesi ile Kaleme Alınan Metinlerin Yapay Zekâya Aktarılmasında Karşılaşılan Güçlükler

Osmanlı Devleti'nin geniş bir coğrafyaya hükmetmesi beraberinde dil ve kültür alanında zengin bir etkileşim alanı oluşturmuş ve bu durum dönemin eserlerine de yansımıştır. Özellikle Arapça, Farsça ve Türkçeden oluşan metinlerde kelimelerin ait olduğu dilin kendi dil bilgisi kuralları esas alınmıştır. Bu yazı dilinde bir kelime Arapça ve Farsça kökenli ise kelimenin yazılış şekli de o dillerdeki imla kurallarına göre yazılmıştır.

Eski Anadolu Türkçesi ile kaleme alınan el yazmalarının ilk dönemine bakıldığında kelimelerin doğru okunması için *harekeleme* diye adlandırılan simgeler kullanılmıştır. Sonraki dönemlerde ise imlanın klişeleşmesi ile birlikte sade yazılan metinlerde kısmen ayet, hadis ve Arapça ibarelerle, birbirine karıştırılması muhtemel olan kelimelerin harekeli olduğu görülmektedir (Ece, 2015: 253). Harekeli metinlerde kelimelerin telaffuz edildiği fonemden farklı bir şekilde yazılması özellikle dijitalde Latinize edilirken standart bir dijital yazı dili oluşturma noktasında problemlerle karşılaşılmıştır. Örneğin, "دُعَا اِتْدِي", "كيجه, وير" "dua itdi, gice, vir", kelimeleri "dua etti, gece, ver" şeklinde Latinize edilmektedir.

Projenin metin eşleştirme kısmında yer alan bir kişinin karşılaştığı ana güçlük -Osmanlıca üzerine çalışan herkesin de kabul edeceği gibi- alfabe sorunu ve metinleri günümüz alfabesine aktarmaktır. Ayrıca projede kullanılan araçların ve materyallerin yetersiz kaldığı noktalar da bulunur. Bu ve benzeri nedenlerle hem yapay zekânın metin tanımlaması hem de standart bir çeviri oluşturma yolu zorlaşmaktadır. Bundan dolayı Osmanlı Türkçesinin bilgisayar algoritmasında tanınır hâle gelmesi için yürütülen çalışmada bazı güçlüklerle karşılaşılmıştır. Bu güçlükler, Osmanlı Türkçesiyle yazılan metinlerde yazım birliğinin ve dilin standartlaşmış gramer yapısının olmamasından kaynaklanmaktadır. Problemlerin en aza indirgenmesi ve birçoğunun üstesinden gelmek için kesbî ve tecrübi bilgiler gerektirdiği aşikârdır. Bu hususta çalışma boyunca karşılaşılan güçlükler/zorluklar aşağıda listelenmiştir:

a. Osmanlı Türkçesi, Arap alfabesine ek olarak Farsçadan üç harfin eklenmesiyle 31 harften oluşan yazı dilidir. Bu yazı dilinde Latin alfabesinde bulunan "a,e,ı,i,o,ö,u,ü" sesli harflerine karşılık olarak Arap

alfabesinde bu sesleri karşılayacak "**ا،و،ى**" harflerinin kullanılması kelimelerin bağlam noktasında doğru şekilde algılanması sorununu ortaya çıkarmıştır. Örneğin metin içerisinde kullanılan "**كوز**" kelimesinin "göz, güz, köz, küz" şeklinde okunması.

b. Arap harfli alfabede kullanılan bazı karakterin Türkiye Türkçesinde karşılıklarının birden fazla sese tekabül etmesi bir başka problemi gündeme getirmektedir. Örneğin **ض** harfi "d,z" seslerine karşılık gelirken "**ط**" harfi "d,t" seslerine denk gelmektedir. Latin alfabesindeki "t" sesine karşılık olarak da "**ت، ط**" harflerinin olması, yine "k" sesi için "**ق،ك**" harflerinin kullanılması metinden metine değişiklik göstermekle birlikte bazı meseleleri de beraberinde getirmektedir. Yine Latin alfabesinde bulunan "s" sesini karşılayacak "**ث،س،ص**" ve "z" sesini karşılayacak "**ض،ظ،ز،ذ**" harfleri ile "h" sesini karşılayan "**ح،خ،ه**" harfleri ile yazılan kelimelerin Latinize edilmesinde de aynı problemler görülmektedir. Örneğin, "**صورت**, **ظهور**، " kelimesi Türkiye Türkçesine aktarılırken **ذوخور، سورەت**" şeklinde yazılabilmektedir.

c. Osmanlı Türkçesi metinleri her ne kadar Türkiye Türkçesi sözdizimi kurallarına göre yazılsa da başta İslam dininden alınmış özel isimler, terim ve kavramlar olmak üzere faklı dillerden kelime ve kelime yapılarını da bünyesinde bulundurmaktadır. Arapça ve Farsçadan alınan kelimeler aslına sadık kalınarak doğrudan alınsa da Rumca, İtalyanca, Ermenice, Yunanca gibi kelimeler fonetik esaslarına göre yazıya geçirilmiştir. Bundan dolayı mezkûr dillerden alınan kelimelerin doğru okunup çevrilmesinde bazı güçlükler meydana gelmiştir. Örneğin bir metinde "Elena demekle marûf ve Mar Sikla ve Sittinâ-yı ve Mar Eftimyus ve Mar Katrina ve Mar Mîkâil ve Mar Yorgi ve Mar Yuhanna" ibaresi geçmekte olup ve buradaki özel isimlerin okunuşlarını bulmak zordur: **النه ديمكله معروف و مار ثقله و ستى نای و مار افتميوس و مار كاترنه و مار مكائل و مار يورگى و مار يوحنا**

d. Türkiye Tükçesinde görülen ünsüz benzeşmesi ses olayının Osmanlı Türkçesinde çoğu metinde görülmediği anlaşılmıştır. Örneğin Osmanlı Türkçesi yazı dilinde "**كيتدى**" şeklinde yazılan "gitti", "**صابتدر**" kelimesi "sabittir" şeklinde Latinize edilmiştir. Bu durumun bilgisayar ortamında algılanması için metinlerin satırlara bölünmesi ve ardından harf, kelime, cümle ve paragraf takipleri proje sürecinde gerçekleşmiştir.

e. Osmanlı Türkçesinde çoğul ekleri, küçültme ekleri (ism-i tasgîr) kimi kelimelerde kendisinden önce birleşen harfler olmasına rağmen ayrı yazılmıştır. Soru edatları ise ayrı yazılması gerekirken bitişik yazılmıştır. Bu durum Osmanlı Türkçesinde yazım ve imla birliğinin olmamasından kaynaklanmaktadır. Örneğin "**حركتلر**" şeklinde yazılması gereken bir ibare "**حركت لر**" şeklinde yazılmıştır.

