ÇEVİRMENLER İÇİN OCR VE DİZGİ REHBERİ



İSTANBUL 17/05/2025

Hazırlayan

Versiyon	Hazırlayan	Açıklama	Tarih
V1	Kemal ISTIL	Genel içeriğın hazırlanması ve düzenlenmesi	20/04/2025

Yardımcı Yazarlar

Versiyon	Yazar	Açıklama	Tarih
V1	Zeynep BEYDEŞ	Belge tasarımı ve içerikleri bilgileri	
	Yeşim	Belge fotoğrafı çekimi bölümü	
	Damla ÖR	OCR alan türleri araştırması	
	Uğur DEMİRKAN	Fiziksel belge güvenliği öğeleri araştırması	
	Nisanur DARDAN	CAT Tool segmentasyonu	

Doküman Geçmişi

Onay Geçmişi

Onaylayan	Onay	Açıklama	Tarih
Burak OSLU	~	Kontrol edildi, onaylandı.	17.05.2025
Kemal ISTIL	~	Kontrol edildi, onaylandı.	17.05.2025

Doküman Revizyon Geçmişi

Revizyon	Versiyon	Açıklama	Tarih
Kemal ISTIL	V1		17.05.2025

İÇİNDEKİLER

1.	GİRİŞ		1
2.	OCR Ö	NCESİ HAZIRLIK	2
	2.1. Bel	ge Yapısı ve Formatları	2
	2.2. Bel	gelerde Bulunan Temel Öğeler	2
	2.3. Bel	genin Güvenlik Öğelerinin Kontrolü	3
	2.3.1.	Fiziksel Güvenlik Önlemleri	3
	2.3.2.	Kimyasal ve Görsel Güvenlik Önlemleri	5
	2.3.3.	Elektronik/Dijital Güvenlik Önlemleri	6
	2.4. Bel	ge ve Görüntü Kalitesi Kontrolü	8
	2.4.1.	Tarama İşlemi	8
	2.4.2.	Fotoğraf Çekimi	9
	2.4.3.	Mobil Tarama Uygulamalarını Kullanma	10
	2.4.4.	Dosya Formatı, Saklama ve Gönderme	11
З.	OCR S	IRASINDA DİKKAT EDİLMESİ GEREKENLER	12
	3.1. OR	C Yazılımı Seçiminde Dikkat Edilmesi Gerekenler	12
	3.1.1.	Farklı OCR Araçları ve Özellikleri	13
	3.2. Gör	rüntü Düzenleme (Image Editing)	13
	3.2.1.	Önerilen Önisleme (Recommended Preprocessing)	14
	3.2.2.	Eğriliği Giderme (Deskew)	14
	3.2.3.	Metin Satırlarını Düzleştirme (Straighten Text Lines)	14
	3.2.4.	Fotoğraf Düzeltme (Photo Correction)	14
	3.2.5.	Yamuk Bozulmalarını Düzeltme (Correct Trapezoid Distortion)	14
	3.2.6.	Döndürme ve Çevirme (Rotate & Flip)	14
	3.2.7.	Bölme (Split)	14
	3.2.8.	Кırpma (Сгор)	14
	3.2.9.	Ters Çevirme (Invert)	14
	3.2.10.	Çözünürlük (Resolution)	15
	3.2.11.	Parlaklık ve Karşıtlık (Brightness & Contrast)	15
	3.2.12.	Seviyeler (Levels)	15
	3.2.13.	Silgi (Eraser)	15
	3.2.14.	Renk Işaretlerini Kaldırma (Remove Color Marks)	15
	3.3. Ala	n Türlerinin Seçimi	15
	3.3.1.	Metin Alanı (Text Area)	15
	3.3.2.	Tablo Alanı (Table Area)	16
	3.3.3.	Görsel Alan (Picture Area)	16
	3.3.4.	Arka Plan Resmi (Background Picture)	16
	3.3.5.	Zorlu Alanlarda Manuel Kontrol	16
4.	OCR S	ONRASI WORD DİZGİ ÇALIŞMASI	17
	4.1. Gör	ünmez Karakterler ve Özel İşaretler	17
	4.1.1.	Paragraf ve Satır Sonu İşaretleri	17

	Sekme, Boşluk ve Hızalama Karakterleri	17
4.1.3.	Sütun, Sayfa ve Son Not İşaretleri	18
4.1.4.	Metin İçi ve Gelişmiş Arama Karakterleri	18
4.1.5.	Grafik, Dipnot ve İçerik Temsilleri	18
4.1.6.	CAT Tool Segmentasyonuna Etkisi	18
4.2. Ta	blolama	18
4.2.1.	Hücre İçeriğini Kontrol Etme	19
4.2.2.	Hücre İçi Düzenleme ve Kısayollar	19
4.2.3.	Tablo Düzenleme Taktikleri	19
4.3. Ce	tvel Kullanımı ve Sekmelerin Yerleştirilmesi	20
4.3.1.	Cetvelin Görünür Hale Getirilmesi	20
4.3.2.	Girintiler	21
4.3.3.	Sekme Durakları	21
4.4. Sü	tunlar	22
4.4.1.	OCR Sonrası Sütun Hataları	23
4.5. Me	tin Kutuları ve Çerçeveler	23
4.5.1.	Teknik Öneriler	24
4.6. Ya	zı Tipi, Boyut ve Standartlaştırma	24
4.6.1.	CAT Araçlarında Segment İçine Tag Olarak Yansıyan Biçim Bozuklukları_	25
4.6.2.	Yapılması Gerekenler	25
4.7. Sa	yfa Yapısı ve Düzeni	26
4.7. Saj 4.8. Re	yfa Yapısı ve Düzeni simler ve Şekiller	26 27
4.7. Sa 4.8. Re 4.8.1.	y fa Yapısı ve Düzeni simler ve Şekiller Görsel Öğelerin Kullanımı: Nerede Yerleştirilir, Nerede Yerleştirilmez?	26 27 27
4.7. Sa 4.8. Re 4.8.1. 4.8.2.	y fa Yapısı ve Düzeni	26 27 27 28
4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3.	y fa Yapısı ve Düzeni	26 27 27 28 29
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 	y fa Yapısı ve Düzeni	26 27 27 28 29 31
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 	yfa Yapısı ve Düzeni	26 27 28 29 31 31
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 5.1. Big 5.2.1. 5.2.2. 5.2.3. 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33 33
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 5.2.3. 5.2.4. 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33 33 34
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33 33 34 35
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 5.1. Big 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 5.3. HT 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33 33 33 34 35 35
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 5.3. HT 5.4. CA 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33 33 33 34 35 35 35
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 5.3. HT 5.4. CA 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 31 33 33 33 33 33 33 33 33 33 35 35 36
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 5.3. HT 5.4. CA 5.4.1. 6. ÇEVİ 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 333 3333 3333 333333
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 5.3. HT 5.4. CA 5.4.1. 6. ÇEVİL 6.1. Be 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 333 33 33 33 333 333 333 333 333 333 333 333 333 333 333 333 333 333 333 3333 3333 333333
 4.7. Say 4.8. Re 4.8.1. 4.8.2. 4.8.3. 5. CAT 2 5.1. Big 5.2. Me 5.2.1. 5.2.2. 5.2.3. 5.2.4. 5.2.5. 5.3. HT 5.4. CA 5.4.1. 6. ÇEVÍA 6.1. Be 6.1.1. 	yfa Yapısı ve Düzeni	26 27 28 29 31 31 32 33 33 33 33 33 33 33 33 33 33 33 35 35 36 37 37 37

6.1.3.	Biçimlendirme ve Stil Tutarlılığı	37
6.2. Me	tin Hatalarının Kontrolü	37
6.2.1.	OCR Kaynaklı Hatalardan Dolayı Oluşan Çeviri Hataları	37
6.2.2.	Yazım ve Dilbilgisi Denetimi	38
6.3. CA	T Sonrası Kontrol	38
6.3.1.	Etiket Hataları ve Sıralama Sorunları	38
6.3.2.	Kaymalar ve Format Bozulmaları	38
6.3.3.	Dil Kodları ve Biçimlendirme Uyumsuzlukları	38
6.3.4.	Tablo ve Alan Bozulmaları	39

1. GİRİŞ

OCR (Optik Karakter Tanıma) fiziksel belgelerin dijital metinlere dönüştürülmesini sağlayan yazılımların kullandığı işlemin genel adıdır. Amaç, geleneksel basılı belgelerin dijital metin formatlarına dönüştürülmesi ve düzenlenmesidir. Bu teknolojiler büyük miktarda veriyi hızlı bir şekilde dijital ortama aktarmayı ve düzenlemeyi mümkün kılarak belge yönetimi ve dijital arşivleme işlemlerini çok daha verimli hale getirir. Ancak OCR ve dizgi işlemleri yalnızca teknolojiye dayalı bir süreç değil, aynı zamanda dikkat, hassasiyet ve titizlik gerektiren bir iş akışıdır. Bu yüzden her aşama doğru bir şekilde uygulanmalı ve her adımda kalitenin korunmasına özen gösterilmelidir.

Kullanılan yazılımın algoritmasına göre karşılaşılan zorluklar metinlerin doğruluğunu ve okunabilirliğini doğrudan etkileyebilir. OCR öncesinde alınabilecek basit önlemler daha doğru sonuçlara ulaşmada süreci hızlandıracaktır. Belgelerin doğru bir şekilde taranması, uygun dosya formatlarının seçilmesi ve metnin açık bir şekilde tanınabilir olması için gerekli ortamın hazırlanması önemlidir. Hiçbir OCR yazılımı tamamen doğru olarak aktaracağına dair herhangi bir garanti vermemektedir. Bu yüzden OCR işlemi sırasında kullanılan yazılımın çıktıyı en az hatayla verebilmesi için dikkat edilmesi gereken çeşitli unsurlar vardır. Bu unsurlar belgenin doğru bir şekilde dijital ortamda tanınması ve veri kaybının önlenmesi için kritik rol oynar.

OCR işlemi tamamlandıktan sonra dizgi aşaması başlar. Birçok metin dosya formatı bulunsa da basılı kaynaklardan yapılan çeviri işlerinde en yaygın kullanılan yazılım MS Word formatıdır. OCR sürecinde oluşan hataların düzeltilmesi metnin düzgün bir şekilde düzenlenmesi ve okunabilirliğin artırılması için önemlidir. OCR'nin neden olduğu yanlış okuma hatalarının veya eksik harflerin düzeltilmesi belge çevirilerinde, özellikle çok hedef dilli çeviri projelerinde, daha hızlı aksiyon alınmasını sağlar. Bu şekilde de doğru hazırlanmış bir dizgi hem çevirmenin dosyayı yeniden hazırlaması sürecinde vakit kaybını önler hem de CAT (bilgisayar destekli çeviri aracı) yazılımlarına aktarımı sırasında segmentasyonu daha doğru hale getirerek eşleşme oranlarını arttırır. Bu sayede, özellikle çok dilli projelerde, çevirmenlerin verimli bir şekilde çalışabilmesine yardımcı olur. Burada amaç çeviri sürecinde yaşanacak dil hatalarını minimize etmek ve çeviri sürecini hızlandırmaktır.

Doğru hazırlanmış bir kaynak dosya ile ilerleyen çeviri projelerinde tüm işlemler tamamlandığında son kontrol ve kalite güvencesi süreci daha hızlı ilerler. En başta hatalı oluşturulmuş dosyaların kalite kontrolü de normalden daha uzun sürer ve hataların tespit edilip düzeltilmesi sürecinde de ayrıca vakit ve iş gücü kaybına yol açar. Bu rehber ile OCR ve dizgi süreçlerinin her bir aşamasını detaylı bir şekilde ele alarak doğru teknikler ve en iyi uygulamalarla hem çevirmenlere dizgi işlemleri üzerinden ek gelir kaynağı oluşturmayı hem de çeviri süreçlerini daha hale getirmeyi hedeflemekteyiz. Her bir adımın önemini ve nasıl etkili bir şekilde gerçekleştirilebileceğini anlamak bu süreci başarıyla tamamlamanızı sağlayacaktır.

2. OCR ÖNCESİ HAZIRLIK

2.1. Belge Yapısı ve Formatları

Dikey Belgeler

Dikey format metin ağırlıklı belgelerde tercih edilir. Belge türlerinden bazıları:

- Resmi yazılar (dilekçeler, mahkeme kararları, raporlar, sözleşmeler vb.)
- Tek sütunlu sertifikalar ve katılım belgeleri
- Transkriptler
- Akademik yazılar
- Ticari yazılar (Ticari Sicil Gazetesi, oda faaliyet belgeleri, yıllık raporlar, beyanlar vb.)
- Formlar vb.

Yatay Belgeler

Yatay format ise görsel ağırlıklı veya çok sütunlu belgelerde tercih edilir. Örnek olarak:

- Özel baskı karton belgeler (Diplomalar, meslek eğitim belgeleri, başarı sertifikaları vb.)
- Tablo ağırlıklı, çok sütunlu yazılar (Aile nüfus kayıtları, maaş bordroları, analiz tabloları, vergi levhası vb.)
- Kağıt tasarrufu için kitapçık şeklinde hazırlanan evraklar

Özel Nitelikli Belgeler ve Kimlikler

Uluslararası kabul görmüş standartlara göre hazırlanan pasaport, sürücü ehliyeti, çipli kimlik, mesleki kimlikler gibi özel nitelikli kişiyi tanımlayıcı belgeler.

Öğe	Açıklama	Genellikle Konumu
Kurumsal Logo	Kurumun logosu veya amblemi	Üst sol, orta veya sağ
Belgeyi Veren Kurum/Kuruluş Anteti	Belgeyi hazırlayan ve onaylayan kurum/kuruluşa ait bilgiler	Üst veya alt bilgi alanı
Başlık	Belgenin türünü belirtir (örn. Sertifika, Diploma)	Kurum/Kuruluş adı altında sol veya ortada
Alt Başlık	Belgenin konusunu veya kapsamını açıklar	Başlığın hemen altında veya sol çaprazında
Alıcı/Belge Sahibi Bilgileri	Belge sahibinin kişisel bilgileri (adı, öğrenci numarası vb.	Orta bölümde belirgin
Metin Bölümü	Açıklama, ödül nedeni, belgeyi veren kurum gibi bilgiler	Kişisel bilgilerden sonra ortalanmış veya sola dayalı
İmza Alanı	Yetkililerin imza atacağı alan	Hiyerarşik artan şekilde alt sağa doğru veya eşit dağılmış şekilde

2.2. Belgelerde Bulunan Temel Öğeler

Kaşe / Mühür	Resmi damga	İmza alanına veya altına
Tarih	Belgenin düzenlenme tarihi	Üst sağ veya alt sol

2.3. Belgenin Güvenlik Öğelerinin Kontrolü

Belgelerin güvenlik öğeleri belgenin dolaşımı sırasında resmiyetini garanti altına almak ve evrakta sahteciliği önlemek amacıyla kullanılır. Tek bir belge türünde aynı anda birden çok güvenlik önlemi bulunabilir ve resmi tasdik istenen, özellikle noter onayı gerektiren çeviri işlemlerinde yasal sorumluluk almadan önce ilgilerle birlikte kontrol edilmesi gerekir. Belgenin taramasının/fotokopisinin alınması sırasında ise bu güvenlik öğelerinin zarar görmemesine özellikle dikkat edilmelidir. Belge güvenliği konusunda yeterli bilgi sahibi olmayan kişiler belgelerin taraması veya fotokopisinin çekilmesi istediğinde belgenin güvenlik öğelerine zarar verebilir. Bu durumda belgenin aslı resmi niteliğini kaybedebilir ve tüm resmi işlem süreci riske girer. Örnek olarak sayfaları tek tek taramayla uğraşmak yerine tarayıcıların otomatik beslemelerini kullanmak için yaprakları ayırırsanız zımba alanında oluşan tahribat belgenin ara sayfalarının değiştirildiği şüphesi yaratabilir. Bu tarz durumlarda güvenlik öğeleri zarar gördüğü için belge artık resmiyetini kaybetmiş sayılır ve teslim edileceği ilgili kurumlar tarafından geçerli olarak kabul edilmeyebilir.

2.3.1. Fiziksel Güvenlik Önlemleri

Soğuk Damga (Embossed Seal)

Belgeleri ayırt etmeye yarayan kabartmalı damgalardır. Özellikle takım halindeki çok yapraklı belgelerde belgelerin aynı takıma ait olup olmadığını incelemek için kullanılır. Eğer yapraklarda soğuk damgalar üst üste eşleşmiyorsa veya yaprakların herhangi birinde kabartma bulunmuyorsa belgede sahtecilik yapıldığı şeklinde algılanır.

- Noter tasdikli belgelerde, diploma ve devlet kurumu merkezli resmi evraklarda kullanılır.
- Kağıdın üzerine basılarak kabartma şeklinde bir iz bırakır. Mürekkep kullanılmaz.
- Çoğunlukla kurum dışında dolaşıma girecek olan şirket belgeleri ve resmi sertifikalar genellikle soğuk damga ile mühürlenir.



Kurdele ve Mum Mührü (Ribbon & Wax Seal)

Belgelerin sayfalarını bir arada tutmak ve çok yapraklı evrakların sayfalarının ayrılmamasını garantilemek için kullanılır.

- Bazı ülkelerde özellikle noter belgelerinde ve mahkeme evraklarında kullanılır.
- Belgeler kırmızı veya mavi kurdeleyle birleştirilir, ardından üzerlerine balmumu veya plastik mühür basılır.
- Günümüzde dayanıklı kağıt malzemelerden üretilmiş özel mühür çıkartmalar da kullanılabilir.
- Genellikle Almanya'dan gelen mahkeme kararları ve noter belgelerinde; Rusya ve Azerbaycan gibi ülkelerden gelen resmi belgelerde yaygın olarak kullanılır.
- Çıkartma veya balmumu mührüne ve kurdeleye zarar vermeden sahtecilik amacıyla evrak sayfalarının değiştirilmesi çok zor olduğundan yaygın olarak kullanılan güvenlik önlemlerinden bir tanesidir.

Hologram (Security Hologram)

Talep eden kurum veya kuruluşlara özel olarak hazırlanan, ışık altında farklı açılarda görülebilen güvenlik öğeleridir. Çeşitli şekillerde belge üzerinde yer alabilir. Genellikle belgeye özel numara vermek ve belgede kalıcı iz bırakmak amacıyla kullanılır.

- Kimlik belgeleri, pasaportlar, sertifikalar, diplomalar ve yüksek güvenlik gerektiren evraklarda bulunur.
- Çıkartma olarak eklenebileceği gibi doğrudan özel tasarlanmış yazıcılarla da basılabilir.
- Çizilme, sıcaklık ve sıyrılmaya karşı dikkat edilmelidir. Mat, solmuş veya soyulmaya başlamış hologramlı belgelerin yenilenmesi talep edilmelidir veya bu durum yetkili makamlara iletilecek şekilde hareket edilmelidir.
- Tarama veya fotoğraf çekimi sırasında büyük oranda yansıma yapar, görsellerde içerik belli olmayabilir, kopyalanması zordur.
- Sökülmesi durumunda belgede leke veya desenli parça bırakacak şekilde üretilir.





Güvenlik İpliği (Security Thread)

Genellikle banknotlar ve ticari değeri olan senetlerde kullanılır. Maliyetli olduğu için basit resmi belgelerde çok yaygın kullanılmasa da kurum veya kuruluşun isteğine göre özel kağıtlara matbaalarca basılabilir. Işığa tutulduğunda görünen veya özel cihazlarla doğrulanan ince ipliklerdir.

- Özel baskı gerektirdiği için yüksek güvenlikli standart belgelerde bulunur.
- Özellikle paralarda yaygın olarak kullanılır.
- Belgeyi veren kurum veya kuruluşun bilgileri şeritte yer alır.

2.3.2. Kimyasal ve Görsel Güvenlik Önlemleri

Mikrobaskı (Microprinting)

Standart yazıcılarla oluşturulamayacak seviyede küçük, insan gözüyle okunması zor olan ancak büyüteçle görülebilen küçük harf veya desenlerin kullanıldığı bir güvenlik önlemidir. Belgelerde, banknotlarda ve sertifikalarda sahteciliği önlemek için kullanılır. Sahte belgelerde bu tür detaylar genellikle bulanık veya düzensiz olur.

- Resmi belgelerde, banknotlarda, pasaportlarda ve yüksek güvenlik gerektiren belgelerde bulunur.
- Genellikle kurumlar arası gizli öğelerdir, devlet kurumları dışında detaylar kurum dışına paylaşılmaz.
- Belgelerin belirli alanlarına entegre edilerek, normal yazı gibi görünse de büyütüldüğünde özel harf veya rakam kombinasyonları içerebilir.

UV Baskı (Ultraviolet Printing)

UV baskı yalnızca ultraviyole ışık altında görülebilen özel mürekkeplerle ve baskı cihazları ile yapılan, yüksek güvenlik gerektiren belgeler için kullanılan bir baskı türüdür. Özellikle kimliklerde sahteciliği önlemek için sıkça kullanılır.

- Kimlik kartları, pasaportlar, diplomalar ve yüksek güvenlikli belgelerde bulunur.
- Belgenin üzerine gözle görülmeyen ancak UV ışık altında belirgin hale gelen yazılar veya desenler basılır.
- Hassas olduğu için uzun süreli olarak güneş altı, nemli, sıcaklığı yüksek ortamlara maruz bırakmak mürekkebin silinmesine yol açabilir.
- Tarama/fotokopi makinesinin otomatik besleyicisi kullanılmamalı ve tarama alanının çok ısınmamasına veya ışığın zarar verici seviyede olmamasına dikkat edilmelidir.

Filigran (Watermark)

Filigran, kâğıt üretimi sırasında içine yerleştirilen, yüksek oranda transparan olarak basılmış özel desenlerdir. Banknotlar, resmi belgeler ve değerli kâğıtlar için yaygın bir güvenlik unsurudur.



- Banknotlarda, pasaportlarda, noter onaylı belgelerde ve resmi sertifikalarda bulunur.
- Genellikle belgenin belirli bir alanında yer alabilir ve belgenin arkasından ışığa tutulduğunda görülebilir.
- Özel sektörde de genellikle akreditasyon veya sertifika kuruluşlarının verdiği belgelerde kullanılır.

Kimyasal Tepkimeye Giren Kâğıt (Chemical Reactive Paper)

Üzerine belirli kimyasal maddeler temas ettiğinde renk değiştiren veya tahrip olan özel belgeler için kullanılır. Bu sayede yetkisiz değişiklikler tespit edilebilir.

- Genellikle yüksek öncelikli resmi belgeler, çekler ve reçetelerde bulunur.
- Mürekkep silme veya değiştirme girişimlerinde belgenin yapısında renk değişimi meydana gelir, bu da sahteciliği ortaya çıkarır.

Lazer Delik (Laser Perforation)

Belge üzerine lazerle küçük delikler açarak numaralar veya desenler oluşturma tekniğidir. Işığa tutulduğunda net bir şekilde görülebilir.

- Pasaportlar, kimlik kartları ve yüksek güvenlik gerektiren belgelerde kullanılır.
- Düzgün bir şekilde taklit edilmesi zordur.

2.3.3. Elektronik/Dijital Güvenlik Önlemleri

Dijital güvenlik önlemleri, belgelerin elektronik ortamda takibi, doğrulanması ve sahteciliğe karşı korunması amacıyla kullanılan teknik çözümlerden oluşur. PDF veya taranmış evraklarda bu güvenlik unsurlarının doğrulandıktan sonra doğrulanmış versiyonların ve çeviriye kaynak olarak kullanılması gerekir. Basılı olarak verilen ancak elektronik olarak imzalandığı belirtilen belgelerde tarama ve OCR yerine doğrudan kaynaktan indirme yoluyla ilerlemek çeviri sürecini daha güvenli ve hızlı hale getirir. Örnek olarak Türkiye'deki kesinleşmiş mahkeme kararlarının orijinal dijital versiyonları UYAP Evrak Doğrulama arayüzü üzerinden indirilebilir.

QR Kod, Barkod veya Link ile Doğrulama (Verification via QR Code, Barcode, or Link)

Belge üzerinde yer alan QR, barkod veya linkler dolaşımdaki belgenin doğruluğunu bir veritabanı aracılığıyla kontrol etmeye yarar. Bu öğeler web sayfasına yönlendirerek belgenin verildiği kurum/kuruluş tarafından teslim edilmiş asıl kaynağını gösterir veya doğrulama ekranı sağlar. Örnek olarak Türkiye'deki e-Devlet belgeleri gösterilebilir.