f. Arapça tamlamalarda muzâf olan kelimenin sonunun ötreli, fethalı ve kesreli okunuşuna göre kelimelerin yazımında Latinize edilirken bu harekelerin okunuşu ses olarak bilgisayar tarafından algılanmayabilmektedir. Örneğin "**بعدالزوال** ba'de'z-zevâl" tamlaması "**بعدالزوال**", "**با امرالله**, bi-emri'llâh", "**بى امرالله**" şeklinde yazılabilmektedir. Tüm kelimeleri Arapça olan cümleler Latinize edilirken birebir çeviride problemler çıkabilmektedir. Örneğin, "**حديث معجز البيان سلطان تختكاه لولاك**" , "hadîs-i mu'cizü'l-beyân-ı sultân-ı tahtgâh-ı Levlâk" Arap harfli cümlenin Türk alfabesiyle yazımında diyakritik yani ayırıcı işaretlerinin yazımda standartlaşmaması çeviride müşkülata sebep olmaktadır.

g. Aynı yazı hattıyla kaleme alınan eserlerde yazarın ya da müstensihin yazı stilindeki farklılıklar ve varakların türünden ve çözünürlüğünden kaynaklanan sorunlardan ötürü metinlerin çevirisinde müşkül durumlarla karşılaşılmıştır.

h. Osmanlı Türkçesinde yazılışları ile okunuşları Türkiye Türkçesinden farklı olan kelimelerde metin birliği sağlamak amacıyla Arap harfleriyle yazılışlarından ziyade bu kelimelerin Türkiye Türkçesi karşılıkları esas alınmıştır. Örnek: gelüb: gelip, nerdübân: merdiven, mutbak: mutfak kettân: keten, hemân: hemen, furça: fırça.

i. Arapça gramer yapısında görülen harf-i cerr yani ön edatların yazımında metin görselinden Arap harfine çeviri kısmında yanlış yazılışlar görülebilmektedir: " **الى يوم القيامه** ilâ-yevmi'l-kıyâme" şeklinde görselde yazılan Arapça bir tamlama "**الا يوم القيامه**" şeklinde çevrilebilmektedir.

j. Osmanlıcada hem matbu hem de el yazısında yaygın olarak karşımıza çıkan önemli problemlerden bir tanesi de kelime bölünmesi ya da bitişmesi problemidir. Aslında bitişik olarak yazılan Arap alfabesini kullanan Osmanlı alfabesinde önceki veya sonraki harfle bitişmeyen harflerden dolayı

kelime içi boşluklar (zero width non-joiner) söz konusudur. Kelime içi boşluklar sık sık kelime arası boşluklarla karıştırılmaktadır. Osmanlıca yazılırken yazıyı yazanlar kelime bölünmesi veya bitişmesini bir problem olarak görmedikleri için yazıda kelime içi ve kelime arası boşlukları verirken belli bir kurala göre hareket etmemektedirler. Bu durum OCR ve HTR yapılırken hatalı sonuçlar üretilmesine sebep olmaktadır. Bu yüzden bir post-processing adımı ile bu hataların ayrıca ele alınması gerekebilir.

k. Tarihî dokümanlarda yıpranma, lekeler, çizikler ve renk solması gibi sorunlar görüntü kalitesini kötüleştirmektedir. Bu durumda yazıdaki karakterler kısmen veya tamamen okunmaz hâle gelebilmektedir. Böyle durumlarda ön işleme adımlarında görüntü kalitesini arttıran temizleme, zenginleştirme veya büyütme (cleaning, enhancing, upscaling) gibi yöntemler kullanılabilir.
l. Osmanlıcada birçok harfin ana gövde şekli aynı olup bu harfler sadece noktalarla ayırt edilmektedir. Örneğin Be, Te, Se, Pe (ب، ت، ث، پ) gibi harflerin ana gövdeleri aynıdır. Bu harfler üst veya alta bir, iki ve üç nokta konularak ayırt edilmektedir. Noktaların silinmesi veya birleştirilmesi karakterlerin tanınmasını zorlaştırmaktadır.
m. Osmanlıca dilinin, sözcüklerin telaffuzu ile yazılışı arasında birebir tutarlı bir ilişki bulunmayan fonetik bir yapıya sahip olması, Türkçe dilinin ise fonetik bir yapıya sahip olması, iki dil arasındaki temel farklardan biridir.
n. Osmanlıca ve Türkçe dillerinin, sözcüklerin yazılışına ilişkin ortografik kuralları bakımından birbirlerinden farklı sistemlere sahip olduğu görülmektedir.
o. Osmanlıca ve Türkçe dillerinde kullanılan ses ve harf envanterleri arasında hem niceliksel hem de niteliksel farklılıklar bulunmaktadır. Bu durum, iki dil arasındaki fonetik ve grafemik sistemlerin birbirinden farklılaşmasına neden olmuştur. Osmanlıca harflerinin, Türkçe harfleriyle birebir değil, çoğunlukla birden fazla harfe karşılık geldiği gözlemlenmektedir.

Dijitale aktarılan metinler üzerinde dil çalışmalarının yapılması noktasında filologların, bilgisayar ve yazılımcıların disiplinler arası çalışmaları ile bu problemlerin çözüme kavuşturulması noktasında da adımlar atılabilir.