- Dolaşımdaki bu tarz belgelerde dijital doğrulama sürecinin nasıl yapılacağı ile ilgili mutlaka bir yönlendirme bulunur ve doğrulama arayüzüne tüm dünyadan erişim sağlanabilir olmalıdır.
- Islak imzalı olsa dahi asıl belge dijital olarak verildiğinden mutlaka doğrulanmış belge kaynak olarak alınmalıdır. Islak imzalı belge ile dijital kopya hukuki bir uyuşmazlığa konu olduğunda hukuki öncelik doğrulama şartlarını karşılayan dijital belgededir.
- Çeviriye başlanmadan önce QR kod, barkod veya link üzerinden doğrudan kaynaktan belge indirilmelidir ve güncelliği kontrol edilmelidir.

- Doğrulanamayan belgeler kullanılmamalıdır veya bu durum ilgili kurum/kuruluşlara bildirilecek şekilde önlemler alınmalıdır.
- Yönlendirilen web sitesinin alan adı mutlaka sadece mevzuatla talep edilebilen güvenilir bir uzantıya (*.gov/gov.tr, *.edu/edu.tr, *org/org.tr vb.) sahip olmalıdır ya da alan adının Whois sorgusu belgeyi veren kurum/kuruluş bilgileriyle örtüşmelidir.
- Doğrulama yapılan web sitesi sahibi kurum/kuruluşla ilgili ülkenin yasalarına göre yasal varlığını kanıtlayabildikleri ve kontrol edilebilen vergi numarası, ticaret sicil bilgileri vb. bilgiler bulunmalıdır.

Belge Kimlik Bilgileri: UUID, Referans No, Blockchain Hash Kodu vb. (Document Identifiers: UUID, Reference Number, Blockchain Hash, etc.)

Resmi belgelerde her belgeye özel bir referans numarası, doğrulama kodu, UUID (evrensel benzersiz tanımlayıcı) veya blokzincir hash kodu bulunur. Bu tarz belge kimlik bilgileri sayesinde belgelerin tekilliği garanti altına alınmış olur ve kurumun sisteminde aslı sorgulanabilir.

- Mümkün olan her durumda bu numaranın kurum/kuruluşun sisteminde kontrol edilmesi gereklidir.
- Mümkün olmayan durumlarda belge geçerliliğinin mutlaka en az 1 fiziksel güvenlik öğesi (belge başka ülkede kullanılacaksa ayrıca Apostil veya Tasdik Şerhi) bulunmalıdır.
- Kodlar, sayılar ve belge kimliğini belirleyici bütün bilgiler olduğu gibi aktarılmalıdır, kod üzerindeki anlamlı yazılar çevrilmemelidir. Genellikle makine çevirisinin dikkatsiz şekilde kullanımında en çok yapılan hatalardandır.
- Hiçbir zaman hiçbir durumda aynı bilgilere sahip aynı belgenin 2 farklı versiyonu dolaşımda bulunmaz. Veriliş günü veya zamanı farklı olsa dahi yeni bir belge kimlik numarası ile tekrar oluşturulur. Aksi halde belgenin sahte olabileceği şeklinde önlemler alınmalıdır.

Dijital Damga, Elektronik İmza/Sertifika ve Zaman Damgası (Digital Seal, Electronic Signature/Certificate, and Timestamp)

e-İmza/Sertifika, dijital belgelerin içeriğinin değiştirilmediğini ve belirli bir kişi veya kurum tarafından onaylandığını gösteren kriptografik güvenlik altyapısıyla oluşturulmuş yüksek güvenlikli dijital imzadır. Zaman damgası ise bu imzanın belirli bir tarih ve saatte atılmasıyla dijital dosyanın yasal olarak o anda oluşturulduğunu teknik olarak garanti eder. e-İmzalı belgeler genellikle .pdf, .xml, .eyp veya .p7m gibi formatlardadır.

- Belgenin e-imzalı olduğuna dair bir açıklama bulunuyorsa imzanın gerçekliği ve belge içeriğinin değiştirilip değiştirilmediği mutlaka kontrol edilmelidir.
- Uluslararası alanda PAdES (PDF Advanced Electronic Signatures), XAdES (XML Advanced Electronic Signatures) ve CAdES (CMS Advanced Electronic Signatures) gibi standartlar kullanılır.
- PDF belgelerinde e-imzalar Adobe Reader, benzeri programlar veya ilgili mevzuata göre hazırlanmış özel yazılımlarla "İmza Paneli" üzerinden kontrol edilebilir. Geçerli bir e-imzada sertifika bilgisi, imzalayan kişi/kurum adı, sertifika sağlayıcısı ve imza zamanı yer alır.

- e-İmza doğrulama sürecinde, nitelikli elektronik sertifikalar kullanılır. Bu sertifikalar, Avrupa'da eIDAS, ABD'de ESIGN ve UETA, Türkiye'de ise 5070 sayılı Elektronik İmza Kanunu kapsamında düzenlenir.
- Türkiye'de bu sertifikalar BTK tarafından yetkili elektronik sertifika hizmet sağlayıcıları (ESHS) tarafından verilir. Avrupa'da benzer görev, Qualified Trust Service Providers (QTSP) tarafından yürütülür.
- Zaman damgası bulunmayan e-imzalı belgelerde oluşturma tarihi manipüle edilebileceğinden geçersiz olarak kabul edilir veya bu durum ilgili kurum/kuruluşa aktarılacak şekilde önlemler alınmalıdır.
- Dijital belgeye herhangi bir müdahalede bulunulması halinde bu teknik sertifikalar belgeden silinir veya hata uyarıları verir.

2.4. Belge ve Görüntü Kalitesi Kontrolü

2.4.1. Tarama İşlemi

Evde ya da ofiste kullanılabilen orta segment düz yataklı tarayıcı özelliğine sahip yazıcılar veya otomatik besleme özelliğine sahip profesyonel yazıcılar OCR işlemlerine uygun yüksek kaliteli belge dijitalleştirme işlemleri için yeterli performansı sağlar. Özellikle metin ağırlıklı belgelerde karakter tanıma oranının yüksek seviyede olması için bu tarz tarama sistemlerinin kullanılması önemlidir. Ayrıca resmi belge çevirisi işlemlerinde noter onayı, yeminli tercüme ve resmi başvuru gerektiren işlem süreçlerinde belgelerin fotokopilerinin çeviriye ek olarak sunulması gerektiğinden profesyonel şekilde taranmış müstenidat ekler sunulan çeviri belgelerin kabul edilme ihtimalini doğrudan etkiler. Tarama işlemlerinde dikkat edilmesi gerekenler;

- Çözünürlük: Cihaz en az 300 dpi tarama kapasitesine sahip olmalı, metin yoğun belgelerde 600 dpi boyutuna kadar kadar çıkabilmelidir. Daha düşük çözünürlükte alınan görüntülerde metin bulanık çıkar ve OCR yazılımı harfleri doğru algılayamaz. Tarayıcı kullanıldığı durumlarda en az 300 dpi ile tarama yapılmalıdır.
- **Renk Desteği:** Gri tonlama ve renkli tarama seçenekleri bulunmalıdır. Renkli mühür, kaşe veya imzaların yer aldığı belgelerde renkli mod kullanılmalıdır. Sadece siyah-beyaz tarama yapan cihazlar ayrıntı kaybına yol açabilir.
- **Tarama Alanı:** Tarayıcı cam yüzeyi, standart A4 belgelerin yanı sıra daha büyük boyutlu özel basım belgeleri, diplomaları veya fiziksel belge güvenlik önlemlerine zarar vermeden alabilecek genişlikte olmalıdır. Minimum tarama alanı A4 (210 x 297 mm) olmalı, tercihen A4+ veya legal boyutları da desteklemelidir. Sayfa dışına taşan belgelerde kenar kaybı olur ve bu durum belge bütünlüğünü bozar.
- Yüzey Teması ve Netlik: Düzgün tarama için belge cam yüzeyle tam temas etmeli, kapağın kapatılmasıyla belge düz şekilde sabitlenmelidir. Kat izi, kıvrılma veya sayfa yükselmesi varsa taramadan önce düzeltilmelidir. Bu sayede metin düzgün hizalanır ve karakterler bozulmaz.
- Çoklu Sayfa Desteği (Varsa): Çok sayfalı belgeler taranacaksa otomatik belge besleme (ADF) özelliği olan cihazlar tercih edilebilir. Ancak ADF'li cihazlarda belge sıkışması, hizalama kaybı gibi sorunlar yaşanabileceğinden dikkatli kullanılmalıdır. Ayrıca çok sayfalı taramalarda belge sayfaları ayrı olması gerektiğinden evrakta sahteciliği önleyen fiziksel güvenlik önlemlerine zarar vermek belgenin resmiyetini tehlikeye sokabilir.

• Dosya Kaydetme ve Format Desteği: Cihaz PDF, JPEG ve PNG formatlarında çıktı verebilmelidir. Çok sayfalı belgeler tek bir PDF dosyası içinde toplayabilmeli, sayfa sırası doğru olmalı veya profesyonel doküman editör yazılımlarıyla uyumlu çalışabilmelidir. Tarama sonrası belge içeriği okunabilirliği, kenarlar vb. kontroller mutlaka yapılmalıdır.

2.4.2. Fotoğraf Çekimi

OCR çıktısının kalitesi kullanılan belge görüntüsünün kalitesine doğrudan bağlıdır. Düşük çözünürlüklü, bozulmuş veya belge dışında öğeler (el, sayfa üzerinde leke, gölge, belge kırışıklıkları vb.) görüntüler OCR yazılımında hatalı tanımlamalara yol açabilir. Özellikle eski belgelerin dijitalleştirilmesi sürecinde yüksek kaliteli görüntülerin elde edilmesi kritik bir adımdır. Bu kısımda belgelerin fotoğraflanması sırasında dikkat edilmesi gereken temel kurallar ve teknik detaylar ele alınmıştır.

Kamerayla çekilen belgeler genellikle .jpg veya .heic (iPhone) uzantılı olur. Ancak bu tür görseller bazı sorunlara yol açabilir:

- Düşük çözünürlük veya bulanıklık: Telefon kamerasıyla belge çekimi hiçbir zaman yazıcı taraması kalitesinde olmayacak. Bu sebeple de OCR işleminde karakterlerin bozulmasına ve yanlış tanınması kaçınılmazdır. Aynı zamanda belgenin fotokopisinin de çeviriye eklenmesi gerektiğinden fotoğraf resmi belgede profesyonel bir etki sağlamaz ve belgenin resmi yapısından şüphe duyulmasına neden olabilir.
- Eğimli veya kırpılmış sayfalar: Kamera açısının ayarlanamaması, sadece yazılara odaklanma veya ihtiyaç olduğu düşünülen kesimi gönderme gibi uygulamalar da bulunmaktadır. Ancak resmi evraklarda eksiksiz aktarım amaçtır ve belgede güven kaybına neden olabilecek her türlü ihtimale karşılık tamamının görüntüde bulunması gereklidir.
- **HEIC format uyumsuzluğu:** iPhone ile çekilen fotoğraflar .heic (live photo) uzantılı olabilir. Bu format bazı bilgisayarlarda veya OCR programlarında tanınmaz. Bu durumda .jpg ya da .png'ye dönüştürülmesi gerekir.

Yapılması gerekenler:

- **Parlama yapmayacak açıyla çekin:** Belgeyi çekerken üzerinde gölge oluşmamasına ve ışığın doğrudan parlamamasına dikkat edin. Özellikle kimlik, ehliyet, diploma gibi parlak yüzeyli belgelerde ışık doğru açıdan gelmezse yazılar okunamaz hale gelir. Kamerayı belgeye doğrudan yukarıdan değil, yanlardan gelecek şekilde ayarlayın. Genellikle OCR yazılımlarında ve belge tarama uygulamalarında görüntü açısını düzeltme özelliği olduğundan perspektif sorunu parlamaya göre kolay düzeltilebilir bir sorundur.
- **Uygun mesafeyi belirleyin:** Telefonu belgeden yaklaşık bir karış (20–30 cm) uzakta tutun. Belge net ve dört köşesi fotoğrafın içinde görünmelidir. Belgenin eğimli kısımları varsa düzleştirin, düz bir zemine koyun. Yamuk veya eğimli görüntülerde OCR yazılımları yazıları veya dizgisini doğru tanıyamaz.
- **Belgeyi düz yerleştirin:** Sayfa buruşuk, katlı veya kıvrılmışsa düzleştirin. Özellikle kenarları kıvrılmış belgelerde kenar bilgileri kaybolabilir. Belgeyi çekmeden önce düzgün hale getirin, kıvrılmış yerleri parmakla bastırarak veya kıvrımı tersine kırarak düzeltebilirsiniz.