5. Araştırmanın Türk Dili ve Edebiyatı Alanına Katkıları

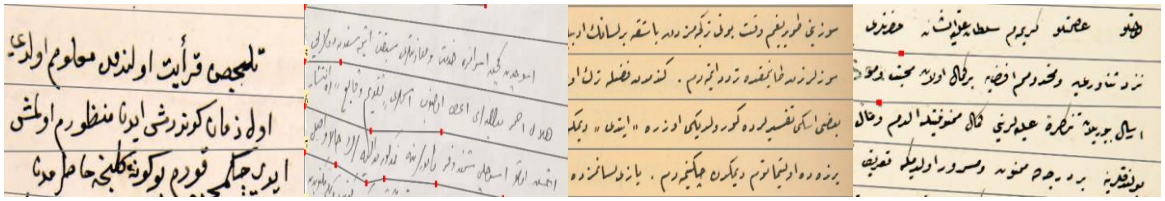
Dil ve edebiyat noktasında elyazmaları ve nadir eserler üzerinde birçok çalışma yapılmıştır/yapılmaktadır. Çeşitli türden eserler akademik bir çalışma ya da hususi çalışmalar olarak bilim dünyasına sunulmuştur. Yapılan çalışmaların ışığında özellikle teknolojinin günbegün ilerlemesi ile birlikte böylesi mühim eserlerin raflarda kalmaması ve daha kolay erişilebilmesi için de çeşitli projeler yapılmaktadır. Bu projelerden biri olan *"Osmanlıca-Türkçe Uçtan Uca Aktarım Sistemi – Osmanlica.com"* sayesinde özellikle işlenmiş eserlerin okunması ve incelenmesi çok daha kolay olacaktır. Projede yukarıda bahsedilen çalışmalar sonucunda bir belgenin sisteme yüklenmesi ve bu belgenin sistem tarafından Arap ve Latin harflere çevrilmesiyle, Osmanlı Türkçesinin öğrenilmesi ve öğretilmesi hususunda bir ilerleme kaydedileceği öngörülmektedir. Edebiyat alanında Osmanlı Türkçesi ile yazılmış metinlerden oluşan bir veri tabanı hazırlamak özellikle hem akademik anlamda hem de üniversitede işlenen dersler için birincil kaynaklara erişme hususunda metin koleksiyonunun olmasına kolaylık tanıyacaktır. Tablo 1'de proje kapsamında kullanılan koleksiyonlardan bazıları verilmiştir. Bu eserler seçilirken veri çeşitliğinin geniş olması gözetilmiş ve bu yüzden hem edebî hem tarihî belgelerden oluşan koleksiyonlar hazırlanmıştır. Metin ve belgelerdeki farklılığın çeşitli olması, sistemin belge tanıma oranındaki başarısını arttıracağı düşünüldüğünden dolayı koleksiyon hazırlanırken bu hususlar göz önünde bulundurulmuştur. Bu yüzden koleksiyon örneğinde Rika OCR ile işlenen bazı Osmanlı Tükçesi ile yazılmış metinlerin listesi, dosya adları ve belgelerin ihtivası hakkında ipuçları verilmiştir. Koleksiyonlar; tarihî, resmî, askerî, coğrafya, şehircilik, dinî ve kültürel metin ile belgelerden oluşmaktadır. Koleksiyon oluşturulurken dikkat edilen noktalardan bir diğeri ise Osmanlı tarihini ve kültürünü anlamak için geniş bir yelpaze sunan

metinleri bir araya getirilerek yapay zekâ destekli metin işleme sisteminde her disipline dair malumatın bulunacağı kapsamlı ağ oluşturmaktır:

Tablo 1. Koleksiyon Örneği

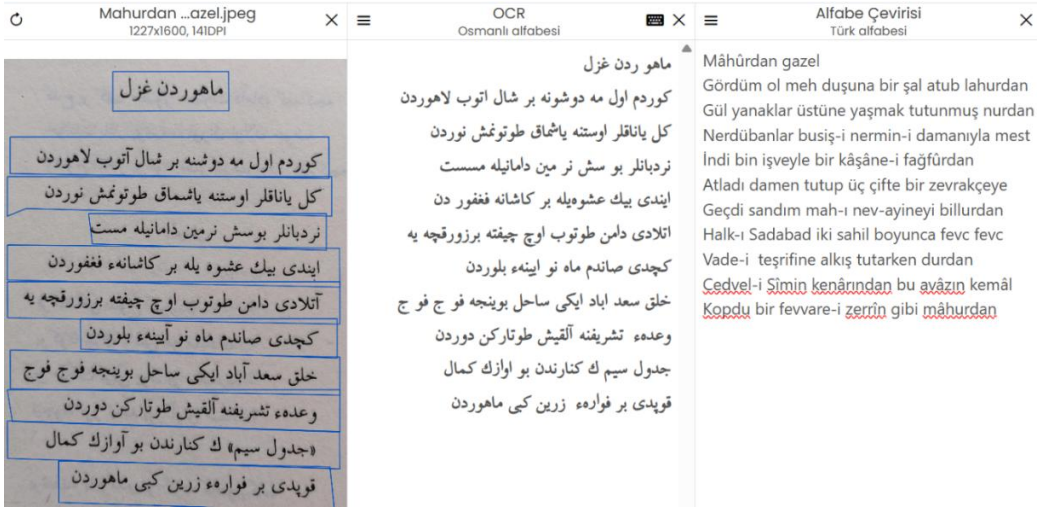
Koleksiyon Adı	Veri Türü	Koleksiyon Adı	Veri Türü
DîvânBâkî	RikaOCR	SarıSaltuk	RikaOCR
DîvânNef'î	RikaOCR	AdanaOlayları	RikaOCR
DîvânNâ'ilî	RikaOCR	1294SavunmaSanayi	RikaOCR
DîvânNâtıkî	RikaOCR	Ankara	RikaOCR
DîvânAvnî	RikaOCR	Azerbaycan	RikaOCR
DîvânNedîm	RikaOCR	TekaudSanduğuBelgeleri	RikaOCR
DîvânŞehrî	RikaOCR	KonyaBeratFermanları	RikaOCR
Süleymânnâme	RikaOCR	MsbSayı-150	RikaOCR
İstiklâlMarşı	RikaOCR	İstanbul'daAfetler	RikaOCR
HakDiniKur'anDili	RikaOCR	TarihSuYolları	RikaOCR
Mecmu'a-287b	RikaOCR	Hazine-iEvrakRehberi	RikaOCR

Çeşitli yüzyıllara ait olan belge koleksiyonları sayesinde yüzyıl içerisindeki üslup değişikliğini anlamak ve dil bilgisi açısından değişimleri metin üzerinde görmek ve incelemek de alana katkı sağlaması adına önem teşkil etmektedir. Koleksiyonların varlığı ayrıca bibliyografya açısından da zengin bir içerik sunacaktır. Osmanlica.com sitesi aracılığıyla eserlerin görüntüleri, okunuşları, Arap ve Latin harflerine aktarımları ve dil çevirilerinin yer aldığı dijital bir kütüphanenin varlığı hedeflenmiştir. Ayrıca belgelerin dijital ortama aktarılması sayesinde aynı yazı stilinde yazılmış farklı metinler görüntüleneceğinden dolayı müstensih farklılıklarından kaynaklanan harflerin ve kelimelerin yazılışları da mukayese edebilme imkânı olacaktır. Kullanıcılar için çeşitli metinleri okuma noktasında kolaylıkların da sağlanması sayesinde arama yapma, indeksleme ve içerik keşfetme süreçleri daha etkin ve verimli bir hale gelmektedir. Örneğin Şekil 4'te Rika hattıyla yazılan farklı türden belgelerin görüntüsü verilmiştir:



Şekil 4. Rik'a yazı stiliyle yazılmış belgelerin görüntüsü

Edebiyat alanında yapılacak araştırmalarda, bir konu hakkında kaynak gösterilirken konu üzerinde yapılan çalışmalara kaynaklık eden ve kronolojik olarak da ilk kaynak niteliği taşıyan Osmanlı Türkçesi ile kaleme alınan eserlerin derlenip tasnif edilerek dijitale aktarılmasıyla, bilgiye ulaşma ve bilgiyi referans etme konularında hızlı bir erişim ağı oluşacaktır. Dil bilgisi başta olmak üzere tematik anlamda da metinlerin analiz edilme kolaylığı, yeni bakış açılarına da kapı aralayacaktır. Özellikle Klasik Türk edebiyatı öğrencileri ve araştırmacıları için hem materyal olarak hem de dijital ortamda veri incelenmesinin yapılması adına oldukça geniş bir dijital alan sunan Osmanlica.com sitesi, manzum ve mensur metinlerin okunması ve tetkik edilmesini sağlamakta ve metinlerin çevriyazısını yapmaktadır. Şekil 5'te Yahya Kemal'e ait Mahurdan Gazel şiirinin görsel metni OCR ile tanınması sağlanarak dijital ortama aktarılmıştır. Görsel metinde yer alan yazı, Osmanlı Türkçesi ve ardından Latin harflerine çevrilerek alfabe çeviri işlemi gerçekleştirilmiştir. Kullanıcılar, bu işlem ile basılı metinler ve el yazması metinler arasındaki farkı kavrayıp metinlerdeki harflerin ve kelimelerin tanınması ile farklı belgeler üzerinde Osmanlı Türkçesini daha rahat öğrenerek araştırma imkânlarından istifade edebilecekler. Ara yüzde bir metnin OCR görüntüsü ve OCR çıktısı ile birlikte alfabe çevrisi görselde verilmiştir:



Şekil 5. Osmanlıca OCR ile Osmanlıca belgenin Türkiye Türkçesine aktarımı

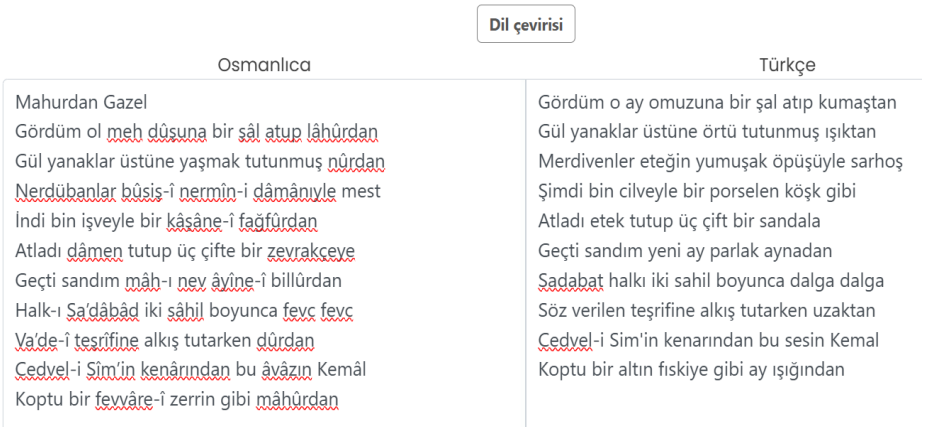
Osmanlı Türkçesinde bir kelimenin Arap harfli yazılışı Türkçeye aktarılırken birden fazla okunuşa sahip olabilmektedir. Bu durumlarda metnin bağlamından hareketle kelimenin doğru okunuşu tahmin edilmektedir. Yapay zekâ aracılığıyla geliştirilen program üzerinde bir belgede yer alan birden fazla okunuş özelliği gösteren kelimeler için "kelime tahmini (n-gram) algoritması" uygulanmaktadır. Bu yöntem ile Arap harfli yazılan bir kelimenin kaç farklı şekilde okunabileceği de görülmektedir. Şekil 6'da görselde site üzerinde geliştirilen çeviri aracını kullanarak örnek metin üzerinden kelime kontrolünün nasıl sağlandığı gösterilmiştir. Görselde anlaşıldığı üzere Arap harfli kelimenin Latin harfli kelime karşılıkları birebir verilmiş ve okuma noktasında doğruluk oranını yükseltmek için bir kelimenin kaç farklı şekilde okunabilirliği sağ tarafta gösterilmiştir. Bu sayede edebiyat alanında metin analizlerinin daha kolay yapılması hedeflenmiştir:



Şekil 6. Osmanlıca-Türkçe alfabe çeviri örneği

Osmanlı Türkçesi bir metnin Türkçe dil çevirisinin yapılması için "kelime grubu tabanlı" bir çeviri sistemi

geliştirilmiştir. Bu sistem "kelime öbeği tabanlı ve mekanik olarak işleyen bir cümleyi öbek öbek soldan sağa doğru tarayarak" (Dölek, 2022: 37) Osmanlı Türkçesinden Türkiye Türkçesine aktaran bir yöntemdir. Bu yöntem sayesinde mensur ve manzum eserlerin dil içi çevirisinin yapılmasıyla birlikte metinlerin anlaşılması daha kolay bir hâle gelecektir. Basit cümlelerde etkin bir şekilde alınan sonuçlar, geliştirilecek programlar ile daha karmaşık cümlelerin de çevirisinde kullanılabileceği kanaatindeyiz. Şekil 7'de görüldüğü üzere geliştirilen çeviri tekniği sayesinde dil içi çevirisi yapılmak istenen Yahya Kemâl'in Mâhurdan Gazel şiiri sisteme yüklenmiş, Arapça ve Farsça kelimelerin günümüz Türkçesi karşılıkları verilerek metin yeniden düzenlenmiştir. Bu çeviri aracı sayesinde edebiyat alanındaki öğrencilerin ve ilgili bireylerin eski Türkçe metinlerindeki muhtevayı ve manayı anlamaları daha kolay hâle geleceği öngörülmektedir:



Şekil 7. Osmanlıca-Türkçe dil çeviri aracı

Türk şiirinde kullanılan aruz ölçüsünü şiirler üzerinde uygulamalı bir şekilde gösteren "Aruz Vezninin Bilgisayarla Çözülmesi" adlı çalışmanın da hizmete sunulduğu Osmanlica.com sitesi üzerinde geliştirilen "Baki" adlı yazılımla, şiirlerde aruz ölçüsünün kolayca bulunması ve Klasik şiirdeki vezin öğretiminde ve veznin anlaşılmasında Yükseköğretim kurumlarına bağlı Türk dili ve edebiyatı dersleri başta olmak üzere lise müfredatında yer alan divan edebiyatı konusunun öğretilmesinde de istifade edilecek bir dijital ortam oluşturulmuştur. Aruz vezninin dijital ortamda çözümlenmesi sayesinde aruz ölçüsünün bulunmasının kolaylaşması, vezin bulmada hata oranlarının azalması, şiirdeki vezin kusurlarının ayrıntılı bir şekilde ortaya konması ve bir divan içindeki çok sayıda şiirin ölçülerinin kısa zamanda bulanabilmesi sağlanmıştır (Atakan, vd., 2007: 6-7).

Eski Türk edebiyatı derslerinde, şiirdeki ahengin bir parçası olan ve metinlerin doğru okunması için ihtiyaç duyulan aruz vezninin öğretilmesinde bilgisayar tabanlı bir eğitim sürecinin olması ile ayrıca aruz veznine dair kuralları gösteren ayrıntılı bir alan oluşturulmuştur. Bu program ayrıca metin tashihi ve tamiri konularında da önemli işlevlere sahiptir. Bir şiirin vezni bulunmadığında veznin bozulduğu yerdeki kelimenin şiirin orijinaline ait olmadığı veya yanlış yazılıp yazılmadığı belli olmaktadır ve vezinlerin taktileri de gösterilmektedir (Atakan, vd., 2007: 21). Aruz veznini bulmak için geliştirilen yazılımda; transkripsiyon harflerinin tanınması, hecelemenin yapılması ve şiirin vezninin doğruluğunu gösteren ayrıntılı görsel web üzerinden kullanıcıya uygun bir şekilde sunulmaktadır. Geliştirilen program kapsamında veri olarak kullanılan şiirler üzerinde "Bâkî" adlı yazılımın şiirde aruz veznini bulma oranının %90'ların üzerinde olduğu tespit edilmiştir (Atakan, vd., 2007: 3). Tablo 2'de Yahya Kemal'in Mâhûrdan Gazel[2] şiiri üzerine vezin bulma işlemi gösterilmiştir ve örnek işlem ilk mısra üzerinden gerçekleştirilmiştir. İlk olarak birinci mısradaki hecelerin açıklık ve kapalılık özellikleri gösterilmiştir. Sonrasında vezinde görülen kusurların hangi hecede olduğu belirtilmiş ve kısaltmalar ile yapılan kusur işlemlerine yer verilmiştir. Kusurlar bulunduktan sonra mısra, vezne uygun şekilde tekrar hecelere ayrılarak vezin ve hece eşleştirilmesi tamamlanmıştır:

Tablo 2. Aruz Vezni Çözüm Aracı

Veri Tabanındaki Şiirlerin Ölçüsü	Gördüm ol meh dûşuna bir şâl atup lâhûrdan Gül yanaklar üstüne yaşmak tutunmuş nûrdan Nerdübanlar bûsiş-i nermîn-i dâmânıyla mest

[2]Şiirin Osmanlı Türkçesi yazımı Coşkun Ak'ın *Osmanlıca* eserinden alınmıştır.

İndi bin işveyle bir kâşâne-i fağfûrdan
Atladı dâmen tutup üç çifte bir zevrakçeye
Halk-ı Sa'dâbâd iki sâhil boyunca fevc fevc
Va'de-i teşrîfine alkış tutarken dûrdan
Cedvel-i Sîm'in kenarından bu âvâzın Kemâl
Koptu bir fevvâre-i zerrîn gibi mâhûrdan

Sanat/Kusur: Ulama (Vasl), Zihaf (Kısaltma), İmale (Uzatma), Med (Ekleme)
Harf Türleri: c: sessiz (consonant), v: sesli (vowel), v̂: uzun sesli (long vowel)
Hece Türleri: – : kapalı/uzun hece, · : açık hece (kısa hece)

Şiirin Ölçüsü **Fâ'ilâtün / Fâ'ilâtün / Fâ'ilâtün / Fâ'ilün**

1. Dize Gördüm ol meh dûşuna bir şâl atup lâhûrdan

Hece No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Önceki Hece	Gör	düm	ol	meh	dû	şu	na	bir	şâl	a	tup	lâ	hûr	dan	
Önceki Yapı	**cvc**	**cvc**	**vc**	**cvc**	**cv̂**	**cv**	**cv**	**cvc**	**cv̂c**	**v**	**cvc**	**cv̂**	**cv̂c**	**cvc**	
Önceki Tür	–	–	–	–	–	–	·	–	–	·	–	–	–	–	
Sanat/Kusur		**U**				**Z**	**İ**								
Sonraki Tür	–	·	–	–	–	·	–	–	–	·	–	–	–	–	
Sonraki Yapı	**cvc**	**cv**	**cvc**	**cvc**	**cv̂**	**cc**	**cv̂**	**cvc**	**ccv̂**	**v**	**cvc**	**cv̂**	**cv̂c**	**cvc**	
Sonraki Hece	Gör	dü	mol	meh	dû	şu	nâ	bir	şâl	a	tup	lâ	hûr	dan	
Dize Ölçüsü	Fâ	i	lâ	tün	fâ	i	lâ	tün	fâ	i	lâ	tün	fâ	lün	
Dize Ölçüsü	Fâ'ilâtün / Fâ'ilâtün / Fâ'ilâtün / Fâ'lün (Şiirin Ölçüsünden farklı!)														

Aruz veznini bulma programı sayesinde aruz kalıplarının anlaşılmasında ve çözümlenmesinde gerek öğrenen gerekse öğreten için teknolojik anlamda ilerlemenin kaydedilmesi, süreci daha eğlenceli ve hızlı bir şekilde değerlendirmeyi beraberinde getirmektedir. Kısaca uygulama, mısra, mısradaki heceler, hecelerin kapalı ve açık, uzun veya kısa durumları, mısraının vezni, mısraının taktili bir şekilde okunuşu ve mısradaki aruza dair kurallar bir bütün şeklinde kullanıcıya sunulmaktadır. Programın geliştirilmesiyle birlikte aruz ölçüsünün yanında hece ölçüsü ile yazılmış şiirlerin incelenmesi, divan şiirindeki aruz noktasında yapılan kusurların istatiksel verilerinin hesaplanması ve tetkik edilmesi ile aruz vezninin tarihsel süreçteki gelişimi hakkında bilgi edinilmesinde yol kat edilebilir ve aruz vezni ile şiir yazmak isteyenler için eğitimin verilmesi de sağlanabilir (Atakan, vd., 2007: 23).

Osmanlica.com sitesinin Türk dili ve edebiyatı alanına katkıları hususunda site üzerinde "Türkçe Sıklık Analizi" yazılımı da bulunmaktadır. Bu program; bir metindeki ses, hece, ek, kelime vb. sıklıkların analizine dair çalışmaların yapılması için geliştirilmiş ve böylelikle hem daha kısa sürede hem de kolay bir şekilde çalışmalar yapılıp veriler elde edileceği öngörülmüştür. "Ergin" adıyla hazırlanan yazılım sayesinde Türkçe öğretimine dair yöntemlerin geliştirilmesi ve dilin tetkik edilmesinde istifade edileceği düşünülmektedir. Çünkü Türk dilinin eğitimi ve öğretiminde birbirinden farklı çalışmalar yapılmakta, çeşitli yöntemler ve teknikler kullanılmaktadır. Elde edilen veriler ise bilgisayar aracılığıyla analiz edilmektedir. Türkçe öğretiminde araştırmacıların üzerinde durdukları çalışmalardan biri "sıklık analizi"dir. Bu çalışmaları önemli kılan ise özellikle dilin öğretilmesinde ve dil bilim araştırmalarında katkı sunacağı görüşüdür. Bu hususta Doğan Aksan, "Okula yeni başlayan, ilkokul ilk sınıflarındaki öğrencilere ya da bir yabancı dili öğrenmeye başlayanlara ilk aşamada öğretilmesi gereken sözcüklerin hangileri olduğu, dilin sözcüklerinin hangi sıraya göre tanıtılması gerektiği saptanırken değişik gereksinimleri karşılayan ana dili ve yabancı dil sözcükleri hazırlanırken dildeki ögelerin sıklıklarının bilinmesi gerekmektedir." (Aksan, 1998: 20) diyerek dillerde sıklık analizinin yapılmasının dil öğretim sürecinde geliştirilecek yeni metotlar ve tavsiyeler için ön koşul olma özelliğine sahip bir araştırma olduğu hakikatine değinmektedir. Bu noktada "Ergin" adlı yazılım, "birden fazla