- Her seferinde tek sayfa kullanın: Birden çok sayfalı belgelerde diğer sayfalar da zemin algılamasını zorlaştırabilir, diğer sayfadaki yazılar da OCR tarafından aynı sayfadaki yazıların bir parçası olarak algılanabilir.
- **Işık ve gölgeyi dengeleyin:** Aydınlatma mümkün olduğunca eşit olmalı. Belgenin bir tarafı aydınlık, diğer tarafı gölgede kalmamalı. Işığın arkadan veya doğrudan kameraya gelmesi parlamaya yol açabilir. Gölge veya parlama varsa çekimi yeniden yapın.
- Netlik kontrolü yapın: Kamera otomatik olarak netlik ayarı yapmıyorsa ekran üzerinde belgeye dokunarak odaklamayı manuel olarak yapın. Yazılar bulanıksa veya net değilse fotoğrafi yeniden çekin. Net olmayan görüntüler OCR tarafından okunamaz.
- Düz ve zıt renkli zemin kullanın: Belgeleri elde değil, düz ve belge kağıdının zıttı renklerde düz zeminlere koyarak çekin. Saydam veya belge ile aynı renk zeminler kameranın odaklanmasını ve kontrast ayarını bozacağından bozulma veya netlik kaybı ortaya çıkabilir.
- **Ellerinizi sabitleyin:** Eliniz titriyorsa belgeyi sabit bir yüzeye koyup dirseklerinizi dayayarak çekim yapın. Gerekirse kitap veya kupa gibi sabit bir nesneyle telefonu destekleyebilirsiniz.
- Yeterli çözünürlükte telefon kamerası kullanın: Yeni nesil bir akıllı telefon (12 MP ve üzeri) genellikle yeterli kaliteyi sağlar. Ön kameralar yeterli kaliteye sahip olmayabilir.

2.4.3. Mobil Tarama Uygulamalarını Kullanma

Mobil tarama uygulamaları, belgeyi doğrudan telefondan PDF veya görsel formata dönüştürmek için pratik bir alternatiftir. Özellikle tarayıcıya erişimin olmadığı durumlarda hızlı çözüm sunar. OCR öncesi belge kalitesinin korunabilmesi için uygulama seçiminde aşağıdaki teknik özelliklere dikkat edilmelidir:

Uygulama Seçerken Dikkat Edilmesi Gerekenler:

- Otomatik kenar algılama ve düzleştirme: Uygulamanın belge kenarlarını algılayıp yamukluğu düzeltebilmesi önemlidir.
- Filtre ve netleştirme seçenekleri: Soluk ya da gölgeli belgelerde okunabilirliği artırmak için kontrast ve parlaklık filtreleri içermelidir.
- Çoklu sayfa desteği ve PDF oluşturma: Sayfaları sırayla tarayıp tek bir PDF dosyası olarak kaydedebilmeli, sayfa sırası kolayca düzenlenebilmelidir.
- **Reklamsız ve filigransız kullanım:** Uygulamanın ücretsiz sürümde belge üzerine filigran eklememesi ve kullanıcıyı reklamla bölmemesi tercih edilir.

Önerilen Uygulamalar:

- **Microsoft Lens:** Ücretsizdir, reklamsız çalışır, PDF veya resim formatında dışa aktarım yapar. Taranan belgeye filigran eklemez.
- iPhone Notlar Uygulaması (iOS): Cihaza önceden yüklüdür, ek uygulama gerektirmez. Taranan belgeyi doğrudan not olarak kaydeder, PDF olarak paylaşılabilir. Filigran içermez.
- Adobe Scan: Ücretsiz sürümünde PDF kaydetme, otomatik kenar düzeltme, gelişmiş filtreleme seçenekleri ve çok sayfalı PDF oluşturma mevcut, ücretli sürümünde metin tanıma (OCR), PDF imzalama ve düzenleme özelliklerine sahiptir. Belgeleri buluta kaydedebilir, dışa aktarma yapılırken filigran eklemez.

Ek Tavsiyeler:

- Uygulama ile tarama yaptıktan sonra her sayfanın netliğini ve kadraja oturup oturmadığını manuel kontrol edin.
- PDF olarak dışa aktarırken dosya adını belge türüne ve tarihe göre düzenleyin.

Mobil uygulamalarla doğru tarama yapıldığında masaüstü tarayıcılara yakın kalitede OCR uyumlu dijital belge elde edilebilir.

2.4.4. Dosya Formatı, Saklama ve Gönderme

- Kamerayla çekilen belgeler mutlaka net, düzgün, yüksek çözünürlükte ve .jpg ya da .png formatında saklanıp aynı şekilde iletilmelidir.
- Belgeye ait dört köşe fotoğraf içinde kalacak şekilde kadraj dışı kalmamalıdır.
- Gerekirse belge düzenleme uygulamaları (Adobe Scan, Microsoft Lens, CamScanner gibi) ile sayfa kırpma ve düzleştirme işlemi yapılabilir.
- Tarayıcıdan alınan belgeler genellikle .pdf, .jpg, .tiff veya .png olarak kaydedilmelidir, işletim sistemlerine özel dosya formatları kullanılmamalıdır.
- Belgeler dili, çeviri ise dil çifti, belge sahibi veya proje adı, versiyon ve içerik bilgisi içerecek şekilde adlandırılmalıdır. Örnek olarak TR-EN_Kemal ISTIL_Mahkeme Kararı_V1 veya TR_Kemal ISTIL_Transkript.pdf
- Belgeler hem bilgisayar hem de -en azından iş teslimine kadar- bulut ortamında (Google Drive, Dropbox, OneDrive vb.) yedeklenmelidir. USB bellek gibi ikinci bir depolama da önerilir.
- Eğer fiziksel belgenin de saklanması ve taşınması gerekiyorsa aslı lazım olmadıkça ofis alanında saklanmalı, taramalar veya fotokopiler kullanılmalıdır. Belgenin taşınması gerekiyorsa dosya poşeti, çanta içi klasör ve su geçirmeyen evrak çantası kullanın. Fiziksel dokümanlara zarar gelmesi halinde tazmin ve yasal yükümlülükler bulunduğu unutulmamalıdır.

3. OCR SIRASINDA DİKKAT EDİLMESİ GEREKENLER

Belge çevirisinde belgenin aslındaki dizgiye sadık kalınması için yasal bir zorunluluk yoktur ancak belgelerin amacına uygun, okunaklı ve estetik şekilde hazırlanması çevirinin dolaşımı ve pazarlanmasını kolaylaştırdığından tercih edilir. Bu da çeviri hizmetlerinde kalite algısını güçlendirerek müşteri sadakatini arttıran etkenlerden biridir.

3.1. ORC Yazılımı Seçiminde Dikkat Edilmesi Gerekenler

OCR (Optik Karakter Tanıma) yazılımı seçerken yalnızca tanıma kalitesi değil, aynı zamanda veri güvenliği, kapsamlı belge desteği, kullanım alanı ve entegrasyon yetenekleri gibi faktörler de dikkate alınmalıdır. Özellikle çeviri hizmet sağlayıcıları açısından, işlenen belgelerde müşteri bilgileri, sözleşmeler, kimlik belgeleri ve özel içerikler yer alabileceğinden gizlilik ihlali riski taşımamak büyük önem taşır.

- Karmaşık belge yapılarında dahi harf ve satır hatalarını en aza indirmelidir.
- Türkçe dahil çok dilli tanıma sağlayan sistemler tercih edilmelidir.
- Belgelerdeki tablo, başlık, numaralandırma ve hizalama yapıları bozulmamalıdır.
- Açık olarak paylaşılmamış metinler için online OCR araçları kullanılmamalıdır.
 - Çevrimiçi (cloud tabanlı) OCR araçları, belge içeriğini dış sunuculara yükleyerek işler.
 - Bu da müşteriye ait kişisel veya ticari verilerin üçüncü taraflarla paylaşılması anlamına gelebilir.
 - Özellikle resmi belgeler, kişisel veriler ve sözleşmeler gibi hassas içeriklerde, bu tür araçların kullanımı veri ihlali riskine yol açar.
- Bu nedenle öncelikle masaüstünde çalışan, veriyi yerel diskte işleyen OCR araçları tercih edilmelidir. Böylece çeviri hizmet sağlayıcısı olarak KVKK ve benzeri veri koruma yükümlülüklerini ihlal etmeden çalışmış olursunuz.

Araç Adı	Çalışma Şekli	Veri Güvenliği Notu	Kullanım Özeti
ABBYY FineReader PDF	Masaüstü	Veriler cihazda kalır, yüksek güvenlik sağlar	Profesyonel kullanım, çok dilli destek
Adobe Acrobat Pro DC	Masaüstü + Bulut	Dosya yerel işlenebilir, bulut seçeneği devre dışı bırakılabilir	PDF düzenleme + OCR
Tesseract OCR	Masaüstü (açık kaynak)	Tamamen yerel çalışır, geliştiriciye uygundur	Özelleştirilebilir OCR çözümleri
Readiris	Masaüstü	Yerel çalışma, dosyalar kullanıcıda kalır	Küçük işletmelere yönelik kullanıcı dostu arayüz
Microsoft OneNote	Masaüstü + Senkronizasyon	Cloud senkronizasyonu devre dışı bırakılabilir	Basit belgeler ve görsel OCR notları
Google Cloud Vision	Çevrimiçi	Veri dış sunucularda işlenir, gizlilik riski vardır	Geliştiricilere yönelik API, toplu tanıma
Amazon Textract	Çevrimiçi	Bulut ortamında çalışır, gizli belgeler için uygun değildir	Form, tablo ve yapılandırılmış belge işleme

3.1.1. Farklı OCR Araçları ve Özellikleri

Not:

Hizmet sağlayıcılar için öncelikli önerimiz ABBYY FineReader, Adobe Acrobat Pro'nun yerel OCR özelliği, Tesseract gibi masaüstü çözümleri tercih etmeleridir. Zorunlu olmadıkça internet tabanlı sistemler kullanılmamalı, kullanılıyorsa da müşteriden açık rıza alınmalı ve kullanım sözleşmelerine bu husus açıkça eklenmelidir.

3.2. Görüntü Düzenleme (Image Editing)

OCR işleminden önce belge görsellerinin uygun şekilde işlenmesi tanıma doğruluğunu doğrudan etkiler. Profesyonel OCR yazılımları tarama veya kamera yoluyla elde edilen belgelerdeki hizalama, bozulma, renk dengesi gibi sorunları düzeltmek için kapsamlı ön işleme araçları sunar. Örnek olarak ABBYY Finereader yazılımı ve görselleri kullanılacaktır.

3.2.1. Önerilen Önişleme (Recommended Preprocessing)

Belge yapısını otomatik olarak analiz ederek en uygun iyileştirme adımlarını uygular. Eğiklik düzeltme, kontrast artırma, döndürme gibi işlemleri tek adımda gerçekleştirir. Hızlı ve güvenli başlangıç için önerilir.

3.2.2. Eğriliği Giderme (Deskew)

Tarama sırasında oluşan sayfa eğikliklerini giderir, metni yatay düzleme hizalar.

3.2.3. Metin Satırlarını Düzleştirme (Straighten Text Lines)

Kitap gibi kavisli taramalarda veya çok sayfalı belgelerde kıvrılan kısımdaki satırlar düzeltilir. Satır hizalarının düzeltilmesi, karakterlerin bozulmadan tanınmasına olanak sağlar.

3.2.4. Fotoğraf Düzeltme (Photo Correction)

Mobil cihazlarla çekilen belge görüntülerinde ortaya çıkan perspektif bozulmalarını düzeltir. Sayfa kıvrımlarını düzleştirerek daha düz bir görüntüye dönüştürür.