metin üzerinde çalışabilir; istenirse metinler birleştirilerek de tek bir metin hâlinde analiz edilip sonuçlar bir pencereden görülebilir. Bunun yanında, metinler birleştirilmeden her biri farklı bir metin şeklinde analiz edilip sonuçlar farklı pencerelerde kullanıcıya gösterilebilir." (Oktay, vd., 2007: 568). "Türkçe Metin Sıklığı Çözümleyicisi" sayfasına metin dosyası yüklendiğinde "harf, hece, kelime, N-grams, cümle ve paragraf sıklıkları"nın bulunması için Şekil 9'daki görülen seçenekler verilmiştir:



Şekil 9: Türkçe metin sıklık çözümleme aracı

Metin dosyası sisteme yüklendikten sonra araştırma ve incelemeye uygun seçenekler seçilip çözümlenmesi istenildiği takdirde sıklıkların çözüm tablosu yüzdelik dilimleriyle verilmektedir. Ahmet Haşim'in Müslüman Saati adlı yazısının "Türkçe Metin Sıklığı" programı kullanılarak harf, hece içi ve kelime içi harf sıklığı analizi Tablo 3, Tablo 4 ve Tablo 5'te verilmiştir. Programa yüklenen metin üzerinden sıklık analizini hangi hususta yapılacağı Şekil 9'daki seçenekler arasında bulunmaktadır. Ayrıntılı olması adına metnin harf, hece içi harf ve kelime içi harf sıklığı analizi ve yüzdelik oranları verilmiştir.

Görüldüğü üzere geliştirilen yazılım, Türkçe yazılmış metinlerin dijital ortamda işlenerek metinlerin; ses, hece, ek, kök, kelime sıklık ve oranları ile metinlerin analizinin yapılması imkânı sağlanarak başta Türkçe ve Türk dili ve edebiyatı alanlarında olmak üzere diğer disiplinlerde de kullanılabilecek özellikler göstermektedir. Osmanlica.com sitesini hülasa edecek olursak Osmanlı Türkçesiyle yazılmış bir belgenin Türkiye Türkçesine çevrilme sürecini ve analiz aşamaları Şekil 10'da akış diyagramı görselinde verilmiştir. Osmanlı Türkçesine dair bir belge dijitalleştirilmek istendiğinde öncelikle "Osmanlıca OCR (Optik Karakter Tanıma)" teknolojisi ile belge dijital alana aktarılmaktadır. Burada kullanılan OCR modeli, metnin kalitesini belirlemede oldukça önemlidir. Belgede taranan karakterlerin metne dönüştürülmesiyle birlikte kontrol aşaması devreye girmekte ve metnin alfabe çevirisi analiz edilmektedir. Çevrilen Arap harfli metin daha sonra Türkiye Türkçesine aktarılmaktadır. Burada sistemde yer alan "Türkçe Sıklık Analizi" programı üzerinden metinde geçen harf, kelime, cümle ve paragraf sıklık analizleri de kontrol edilebilmektedir. Ayrıca eğer bir şiir metin belgesi

üzerinde işlem yapılıyorsa istenildiği takdirde "Aruz Vezni" incelemesi de yapılabilmektedir. Metinlerin anlaşılır hâle gelmesi için "Dil Çeviri Aracı" kullanılmakta ve yeniden bir kontrol aşaması süreci başlamaktadır. Metinlerin Türkiye Türkçesine çeviri aşaması tamamlandığında sadeleşmiş metin üzerinde gerektiğinde tekrardan "Türkçe Sıklık Analizi" yapılabilmektedir:

Tablo 3. Harf Sıklığı

Sıra	Harf	Sıklık	%
1	a	558	12.31
2	e	450	9.93
3	i	439	9.68
4	n	330	7.28
5	r	314	6.93
6	l	278	6.13
7	ı	209	4.91
8	t	202	4.46
9	k	171	3.77
10	m	161	3.55
11	d	155	3.42
12	s	149	3.29
13	z	112	2.47
14	u	112	2.47
15	ü	107	2.36
16	b	107	2.36
17	y	104	2.29
18	g	81	1.79
19	ş	79	1.74
20	v	75	1.65
21	h	70	1.54
22	o	63	1.39
23	ğ	45	0.99
24	c	44	0.97
25	ç	42	0.93
26	ö	26	0.57
27	p	25	0.55
28	f	25	0.55
Toplam		45333	100

Tablo 4. Hece İçi Harf Sıklığı

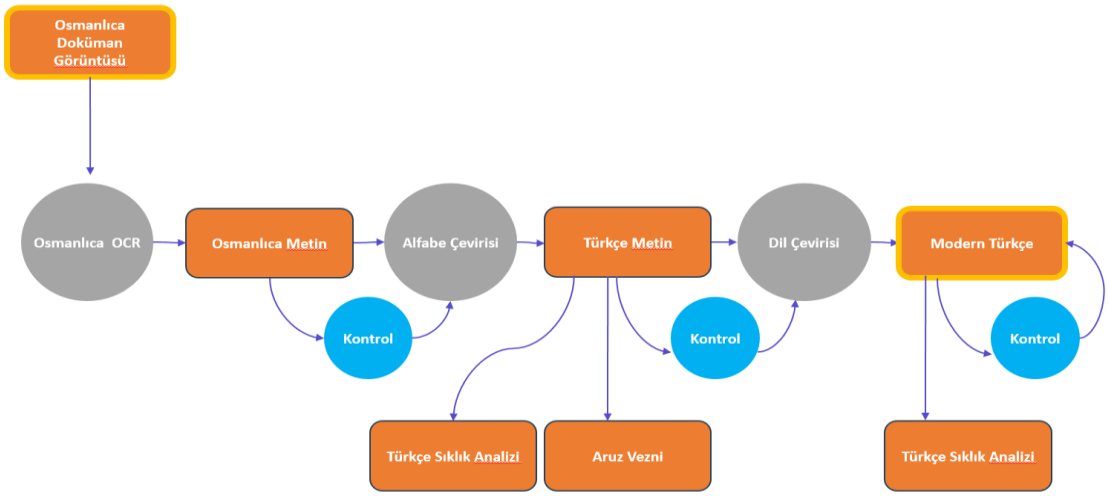
Sıra	Hece	Sıklık	%	1.Harf	2.Harf	3.Harf	4.Harf
1	a	558	12.31	79(%14)	479(%85)	0%	0%
2	c	44	0.97	39(%88)	1(%2)	4(%9)	0%
3	y	104	2.29	87(%83)	3(%2)	14(%13)	0%
4	i	439	9.68	40(%9)	399(%90)	0%	0%
5	p	25	0.55	11(%44)	2(%8)	12(%48)	0%
6	ı	209	4.61	7(%3)	202(%96)	0%	0%
7	d	155	3.42	154(%99)	0%	1(%0)	0%
8	t	202	4.46	121(%59)	36(%17)	36(%17)	9(%4)
9	h	70	1.54	52(%74)	1(%1)	17(%24)	0%