3.2.5. Yamuk Bozulmalarını Düzeltme (Correct Trapezoid Distortion)

Belge görseli eğimli veya yamuksa dört köşeden yeniden hizalayarak sayfayı düzleştirir.

3.2.6. Döndürme ve Çevirme (Rotate & Flip)

Belge yönü doğru ayarlanmamış veya aynalı şekilde çekilmiş/taranmış ise doğru şekilde yönlendirmeye olanak sağlar.

3.2.7. Bölme (Split)

Birden fazla sayfanın yer aldığı görsellerde ilgili sayfaları dikey ve yatay düzlemde ayırır.

3.2.8. Kırpma (Crop)

Belgenin dışında kalan boşluklar, masa kenarları veya tarayıcı çerçevesi gibi belge dışında kalan unsurları siler. Benzer bir işlem Yamuk Bozulmalarını Düzeltme ile de yapılabilir.

3.2.9. Ters Çevirme (Invert)

Negatif (beyaz yazı – siyah zemin) görüntülerde renkleri tersine çevirerek OCR'ye uygun hale getirir.

- <u>'-</u> *-	<u>Ö</u> nerilen Önişleme
	<u>Eğriliği giderme</u>
$\stackrel{\rm T}{\sim}$	<u>M</u> etin Satırlarını Düzleştir
0	<u>F</u> otoğraf Düzeltme
₽	Yamuk Bo <u>z</u> ulmalarını Düzelt
20	<u>D</u> öndür ve Çevir
1	Böl
4	<u>K</u> ırp
Ø	Ters Çe <u>v</u> ir
H	Çözünürlük
:0:	<u>P</u> arlaklık ve Karşıtlık
μ.	Seviyeler
\Diamond	<u>S</u> ilgi
8	Renk <u>İş</u> aretlerini Kaldır

3.2.10. Çözünürlük (Resolution)

Belgenin DPI (dot per inch) değeri kontrol edilip gerekirse ayarlanabilir. Görüntüleri küçültmede etkilidir ancak düşük çözünürlüklü görüntüleri iyileştirmede kullanılamaz.

3.2.11. Parlaklık ve Karşıtlık (Brightness & Contrast)

Soluk veya aşırı koyu görsellerde parlaklık ve karşıtlık ayarlayarak metni arka plandan ayırır.

3.2.12. Seviyeler (Levels)

Gri tonlamalı görüntülerde açık ve koyu tonlar arasındaki dağılımı dengeler.

3.2.13. Silgi (Eraser)

Kullanıcının, belge üzerinde bulunan istenmeyen lekeler, çizgiler, el yazısı notlar vb. öğeleri silmesini sağlar. Silgi alanı seçildiği alanın yoğunluğundaki arka plan rengine uyumlu şekilde değişir.

3.2.14. Renk İşaretlerini Kaldırma (Remove Color Marks)

Fosforlu kalem, arka plan rengi veya renkli damga gibi OCR'yi engelleyebilecek öğeleri filtreleyerek metnin net biçimde kalmasını sağlar.

3.3. Alan Türlerinin Seçimi

Metin tanıma alanlarının doğru seçilmesi, oluşturulan metin dosyasında düzenleme sırasında dağılma ve kaymalar sebebiyle oluşacak vakit kaybını önler. Alan türleri ABBYY FineReader gibi profesyonel OCR yazılımlarında kullanıcı tarafından manuel olarak seçilebilir veya otomatik olarak yazılım tarafından tanımlanabilir.



Görsel ABBYY Finereader 16 OCR Editör sürümünden alınmıştır.

3.3.1. Metin Alam (Text Area)

OCR yazılımının çıktı olarak düzenlenebilir metin formatına dönüştüreceği bölgeleri tanımlar. Düz yazılar, alt bilgi ve üst bilgi, maddeler, başlıklar vb. tüm metinsel içerikler bu alana dahil edilir. OCR aracı tanımlanan metin alanları üzerinden satır hizalarını ve paragraf yapılarını analiz ederek karakterleri dijital metne dönüştürür.

- Yazı tipleri, boyutlar, hizalama ve paragraf yapısı OCR çıktısına yansıtılır.
- Metin alanı dışında kalan yazılar tanınmaz.
- Doğru dil tanımı ve OCR dil ayarı ile birlikte karakter tanıma başarısı önemli ölçüde artar.
- Metin alanları, çıktı biçimine göre düzenlenebilir formatta (DOCX, TXT vb.) kaydedilir.

3.3.2. Tablo Alanı (Table Area)

Yapısal olarak satır ve sütunlardan oluşan veri bloklarının tanınmasını sağlar. OCR yazılımı bu alanlarda hücre bölünmelerini algılayarak verileri düzenli bir tablo yapısı içinde çıkarır.

- Satır ve sütun yapısı korunarak çıktı alınır (örnek olarak Excel veya düzenlenebilir tablo içeren DOCX).
- Tablo tanımlaması yapılmazsa veriler metin gibi algılanır ve düzen bozularak metin kutuları veya sayfa sütunları içine yayılabilir. Bu durum CAT Tool segmentasyonunda da büyük sorunlara neden olur.
- Kenarlıksız, sütunlu yazılar vb. içerikler de metin hizasına göre yapay olarak tablolama yoluyla düzen korunarak ayrıştırılabilir.
- Finansal belgeler, formlar, ders çizelgeleri gibi alanlar için süreci fark edilir ölçüde kolaylaştırır.

3.3.3. Görsel Alan (Picture Area)

OCR işlemi yapılmayacak, yalnızca görsel olarak korunması istenen alanlar için kullanılır. Bu alanlar içinde kalan içerikler (fotoğraflar, mühürler, imzalar, grafikler) OCR'ye tabi tutulmaz, sadece görüntü olarak aktarılır.

- OCR yazılımı bu alanları atlayarak doğrudan çıktı dosyasına görsel olarak ekler.
- PDF formatında görüntülenebilir; DOCX gibi düzenlenebilir formatlarda taşınabilir öğe olarak yer alır.
- Logolar, şekiller, grafikler bu şekilde ayrılmalıdır.
- Katalog, teknik çizim, vb. grafik tasarım kapsamına girmeyen resmi belge işlerinde görüntülerin alınmasına gerek yoktur.

3.3.4. Arka Plan Resmi (Background Picture)

OCR yapılmayacak, ancak belge yapısında görsel olarak korunması gereken sabit alanları tanımlar. Genellikle sayfa desenleri, filigranlar, belge şablonları veya güvenlik öğeleri bu sınıfa girer.

- Tanıma işlemi yapılmaz, yalnızca görsel olarak belgeye entegre edilir.
- OCR sırasında görsel gürültü oluşturabilecek arka planlar bu şekilde yalıtılarak tanıma doğruluğu artırılır.
- Özellikle damgalı belgeler, arka planında metin bulunan evraklar için kullanılır.
- Arka Plan Resmi alanı üzerine manuel olarak Metin Alanı (Text Area) eklenerek görsel üzerindeki yazıların OCR'ye dâhil edilmesi sağlanabilir.

3.3.5. Zorlu Alanlarda Manuel Kontrol

- El Yazısı veya Karmaşık Düzenler: El yazısı notlar, çizimler veya farklı yönlere yerleştirilmiş metinler (örnek olarak SmartArt, şema vb.) OCR sürecinde hatalara neden olabilir. Gerekirse bu kısımları manuel düzeltin veya ayrıca işaretleyin.
- Ters Çevrilmiş Metinler: Tarama sırasında sayfa ters çevrilmiş veya baş aşağı kalmışsa ve OCR aracı bunu otomatik düzeltmediyse yanlış algılayabilir. Ayrıca dikey basılmış ancak yatay hazırlanmış sayfalarda da farklı yönde metinler tespit edilmelidir.

4. OCR SONRASI WORD DİZGİ ÇALIŞMASI

4.1. Görünmez Karakterler ve Özel İşaretler

K 7 <u>A</u> ~ ab	<u>P</u> aragraf İşareti	 ✓ tiller 	Düzenleme	Adobe PDF	İmza	
⁸ 🗛 - 🖌 - 🗛 -	Sek <u>m</u> e Karakteri	*	×	Oluştur ve Paylaş	İste	
Yazı Tiş	<u>H</u> erhangi Bir Karakter	ler 🗔		Adobe Acrob	at	
Bul ve Değiştir	He <u>r</u> hangi Bir Basamak				?	×
	H <u>e</u> rhangi Bir Harf					
B <u>u</u> l <u>D</u> eğiştir	Şap <u>k</u> a Karakteri					
Ara <u>n</u> an: ->	§ Bölüm K <u>a</u> rakteri					\sim
	1 Paragraf Karakteri					
	Süt <u>u</u> n Sonu					
Yeni değer: ->	U <u>z</u> un Tire					~
	Tire					
Contraction Contraction	Son Not İşareti	Tûmû	nii Dožictic	Conrolvini Rul	intal	
Arama Sasanaklari	Ala <u>n</u>	ru <u>m</u> u	inu Degiştir		iptai	
Arama seçenekleri	<u>D</u> ipnot İşareti					
Arama yonu: Tumu	Graf <u>i</u> k					
Valnızca tam sözci	El ile Satır S <u>o</u> nu			onek eş <u>l</u> eştir		
Joker karakter kull	El ile Say <u>f</u> a Sonu			one <u>k</u> eşleşti		
Benzer ok <u>u</u> nuşlar (Böl <u>ü</u> nemez Kısa Çizgi			Joktalama karakterlerir	ni <u>v</u> oksay	
T <u>ü</u> m sözcük formla	B <u>ö</u> lünemez Boşluk		В	oşluk karakterlerini yo	ksay	
	<u>S</u> eçmeli Kısa Çizgi					
	<u>B</u> ölüm Sonu					
Değiştir	Boş <u>l</u> uk					
<u>B</u> içim ▼	<u>Ö</u> zel▼ Biçimlendirme	Yok				

OCR yazılımları metni işledikten sonra Word (DOCX) formatına aktarırken satır sonları, sekmeler, boşluklar gibi biçimlendirme işaretlerini de belgeye dahil eder ama bu karakterler Word'de görünmez. Biçimlendirme işaretleri Word düzeninin arka planında hangi metnin nasıl yerleştirildiğini gösterir ve doğru görüntülenmediklerinde aşağıdaki sorunlara yol açabilir. CAT Tool segmentasyonunda bu işaretler de etiket (tag) veya ayrı sekme olarak yer alır.

4.1.1. Paragraf ve Satır Sonu İşaretleri

- Paragraf İşareti (¶): Enter ile oluşur. Yeni paragraf başlatır, segmenti keser.
- El ile Satır Sonu (¬): Shift+Enter ile oluşur. Aynı paragraf içinde alt satıra geçer, segmenti bölmez. CAT araçlarında birleşik algılanır.
- Bölüm Sonu: Sayfa yapısını dosyanın sadece ilgili kısmı için değiştirir (Örnek olarak tek sütun ve çift sütun arka arkaya eklenebilir; yatay ve dikey sayfalar aynı belgede oluşturulabilir).
- El ile Sayfa Sonu: Ctrl+Enter ile oluşturulur, belgeyi yeni bölüme ayırmada sonraki sayfada başlatır.

Daha temiz CAT Tool segmentasyonu açısından gereksiz satır sonları silinmeli; paragraf sonları doğru yerde bırakılmalıdır.

4.1.2. Sekme, Boşluk ve Hizalama Karakterleri

- Sekme Karakteri (\rightarrow): Tab tuşuyla oluşturulur, hizalama amacıyla kullanılır.
- Boşluk (·): Normal boşluk karakteridir. Fazla kullanımı hizalama ve segment yapısını bozabilir.
- Bölünemez Boşluk (°): Ctrl+Shift+Boşluk tuşuyla oluşur. İki kelimenin satır sonunda ayrılmasını engeller (örn: "No°123").

- Bölünemez Kısa Çizgi (-): Ctrl+Shift+Kısa Çizgi tuşlarıyla oluşur. Tireli ifadelerin (örn: e-posta adresi) satır sonunda ayrılmasını engeller.
- Seçmeli Kısa Çizgi: Gerektiğinde bölünebilen, yazımda estetik için kullanılan kısa çizgidir.

OCR sonrası çoklu boşluk karakterleri genellikle sekme niyetine kullanılır; bunlar mutlaka temizlenmeli ve yerine gerçek sekmeler yerleştirilmelidir.