10	m	161	3.55	136(%84)	0%	25(%15)	0%
11	k	171	3.77	99(%57)	11(%6)	58(%33)	3(%1)
12	ş	79	1.74	48(%60)	1(%1)	30(%37)	0%
13	n	330	7.28	97(%29)	19(%5)	214(%64)	0%
14	l	278	6.13	217(%78)	21(%7)	40(%14)	0%
15	e	450	9.93	45(%10)	405(%90)	0%	0%
16	r	314	6.93	160(%50)	8(%2)	146(%46)	0%
17	z	112	2.47	60(%53)	5(%4)	0%	1(%0)
18	u	112	2.47	13(%11)	99(%88)	3(%7)	0%
19	ç	42	0.93	39(%92)	0%	3(%2)	0%
20	ğ	45	0.99	40(%88)	2(%4)	0%	0%
21	b	107	2.36	104(%97)	0%	23(%15)	0%
22	g	81	1.79	81(%100)	0%	0%	0%
23	s	149	3.29	111(%74)	14(%9)	23(%15)	1(%0)
24	o	63	1.39	28(%44)	35(%55)	0%	0%
25	v	75	1.65	66(%88)	5(%6)	4(%5)	0%
26	ö	26	0.57	6(%23)	20(%76)	0%	0%
27	ü	107	2.36	8(%7)	99(%92)	0%	0%
28	f	25	0.55	21(%84)	0%	4(%16)	0%
Toplam		4533	100	1969	1867	683	14

Tablo 5. Kelime İçi Harf Sıklığı

Sıra	Harf	Sıklık	%	1.Harf	2.Harf	3.Harf	4.Harf	5.Harf	6.Harf	7.Harf	8.Harf	9.Harf
1	a	558	12.31	41(%7)	182(%32)	60(%10)	123(%22)	30(%5)	37(%6)	39(%6)	11(%1)	10(%1)
2	c	44	0.97	4(%9)	5(%11)	25(%56)	2(%4)	0%	4(%9)	0%	0%	2(%4)
3	y	104	2.29	38(%36)	5(%4)	32(%30)	1(%0)	14(%13)	5(%4)	3(%2)	2(%1)	1(%0)
4	i	439	9.68	35(%7)	90(%20)	29(%6)	87(%19)	43(%9)	39(%8)	39(%8)	28(%6)	21(%4)
5	p	25	0.55	3(%12)	3(%12)	7(%28)	1(%4)	8(%32)	1(%4)	0%	2(%8)	0%
6	ı	209	4.61	7(%3)	23(%11)	5(%2)	33(%15)	21(%10)	28(%13)	28(%13)	21(%10)	17(%8)
7	d	155	3.42	24(%15)	6(%3)	16(%10)	27(%17)	19(%12)	29(%18)	9(%5)	6(%3)	3(%1)
8	t	202	4.46	25(%12)	13(%6)	41(%20)	50(%24)	28(%13)	19(%9)	11(%5)	9(%4)	3(%1)
9	h	70	1.54	34(%48)	1(%1)	26(%37)	6(%8)	0%	0%	0%	1(%1)	2(%2)
10	m	161	3.55	52(%32)	2(%1)	33(%20)	12(%7)	19(%11)	22(%13)	10(%6)	3(%1)	4(%2)
11	k	171	3.77	42(%24)	18(%10)	39(%22)	13(%7)	30(%17)	13(%7)	7(%4)	1(%0)	3(%1)
12	ş	79	1.74	13(%16)	4(%5)	27(%34)	9(%11)	13(%16)	0%	5(%6)	3(%3)	2(%2)
13	n	330	7.28	11(%3)	23(%6)	65(%19)	32(%9)	50(%15)	34(%10)	28(%8)	26(%7)	24(%7)
14	l	278	6.13	10(%3)	35(%12)	41(%14)	39(%14)	70(%25)	39(%14)	17(%6)	11(%3)	8(%2)
15	e	450	9.93	41(%9)	152(%	18(%4)	59(%13)	45(%10)	57(%12)	25(%5)	20(%4)	7(%1)
16	r	314	6.93	10(%3)	9(%2)	94(%29)	21(%6)	36(%11)	22(%7)	48(%15)	34(%10)	9(%2)
17	z	112	2.47	20(%17)	9(%8)	21(%18)	15(%13)	6(%5)	11(%9)	6(%5)	4(%3)	7(%6)
18	u	112	2.47	13(%11)	25(%22)	9(%8)	27(%24)	16(%14)	10(%8)	5(%4)	4(%3)	3(%2)
19	ç	42	0.93	15(%35)	14(%33)	9(%21)	4(%9)	0%	0%	0%	0%	0%
20	ğ	45	0.99	0%	3(%6)	14(%31)	1(%2)	10(%22)	4(%8)	2(%4)	11(%24)	0%
21	b	107	2.36	75(%70)	1(%0)	16(%14)	6(%5)	6(%5)	1(%0)	1(%0)	0%	0%
22	g	81	1.79	70(%86)	0%	0%	7(%8)	0%	0%	1(%1)	0%	0%
23	s	149	3.29	69(%46)	14(%9)	32(%21)	10(%6)	10(%6)	3(%2)	4(%2)	6(%4)	0%
24	o	63	1.39	28(%44)	27(%42)	0%	1(%1)	4(%6)	2(%3)	0%	0%	1(%1)
25	v	75	1.65	50(%66)	6(%8)	10(%13)	1(%1)	6(%8)	1(%1)	0%	0%	0%
26	ö	26	0.57	6(%23)	18(%69)	0%	0%	0%	0%	0%	0%	0%

27	ü	107	2.36	8(%7)	58(%54)	0%	18(%16)	15(%14)	4(%3)	2(%1)	1(%0)	0%
28	f	25	0.55	15(%60)	0%	3(%12)	1(%4)	4(%16)	1(%4)	0%	1(%4)	0%
Toplam		4533	100	759	746	672	606	503	386	290	205	127



Şekil 10: Osmanlica.com sitesinin akış diyagramı

6. Sonuç

Projenin hedefine ulaşması için birtakım görüntü işleme yöntemleri, derin sinir ağları kullanılmıştır. Geliştirilen sistemin kolay erişilmesi ve pratik bir şekilde kullanılması için interaktif bir çeviri ortamı olan Osmalica.com sitesi tasarlanmıştır. Projenin başarı seviyesini arttırmak için çeşitli yazı stilleri, varak şekilleri ve farklı tip belge ve metinler kullanılmıştır.