4.1.3. Sütun, Sayfa ve Son Not İşaretleri

- Sütun Sonu: Çok sütunlu belgelerde bir sütunu bitirir ve diğerine geçer.
- Bölüm Karakteri (§): Özel stil kullanımlarında tercih edilir; nadiren kullanılır.
- Son Not İşareti: Belge sonunda yer alan açıklamalarda kullanılır.

4.1.4. Metin İçi ve Gelişmiş Arama Karakterleri

Bu karakterler özellikle Bul ve Değiştir işlemlerinde kullanılır:

- Herhangi Bir Karakter/Harf/Basamak: Belirsiz içerikleri eşleştirmek için joker karakterlerdir.
- Şapka Karakteri (^): Bul ve Değiştir penceresinde özel karakterleri çağırmak için kullanılır. Örnek olarak "^p" ile arama yapılırsa paragraf işaretlerini, "^t" ile arama yapılırsa yapılırsa sekme (tab) işaretlerini seçer.
- Alan: Otomatik alan yer tutucularını temsil eder (örn. sayfa numarası, tarih).

4.1.5. Grafik, Dipnot ve İçerik Temsilleri

- Grafik: Resim, şekil, simge gibi nesneleri temsil eder.
- Dipnot İşareti: Metne bağlı açıklamaları alt bilgi olarak gösterir.

4.1.6. CAT Tool Segmentasyonuna Etkisi

Çeviri araçları (CAT Tools) metni çözümlemeye başlamadan önce bu işaretler üzerinden otomatik olarak segmentlere (cümle/paragraf bloklarına) ayırır, daha sonrasında noktalama işaretleri, sayılar ve vb. metin ayrıştırıcılarına göre segmentlere ayırır. Aşağıdaki hatalar segmentasyon dengesini bozar:

- Shift+Enter ile manüel satır kırmaları segmenti bölmez ama CAT içinde tag olarak görünür.
- Boşluk karakterleriyle hizalanan metinler CAT araçlarında kalite hatası olarak listelenir ve kalite kontrol sürecini uzatır.
- Paragraf yerine sekme veya çoklu boşluk kullanımı cümle yapısını tanımlanamaz hale getirir, CAT içinde fazladan tag oluşturarak çeviri süresini uzatır.
- Paragrafların yanlış yerde kesilmesi CAT aracının çeviri belleğiyle eşleşmesini bozar.

Belgeye doğru noktalarda Paragraf, Sayfa Sonu, Sekme kullanarak doğal metin akışı sağlanmalıdır. Gereksiz satır sonları toplu biçimde kaldırılmalıdır (Ctrl+H \rightarrow El ile satır sonu).

4.2. Tablolama

OCR sonrası, tablo düzeni çeviri işlemlerinde ve belge düzeninde en çok hata çıkan alanlardan biridir. Satır ve sütun hizalamaları, hücre yüksekliği, metin taşmaları ve görünmeyen içerikler doğru biçimde kontrol edilmelidir. Tablo seçildiğinde üstte iki menü görünür:

- Tablo Tasarımı: Çerçeve çizgileri, gölgelendirme ve tablo stili ayarları
- Düzen: Hücre ekleme/silme, hizalama ve boyutlandırma işlemleri

4.2.1. Hücre İçeriğini Kontrol Etme

OCR ile oluşturulan hücreler sabit satır yüksekliğiyle gelir. Metnin yazı tipi ayarlarında değişiklik yapılması veya çeviri sonrası kelime sayısının değişmesi sonrası tablo sınırına sığmayan içerikler görünmez hale gelir. Görünmeyen metinler CAT Tool tarafında sorunsuz ortaya çıkarken Word tarafında fark edilmesi zordur. Tablolarda bu tarz kesilmeleri anlamak için:

- Biçimlendirme işaretleri açıkken hücrede ¤ (hücre işareti) görünmelidir.
- Eğer görünmüyorsa ve imleç hücredeyken aşağı kayıyorsa kesilmiş içerik vardır ama görünmüyordur.

4.2.2. Hücre İçi Düzenleme ve Kısayollar

Kısayol	İşlevi
Ctrl+Tab	Hücre içinde sekme oluşturur (hücre değiştirmeden).
Tab	Sonraki hücreye geçer.
Shift+Tab	Önceki hücreye geçer.
Shift+Enter	Aynı hücre içinde alt satır oluşturur.
Ctrl+Shift+Enter	Tabloyu bulunduğu satırdan böler ve ayrı tablolar haline getirir.

4.2.3. Tablo Düzenleme Taktikleri

OCR sonrası oluşan tabloların düzen ayarlarını yaparak görünmeyen içerikleri açığa çıkarabilir, hizalama sorunlarını giderebilir ve sayfa yapısındaki bozulmaları önleyebilirsiniz. Bunun için aşağıdaki adımları uygulayın:

- Hücre içeriğini hizalayın:
 - Hücreleri seçin \rightarrow Düzen sekmesi \rightarrow Hizalama bölümünden üst, orta, alt ve sola, ortaya, sağa hizalama seçeneklerinden birini uygulayın.
 - Başlık hücrelerinde genellikle yatay ve dikey orta hizalı, içerik hücrelerinde üst ve sola yaslı hizalama tercih edin.
- Tabloyu içeriğe ve sayfa genişliğine göre ölçeklendirin:
 - Tablonun tamamını seçin \rightarrow Düzen \rightarrow Otomatik Sığdır \rightarrow İçeriğe Uydur komutunu uygulayın.
 - O Ardından tekrar Düzen → Otomatik Sığdır → Pencereye Sığdır adımını da uygulayarak tabloyu sayfa genişliğiyle hizalayın.
 - veya
 - Tablonun tamamını veya düzenlemek istediğiniz sütunları seçin → Düzen sekmesi → Sütunları Eşit Dağıt komutunu çalıştırın.

Bu şekilde OCR kaynaklı orantısız sütunları dengelemiş olursunuz.

- Satır yüksekliğini serbest bırakın:
 - \circ Tabloya sağ tıklayın → Tablo Özellikleri → Satır sekmesi → "Yüksekliği belirt" kutusunu işaretleyin.
 - Türü "En az" olarak ayarlayın ve değeri 0 cm yazın.

- Tüm metni seçin \rightarrow Giriş \rightarrow Paragraf \rightarrow Satır Aralığı bölümünden En az 1 satır seçeneğini uygulayın.
- Bu sayede OCR'de 1 satırın altında kalmış metinleri görünür hale getirin.
- Metin kaydırmayı kapatın:
 - Tabloya sağ tıklayın → Tablo Özellikleri → Metin Kaydırma sekmesi → "Yok" seçeneğini seçin.
 - Böylece tabloyu belge akışına entegre edin, çevresindeki metin kaymalarını önleyin.
- Sayfa sonunda oluşan silinemeyen boş sayfalardan kurtulun:
 - Tablo sonrasındaki zorunlu paragraf işaretine tıklayın.
 - Yazı tipi boyutunu 1 pt yapın.
 - Paragraf ayarlarını satır aralığı: Tam, değer: 1 satır, önce/sonra boşluk: 0 nk olarak ayarlayın.

Bu adımlarla gereksiz sayfa taşmalarını engelleyin.

- Satır içeriğini bölünmeden tutun:
 - Çok satırlı hücreleri seçin → Giriş → Paragraf iletişim kutusu → Satır ve Sayfa Sonları sekmesi üzerinden "Satırları birlikte tut" seçeneğini işaretleyin.
 - Böylece hücre içeriğini sayfa sonunda parçalanmadan durur.
- Küçük tabloların sayfa sonunda bölünmesini engelleyin:
 - Tabloya sağ tıklayın → Tablo Özellikleri → Satır sekmesi → "Satırın sayfa sonunda bölünmesine izin ver" kutusunun işaretini kaldırın.

Bu işlemleri uygulayarak tabloları yalnızca görsel olarak değil, teknik olarak da düzenli ve çeviriye hazır hale getirin. Belge içeriğinin eksik görünmesini, segment kaymalarını ve sayfa taşmalarını önleyin. Tablolarınız hem Word hem CAT ortamında sorunsuz çalışsın.

• Ayrıca tüm metnin paragraf ayarlarında satır aralığı da en az 1 satır olarak tanımlanmalıdır. Bu şekilde OCR'den 1 satır altında çıkmış metinler de görünür hale gelecektir.

4.3. Cetvel Kullanımı ve Sekmelerin Yerleştirilmesi

OCR sonrası ortaya çıkan Word belgelerinde hizalama sorunları sık görülür. Bu sorunlar genellikle görsel olarak hizalı görünen ancak teknik olarak bozuk yapıdaki boşluk ve sekme karakterlerinden kaynaklanır. Özellikle çeviri sırasında metin kaymaları, segment bozulmaları ve stil hataları yaşanabilir.

4.3.1. Cetvelin Görünür Hale Getirilmesi



Word'de cetvel varsayılan olarak kapalı olabilir. Görüntülemek için: Görünüm sekmesi → Cetvel kutucuğunu işaretleyin.

4.3.2. Girintiler

Sol girinti, paragrafın sol tarafının konumunu ayarlar. Sol girintiyi taşıdığınızda, ilk satır girintiniz veya asılı girintiniz eşitlenmiş olarak taşınır.

- Sol girintiyi ayarlamak istediğiniz metni seçin.
- Cetvelde, cetvelin sol alt tarafındaki kare işaretçisini paragrafın sol kenarında olmasını istediğiniz yere sürükleyin.

Sağ girinti işaretçisi paragrafın sağ tarafının konumunu denetler.

- Doğru girintiyi ayarlamak istediğiniz metni seçin.
- Cetvelde, cetvelin sağ alt tarafındaki üçgen işaretçiyi paragrafın sağ kenarında olmasını istediğiniz yere sürükleyin.

İlk satır girintisiyle, paragrafın ilk satırı girintilenir ve paragrafın aşağıdaki satırları girintilenmez.

- İlk satır girintisini eklemek istediğiniz metni seçin.
- Cetvelde, cetvelin sol üst kısmındaki üçgen işaretçiyi, girintinin başlamasını istediğiniz yere (örnek olarak 1" işaretine) sürükleyin.

Asılı girinti kullanıldığında, bir paragrafın ikinci ve sonraki tüm satırları ilk satırdan daha girintili olur.

- Asılı girinti eklemek istediğiniz metni seçin.
- Cetvelde, alt işaretçinin üst üçgen kısmını, girintinin başlamasını istediğiniz yere, örnek olarak 1" işaretine sürükleyin.

4.3.3. Sekme Durakları

Metni hizalamak için çok sayıda boşluk veya tab kullanmak yerine, cetvel yardımıyla sekme durakları (tab noktaları) tanımlanmalıdır. Bu sayede satırlar hem hizalı hem de teknik olarak düzenli olur. Genellikle belge içindeki bilgileri yerleştirirken kullanılır. Bu şekilde alt alta satırlarda metni sayfanın içine daha hızlı şekilde yayabilir ve uzun süren düzen bozuklukları ve kaymaların önüne geçer. Aynı zamanda CAT araçlarında da fazladan tab veya boşluklardan kaynaklı düzeltilemeyecek QA analizlerinin önüne geçer.



Sekme Türleri:

I Sol Solema Duroği	Metin satırının sol ucunu ayarlar. Siz yazdıkça, metin sağa doğru				
L Sol Sekille Dulagi	doldurulur.				
J Sağ Sekme Durağı	Metin satırının sağ ucunu ayarlar. Siz yazdıkça, metin sola doğru				
	doldurulur.				
I Orto Salema Durači	Metin satırının ortasındaki konumu ayarlar. Siz yazdıkça, metin bu				
- Orta Sekme Duragi	konumda ortalanır.				
L Ondelik Seleme Dureži	Sayıları ondalık ayırıcıya hizalar. Basamak sayısına bakılmaksızın,				
	ondalık virgülü hep aynı konumda kalır.				
	Metni konumlandırmaz. Sekme konumda bir dikey çubuk ekler. Diğer				
L Cubult Salema Durači	sekmelerden farklı olarak, cetveli tıklattığınız anda çubuk sekme metne				
	eklenir. Belgenizi yazdırmadan önce çubuk sekme durağını				
	temizlemezseniz, dikey çizgi yazdırılır.				