Osmanlı Türkçesiyle kaleme alınmış eserlerin OCR işlemi ile dijital ortama aktarılmasının Türk dili ve edebiyatı başta olmak üzere sosyal ve beşeri bilimler alanında birçok faydayı beraberinde getirecektir. Öncelikle el yazmaların ve nadir eserlerin dijital ortamda saklanıp korunması için bu sistemin varlığı önemlidir. Dijital bir ortama aktarılan eserlere erişilme kolaylığını da beraberinde getiren bu çalışma, araştırmacılara ve öğrencilere de rahat bir şekilde analiz yapma imkânı sunmaktadır. Özellikle dilbilimciler için bir metni harf harf, kelime kelime kontrol edebilme imkânı sağlayan bu sistem, dilin işlenmesinin ayrıntısını da gözler önüne sermektedir.

OCR, bir nevi stilometrik tespitlerin yapılmasını da sağlar ve yazarın üslubunu tespit etme, metinleri karşılaştırma, bir konu dâhilinde yapılan çalışmalara kolaylıklar sağlayarak kaynaklık etme bakımından da araştırmacılara ve okurlara katkılar sunmaktadır. Ayrıca, Osmanlı Türkçesi derslerinde eğitim materyali olarak kullanılacağı gibi bu dili öğrenmek isteyenlere kaynak metinler de sunmaktadır. Okullarda öğrenciler ve hocalar tarafından veri setleri oluşturularak Osmanlı Türkçesi derslerinin içeriği bu sayede zenginleştirilebilir. Sistem üzerinden metin çeşitliği arttıkça sosyal bilimler başta olmak üzere tıp, eczacılık, iktisat, hukuk gibi alanlarla ilgilenenler de Osmanlı Türkçesi metinleri üzerinde kolay bir şekilde literatür taraması gerçekleştirebilirler. Osmanlı Türkçesiyle kaleme alınan eserlerin teknolojik gelişmelerle birlikte sundukları verilerin zenginliğini değerlendirmek, kültürel mirasa sahip çıkmak demektir. Ayrıca beşerî ve sosyal bilimlerin mühendislik vb. disiplinlerle ortaklaşa yapacakları çalışmalar, dil, tarih ve kültüre ilişkin mevzularda çok yönlü bir bakış açısına sahip olunmasına ve yeniden yorumlanmasına imkân sağlayacaktır.

Özellikle Eski Türk edebiyatı alanında eğitim gören öğrencilerin Osmanlı Türkçesine vakıf olmaları için dijital

bir platformun geliştiriliyor olması, bu öğrencilere eğitim aşamasında katkı sağlayacağı düşünülmektedir. Özellikle *Osmanlica.com* sitesinde yer alan "Bâkî" isimli yazılımla oluşturulan "Aruz Vezninin Bilgisayarla Çözümlenmesi" ve "Ergin" isimli yazılımla oluşturulan "Türkçe İçin Bir Sıklık Analizi Programı" sayesinde Türk dili ve edebiyatı alanında yapılacak çalışmalarda verimi artıracak zeminin oluşması sağlanmaya çalışılmıştır. Daha verimli ve kullanışlı bir web sayfasının varlığı için çalışmalar zenginleştirilerek devam etmektedir.

Kaynaklar

Ak, C. (2012). *Osmanlıca*. Bursa: Alfa Akademi Basım Yayım Dağıtım.

Aksan, D. (1998). *Her Yönüyle Dil Ana Çizgileriyle Dilbilim*. Ankara: Türk Dil Kurumu Yayınları.

Çetin, N. M. (1989). Aklâm-ı Sitte. *Türkiye Diyanet Vakfı İslam Ansiklopedisi* (C. 16, s. 276-280). Ankara: Türkiye Diyanet Vakfı Yayınları.

Derman, M. U. (1997). Hat. *Türkiye Diyanet Vakfı İslam Ansiklopedisi* (C. 16, s. 427-437). Ankara: Türkiye Diyanet Vakfı Yayınları.

Dölek, İ. (2022). *Osmanlıcadan Modern Türkçeye Uçtan Uca Aktarım Sistemi* (Doktora Tezi, İstanbul Üniversitesi).

Dölek, İ., ve Kurt, A. (2022). A deep learning model for Ottoman OCR. *Concurrency and Computation: Practice and Experience 34.20*, e6937.

Dölek, İ., ve Kurt, A. (2021). Ottoman OCR: Printed naskh font. *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE.

Dölek, İ., ve Kurt, A. (2022). Osmanlıcadan Türkçeye Uçtan Uca Aktarım. *Journal of Smart Systems Research 3*(1), 1-10.

Ece, S. (2015). *Klasik Türk Edebiyatı Araştırma Yöntemleri (I-II)*, Erzurum: Eser Basım Yayın Dağıtım Matbaacılık.

Haşim, A. (2011). *Gurebâhâne-i Laklakan*. (haz. Nâzım Hikmet Polat), İstanbul: Yapı Kredi Yayınları.

Kurt, A., ve Kara, Mehmet. (2007). Aruz Vezninin Bilgisayarla Çözümlenmesi (Proje No. 106K270). *TÜBİTAK, Sosyal ve Beşeri Bilimler Araştırma Grubu (SOBAG)*.

Oktay, M., Kurt, A., ve Kara, M., (2007). Türkçe İçin Bir Sıklık Analizi Programı. *38. Uluslararası Asya ve Kuzey Afrika Çalışmaları Kongresi (ICANAS)*, Ankara, 566-576.

Özkan, M. (2007). Osmanlı Türkçesi. *Türkiye Diyanet Vakfı İslam Ansiklopedisi* (C. 33, s. 483-485). Ankara: Türkiye Diyanet Vakfı Yayınları.

Tulum, M. (2011). *Osmanlı Türkçesine Giriş*. Eskişehir: Anadolu Üniversitesi Yayınları.

Türkiye Büyük Millet Meclisi. (2024). *TBMM Ana Sayfa*. Erişim Adresi: https://www5.tbmm.gov.tr// (03.09.2024).

Etik, Beyan ve Açıklamalar

1. Etik Kurul izni ile ilgili;
☑**.** Bu çalışmanın yazar/yazarları, Etik Kurul İznine gerek olmadığını beyan etmektedir.
2. Bu çalışmanın yazar/yazarları, araştırma ve yayın etiği ilkelerine uyduklarını kabul etmektedir.
3. Bu çalışmanın yazar/yazarları kullanmış oldukları resim, şekil, fotoğraf ve benzeri belgelerin kullanımında tüm sorumlulukları kabul etmektedir.
4. Bu çalışmanın benzerlik raporu bulunmaktadır.