Sekme Nasıl Yerleştirilir?

- Sekme türünü seçmek için cetvelin sol üst köşesindeki simgeye tıklayın. (Her tıklamada sekme türü değişir.)
- Metni hizalayacağınız paragrafi seçin.
- Cetvelin üzerine tıklayarak sekme noktası bırakın.
- Klavyeden Tab tuşuna bastığınızda metin o sekme hizasına gider.

Not: Her paragrafın sekmeleri bağımsızdır. Yeni paragraf açıldığında sekme konumları kopyalanmaz.

Yanlış yerleştirilmiş sekmeler şunlara neden olur:

- Satır hizalarının bozulması
- Yanlış hizalanmış metin blokları
- Biçim olarak düzgün görünen ama teknik olarak bozuk içerikler

Sekme duraklarını menüden düzenlemek için:

Paragrafi seçin \rightarrow Sağ tıklayın \rightarrow Paragraf \rightarrow Sekmeler veya cetveldeki yerleşik sekme işaretine çift tıklayın.

Açılan pencereden:

- Seçili paragrafa ait tüm sekme noktaları listelenir
- Konum değiştirilebilir
- Hizalama türü ve noktalama karakterleri (nokta, tire, boşluk vb.) ayarlanabilir

Sekmeler		? ×
Se <u>k</u> me durak yeri:		<u>V</u> arsayılan sekme durakla
5 cm		1,27 cm
5 cm	•	Temizlenecek durak:
Hizalama		l
🗿 S <u>o</u> l	◯ O <u>r</u> ta	○ <u>S</u> ağ
Ondalık	◯ Ç <u>u</u> buk	
Öncü		
Ο <u>1</u> Yok	<u>2</u>	<u>3</u>
<u> </u>		
<u>A</u> yarla	<u>T</u> emizle	e Tü <u>m</u> ünü Temizle
	Tama	mintal

4.4. Sütunlar

OCR yazılımları çok sütunlu metinlerde karmaşa yaratabilir. Hatta tablolaması gereken metinleri sütunlu metin olarak algılayabilir. Dizgi çalışmasında sütunları korumanız gerekiyorsa Word'ün "Sütunlar" özelliğini doğru ayarladığınızdan emin olun.

iz	Tasarıı	n Düz	en	Başvurular	Posta Gönder	ileri Gözde	en Geçir	Görüni
		Sütunlar					?	×
ıt Sü Yapıs 1	ütunlar	Önceden B	elirlene	inler	<u> </u>	Sola	<u>S</u> ağ	a
	5	Gütun s <u>a</u> yısı: Genişlik ve Sü <u>t</u> un: 1: 	1 aralık <u>G</u> enişlil 17,19 tun ger	c A cm • [• [•]	Agalik:	Araya Önizlem	çizgi koy e	
	Ţ	<u>J</u> ygulama ye	eri: T	üm belgeye	<u> </u>	Veni s	ütuna başla İpt a	al

4.4.1. OCR Sonrası Sütun Hataları

OCR yazılımları çok sütunlu belgelerde metin yapısını algılarken çeşitli hatalar yapabilir. Bu hatalar özellikle belge içinde sütun gibi görünen ancak aslında liste, açıklama kutusu veya tarih/sayı bilgisi gibi tek paragraflık satırlar için yaygındır. Sistem, bu yapıları otomatik olarak ayrı sütunlar gibi algılayabilir ve metni iki ya da daha fazla kolona ayırabilir. Sonuç olarak dizgi sırasında metinler bölünür, hizalamalar kayar ve segment bütünlüğü bozulur.

Örnek olarak madde işaretli veya numaralı listeler, metin yoğunluğu ve hizalama düzeni nedeniyle "çok sütunlu" bir yapı gibi yorumlanabilir. Benzer şekilde, bir paragrafin sağına yerleştirilmiş tarih, açıklama, belge numarası veya iletişim bilgisi gibi kısa bilgiler de OCR algoritması tarafından ayrı sütun sanılabilir. Bu da Word belgesinde paragrafların karşılıklı iki blok halinde hizalanmasına yol açar.

Belge düzenini sağlam tutmak için aşağıdaki teknik kurallara dikkat edilmelidir:

- Biçimlendirme işaretleri ve cetvel açık şekilde varsa "Bölüm Sonu" işaretleri ve bulundukları paragrafların çok sütunlu olup olmadığı kontrol edilmelidir.
- Sütunlar belgenin yalnızca bir kısmındaysa "Bölüm Sonu" işaretinin silinmesi tüm belgeyi etkileyebilir. Böyle bir durumda OCR yeniden yapılmalı veya bölüm tamamen silinerek tekrardan doğru şekilde düzenlenmelidir.
- Sütun gerekmiyorsa tüm belge tek sütuna dönüştürülmeli ve hizalamalar tablo, sekme veya kenar boşluklarıyla sağlanmalıdır.
- Özellikle OCR sonrası yanlış sütunlama sonucu satır veya pagraf görsel olarak doğru hizalanacak şekilde düzenlense bile CAT bunları ayrı içerikler olarak algılayacaktır.

OCR işlemi sonrasında metnin "çok sütunlu" gibi görünmesi tek başına yeterli gerekçe değildir. Metnin gerçekten sütunlu yapıya ihtiyaç duyup duymadığı, içerik ve biçim bağlamında değerlendirilmelidir. Aksi halde çeviri sırasında metin akışı dağılır ve son düzenlemede ciddi zaman kaybı yaşanabilir.

4.5. Metin Kutuları ve Çerçeveler

OCR yazılımlarının "Tam Kopya" veya "Biçimlendirilmiş Metin" modları kullanıldığında sayfa görünümünü korumak amacıyla belgelerdeki tüm içerik (başlıklar, paragraflar, tablolar, numaralar vb.)

genellikle metin kutuları veya çerçeveler içinde tanımlanır. Bu yapıların görsel olarak avantajı olsa da metin dizgisi ve çeviri süreci açısından ciddi sorunlara yol açabilir.

OCR çıktılarında bu öğelerin ortaya çıkardığı başlıca sorunlar:

- İçerik akış dışı kalır. Metin kutuları ve çerçeveler belge akışına dahil değildir ve çeviri araçlarında içerik sıralaması bozulur veya kutu içindeki metin hiç görünmez.
- Tablolar çerçeveye dönüşebilir. Tablo olarak algılanmamış yapılar OCR tarafından çerçeve içine alınmış metin grupları olarak tanımlanabilir. Bu durumda tablo işlevselliği kaybolur, hücreler ayrı metin kutuları veya çerçeveler içinde kalır.
- Kutu veya çerçeve içindeki metinler bağımsız parça olarak algılanır ve belge bütünlüğü parçalanır.
- Sabit satır veya sınır dışı girinti seçiliyse metinler metin kutusu veya çerçeve dışında kalacağından görünmeyebilir. Bu da yazdırırken ya da PDF gibi dışa aktarımda metin kaybına neden olabilir.
- Altbilgi veya üstbilgi kısmında metin kutularında yer alan metinler Word içindeki karakter/kelime sayımında yer almaz. Bu şekilde karakter hesaplamalarında bu kısımlar hesap dışı kalmış olur.

4.5.1. Teknik Öneriler

- Kutu ve çerçeveler metin akışına entegre edilmelidir veya metinler buralardan kesilip düz paragraflar halinde belgeye yapıştırılmalıdır.
- OCR yazılımı ayarlarında biçim koruma azaltılmalı, "Metin İşlevi" seçimi gibi metnin doğrudan yerleşmesine yönelik ayarlamalar yapılmalıdır.
- Tablo içeriği çerçeve içindeyse tablonun çerçeveden çıkarılıp normal tablo olarak yeniden yerleştirilmesi gerekir veya çerçeveye çift tıklanarak açılan menüden "Çerçeveyi Kaldır" seçilmelidir.
- Çeviri araçlarında içerik görünmüyorsa metin büyük ihtimalle bir kutu içindedir. Görselde görünse bile çeviri belleği tanımayabilir.

4.6. Yazı Tipi, Boyut ve Standartlaştırma

OCR ile elde edilen Word belgeleri dışarıdan bakıldığında düzgün ve okunabilir görünebilir. Ancak belgenin arka planında farklı yazı tipi, yazı boyutu, renk ve biçim varyasyonları bulunabilir. Bu farklar çoğunlukla tarama kalitesi, OCR yazılımının yorumlama biçimi ve belgenin fiziksel durumu gibi faktörlerden kaynaklanır.

Özellikle yamuk çekilmiş/taranmış belgelerde, lekeli, yıpranmış veya kıvrılmış kısımlarda OCR algoritmaları aynı paragraf içinde yer alan farklı karakterleri farklı yazı tipi veya boyutta tanıma eğilimindedir. Bu yüzden de Word'e aktarıldığında görsel olarak çok belirgin olmasa bile CAT aracıyla açıldığında segment başına fazladan etiket (tag) eklenmesine neden olur.

Bu etiketler de çeviri esnasında ciddi vakit kayıplarına, çeviri hatalarına ve gereksiz QA uyarılarına sebep olabilir. Segment içinde farklı yazı tipiyle biçimlendirilmiş bir kelime bile olsa sistem bu farklılığı

 (ept>, <ph> gibi etiketlerle işaretler. Fazla tag segmentin iç yapısını karmaşıklaştırır, çeviri kalitesini ve teslimdeki bütünlüğü olumsuz etkiler.

4.6.1. CAT Araçlarında Segment İçine Tag Olarak Yansıyan Biçim Bozuklukları

- Aynı paragraf içinde bazı kelimelerin diğerlerinden farklı olarak benzer başka yazı tipiyle (örnek olarak bir kısmının Arial bir kısmının Calibri vb. olarak) algılanması gibi farklılıklar,
- Satırın bir kısmında tarama görselinin açısı, eğimi, kıvrımına geliyorsa diğer kısmından küçük veya büyük olarak algılanabilir. Bu küçük fark bile yeni tag oluşturur.
- Belgenin orijinalinde vurgulanmamış olmasına rağmen tarama, lekeler veya kıvrımlardan kaynaklı OCR sırasında karakter kalınlığındaki farklar kalın, italik, altı çizili olarak algılanabilir. Bu da temel metin düzenleme özelliği olmayan CAT yazılımlarında fazladan tag olarak görünür. MemoQ, Trados vb. yazılımlarda segment içinde metin kalın, italik, altı çizili ayarlanabildiğinden etiket olarak görünmez ancak düzeltmek istenildiğinde QA hatası olarak görünür.
- Gri arka planlı veya düşük kontrastlı bölgelerde karakterler koyu mavi, gri ya da kahverengi gibi algılanabilir. Bu, metin segmentinde renk etiketi oluşturur.
- Yanlışlıkla küçük puntolu algılanan yazılar otomatik olarak alt indisli karakterler (örn. dipnot numaraları) veya üst simgeler olarak tanımlanabilir. Bunlar hem fazladan tag oluşmasına hem de metnin yanlış algılanmasına neden olabilir.
- Aynı paragraf içinde birden fazla stil kalıntısı (örnek olarak başlık + gövde metni) varsa stil geçişlerini tag olarak aktarır.
- Segmentin başında ve sonunda tek harflik bile yazı tipi farklıysa, sistem tüm segmenti parçalayabilir ya da kodla boğabilir.

4.6.2. Yapılması Gerekenler

Ctrl + A ile tüm metni seçin ve ardından,

- Sağ Tıklama \rightarrow Paragraf menüsünü açın,
- Anahat Düzeyi → Gövde Metni şeklinde ayarlayın. Değişmesi gerekenleri kontrol ve son okuma sırasında Stillerle değiştirebilirsiniz.
- Paragraf öncesi ve sonrası için boşlukları sıfırlayın.
- Satır Aralığı'nı en az 1 olarak ayarlayın.
- Girinti kısmında girintileri kaldırın, varsa "Asılı Girinti" ayarını kapatın.
- "Satır ve Sayfa Sonu" ayarlarında yalnızca "Tek kalan satırları önle" ve "Satırları birlikte tut" kutucukları işaretli kalsın.
- Ardından Yazı Tipi menüsünü açın ve tüm yazı tiplerini tek bir boyuta getirin. Genellikle 11 veya 12 pt kullanılır.
- Aşağıdaki şekilde yazı üzerindeki tüm ayarları kaldırın. Bu da lekelerden ve eğimlerden kaynaklı hataları ortadan kaldıracaktır.

Yazı ti <u>p</u> i rengi:	Alt çizgi <u>s</u> tili: Alt çizgi rengi:			
Otomatik 🗸	(yok) 🗸	Otomatik 🗸		
Efektler				
🔲 <u>Ü</u> stü çizili		🗌 Küçük büyük		
🗌 Üstü çi <u>f</u> t çizili		🗌 Tü <u>m</u> ü büyük		
🗌 Üst simg <u>e</u>		🗌 <u>G</u> izli		
🗌 Al <u>t</u> simge				
Önizleme				
ayarlarında				
Bu bir TrueType yazı tipidir. Aynı yazı tipi hem yazıcıda, hem de ekranda kullanılacak.				

• Sonrasında Gelişmiş sekmesini açarak aşağıdaki ayarları yapın.

Ya <u>z</u> ı Tipi	Ge <u>l</u> işmiş			
Karakter Aralı	ğı			
Ölç <u>e</u> k: 9	6100	\sim		
<u>A</u> ralık: 1	Normal	~	Değer:	▲ ▼
Konum:	Normal 🗸		Değe <u>r</u> :	▲ ▼
🗌 Yazı <u>t</u> ipi	aralığı:	Nokt	a ve y <u>u</u> karıs	ii
OpenType Öz	ellikleri			
Ligatürler:	Yok	~]	
<u>S</u> ayı aralığı:	Varsayılan	~]	
Sayı <u>f</u> ormlar	rı: Varsayılan	~]	
Stil kü <u>m</u> eler	i: Varsayılan	~]	
🗌 Bağlama	a Göre <u>D</u> iğerlerini Kullan			

Bu sayede sadece görsel düzen değil, teknik yapı da standardize olur ve CAT yazılımlarına yüklemeden önce olası tag kirliliğini büyük oranda engellemiş olur.

• Metin rengi Otomatik veya Siyah olarak standartlaştırılır. En küçük ton farklılıkları bile CAT içeriğinde farklı bir tag olarak algılanabilir. Bu da çoğunlukla belge ışığı, kontrast, kıvrım, buruşukluk vb. alanlara gelen metinlerde daha silik görünmesinden kaynaklıdır.

4.7. Sayfa Yapısı ve Düzeni

OCR sonrası belgelerde sayfa yapısı dağınık olabilir. Bu nedenle belgeyi çeviriye hazırlarken aşağıdaki ayarları uygulayın:

- Sayfa boyutunu kontrol edin:
- Düzen sekmesi \rightarrow Boyut \rightarrow A4 (21x29.7 cm) seçeneğini seçin.
- OCR yazılımları bazen özel veya varsayılan olmayan ölçülerde çıktı verir. A4 dışında kalan tüm boyutları A4'e çevirin.
- Kenar boşluklarını belgenin türüne göre ayarlayın:
- Düzen sekmesi \rightarrow Kenar Boşlukları \rightarrow Özel Kenar Boşlukları yolunu izleyin.

- Çeviri belgenin son halinde üst kısımda şirket bilgileri, noter kaşeleri, başlık ve antet öğeleri gibi içeriklerin yer alabilmesi için yeterli alan bırakın.
- Alt kısımda ise yaklaşık 4–5 satırlık bir not, açıklama veya imza alanı kalacak şekilde boşluk ayarlayın.
- Sayfa düzenine göre gereksiz geniş boşlukları daraltarak sayfa ekonomisini koruyun.
- Paragraf girintileri ve aralıklarını temizleyin:
- Paragraf hizalamasını seçin → Giriş sekmesi → Paragraf grubundan girintileri 0 cm, aralıkları önce: 0 nk / sonra: 0-3 nk olarak ayarlayın.
- OCR çıktılarında girintiler, satır sonu boşlukları ve rastgele eklenmiş aralıklar segment kaymalarına yol açabilir.
- Gereksiz sayfa sonlarını ve "Kesme" işaretlerini temizleyin:
- Word'ün "Sayfa Sonu" veya "Kesme" işaretlerini gösteren ¶ görünmez karakterler açıkken kontrol edin.
- Sayfa içinde uygunsuz yerleştirilmiş kesme işaretlerinin ikisini birden seçerek silin.
- Eğer ayrı ayrı silmeyi denerseniz bölüme ait özellikler belgenin öncesine veya sonrasına yansıyıp dağılmaya neden olacaktır.
- Aynı belge içinde hem yatay hem dikey sayfalar gerekiyorsa:
- Düzen sekmesi → Kesmeler sekmesinden ilgili sayfanın başına ve sonuna "Bölüm Sonu" ekleyin.
- Sayfanın yönünü değiştirmek için Yönlendirme → Dikey / Yatay arasında geçiş yapın.
- Bu yöntem özellikle geniş tabloların çevirisinde veya sütunlu resmi belgelerde kullanılır.
- Sütunlu yapılar için de "Bölüm Sonu" kullanın:
- Çok sütunlu metinlerde sadece metne değil, paragrafa da bölüm sonu ekleyin.
- Düzen \rightarrow Sütunlar \rightarrow Daha Fazla Sütun \rightarrow Bu noktadan itibaren uygula şeklinde ayarlayın.
- OCR yazılımları bu işlemleri otomatik yapabilir ancak belgeyi kontrol edip gereksiz bölüm sonlarını kaldırın.

Bu adımlar sayesinde belge yapısında tutarlılık sağlarsınız. Aynı belge içinde farklı hizalama, yönlendirme ve kenar boşlukları varsa çeviri sürecinde metinlerin kayması veya yanlış hizalanması gibi sorunları bu düzenlemelerle önleyin. OCR yazılımları sayfa içinde farklı sayfa özellikleri imkanı sayesinde sayfa dizgisini yaparken genellikle Bölüm Sonu işaretlerini kullanır.

4.8. Resimler ve Şekiller

OCR yazılımları, metin dışı görsel alanları resim olarak tanımlar ve çıktı dosyasına gömebilir. Ancak bu görsellerin belgeye sonradan elle yerleştirilmesi gerekebilir. Bu durumda sayfa düzenini bozmamak ve metin yapısını etkilememek dizgi veya grafik tasarım sürecini hızlandıracaktır.

4.8.1. Görsel Öğelerin Kullanımı: Nerede Yerleştirilir, Nerede Yerleştirilmez?

Belge türüne göre resim, imza, kaşe ve benzeri görsel öğelerin kullanımına ilişkin karar dikkatli verilmelidir. Tercüme yapılan belgenin hukuki bağlayıcılığı, kullanıldığı ortam ve müşteri talebi doğrultusunda aşağıdaki kuralları uygulayın:

Görsel öğelerin yerleştirilebileceği işler:

• Ürün katalogları

- Web sitesi içerikleri
- Tanıtım broşürleri
- Basın bültenleri
- Sosyal medya gönderileri
- Teknik dökümantasyon (görsel açıklamalı, kullanıcı rehberi niteliğinde)

Bu tür belgeler genellikle yerelleştirme, sunum veya marka uyarlaması amacıyla hazırlanır. Yasal bağlayıcılığı yoktur. Müşteri talebi doğrultusunda, imza veya kaşe içermeyen resimler (ürün görselleri, grafikler, diyagramlar, ekran görüntüleri) yerleştirilebilir. Ancak yine de telif haklarına dikkat edin ve resimlerin orijinal kaynağını kontrol edin.

Görsel öğelerin kesinlikle aktarılmaması gereken işler:

- Resmi belgeler (diploma, noter onayı, kimlik, vekaletname, doğum belgesi vb.)
- Hukuki belgeler (mahkeme kararları, yeminli beyanlar, protokoller)
- Eğitim belgeleri (transkript, sertifika, mezuniyet belgesi)
- Sözleşmeler ve beyan formları
- Göçmenlik veya vize başvurularında kullanılan belgeler

Bu tür belgelerde özellikle imza, mühür, kaşe, barkod ve karekod gibi öğelerin kopyalanması sakıncalı durumlar meydana getirebilir. Bu öğelerin dijital ortama taşınması belgede sahtecilik şüphesi oluşturabilir ve çevirmen açısından hukuki sorumluluk doğurabilir.

İmza ve kaşe öğelerine özel dikkat:

- İmza görsellerini kesinlikle yerleştirmeyin.
- Kaşe içerikleri (örneğin noter adı, sicil no, tarih) açıklama notu olarak metne yazılabilir. Örnek: <<KAŞE: İstanbul 23. Noterliği, Sicil No: 123456>>
- Görsel öğenin yerine içerik bilgisi yazı ile ifade edilmeli, gerektiğinde kalın ve büyük harflerle placeholder (yer tutucu) kullanılmalıdır:

<<İMZA>>, <<KAŞE>>, <<FOTOĞRAF>>, <<BARKOD>>, <<KAREKOD>>

Bu yöntemle hem belgenin bütünlüğü korunur hem de çevirmenin yasal sorumluluğu sınırlandırılmış olur. Çeviri alanı ne olursa olsun belgenin bağlamı doğrultusunda görsel öğenin yeri var mı yok mu dikkatle değerlendirin. Her durumda şeffaf ve belgelendirilebilir bir tercih yapın.

4.8.2. Teknik Yerleştirme

OCR sonrası veya manuel olarak belgelere yerleştirilen görsel öğeler, belge düzenini bozabilir. Görsel yerleştirirken bazı teknik ayarların yapılması gerekir.

Görsel ekledikten sonra mutlaka Metin Kaydırma ayarını kontrol edin:

- Resme sağ tıklayın → Metin Kaydırma → Metnin Önüne: Görsel serbest biçimde konumlandırılır, metin akışını etkilemez.
- Metin Kaydırma → Satırla Aynı Hizaya: Görsel, metinle birlikte hareket eder ama bir paragraf gibi davranır.
- Metin Kaydırma → Sıkı / Kare / Üst ve Alt gibi seçenekler yalnızca broşür, katalog gibi özel tasarımlar için uygundur.

Görselin boyutunu ve konumunu sabitleyin:

- Resim Biçimi \rightarrow Konum \rightarrow Sayfaya göre sabitle seçeneğini kullanın.
- Bu ayar özellikle çok sayfalı belgelerde fotoğraf, grafik gibi öğelerin konumunun değişmemesini sağlar.
- Yüksek çözünürlüklü (300 dpi ve üzeri) ve belgeye uygun orantıda görsel kullanın. Aşırı büyük ya da çok küçük görseller belge bütünlüğünü bozar.

4.8.3. Harflerin Tanınmasından Kaynaklı OCR Hataları

OCR yazılımları metni tanırken, görsel veriyi önceden tanımlı karakter sözlükleri ve istatistiksel modellerle karşılaştırarak çözümleme yapar. Bu süreçte karakterlerin şekli, yüksekliği, genişliği, uç yapıları ve çevresindeki boşluklara göre en olası harf tahmin edilir. Ancak belge üzerindeki lekeler, tarama eğikliği, düşük çözünürlük, sıkışık yazı tipleri veya yazıların silik olması gibi nedenlerle birbirine benzeyen harfler yanlış tanınabilir ya da iki ayrı karakter birleşerek yeni bir karakter gibi algılanabilir.

Bu durum özellikle serif yazı tiplerinde (ör. Times New Roman, Bookman, Lucida Bright) yaygındır çünkü bu yazı tiplerinin karakter uçlarında "tırnak" yapıları bulunur. Bu tırnaklar yakın duran harflerin birleşmesine ve hatalı karakter üretimine neden olabilir.

Harf + Harf Birleşmelerine Dayalı Hatalı Tanıma Örnekleri

Aşağıdaki tablo, OCR hatalarının doğrudan harf şekillerinin birbirine karışmasından kaynaklandığı örnekleri gösterir:

Tannş OCK Çiktisi Dirkşen Harner		
samıç	r + n	m
minirnum	$m \rightarrow rn$	rn
tahmat	l + i	h
taaliyet	$f \rightarrow t$	t
öclendi	$d \rightarrow cl$	cl
proieler	$j \rightarrow i$	i
güvenlik	$n \rightarrow i$	11
veliler	r + i	1
	samıç minirnum tahmat taaliyet öclendi proieler güvenlik veliler	samıç $r + n$ minirnum $m \rightarrow rn$ tahmat $l + i$ taaliyet $f \rightarrow t$ öclendi $d \rightarrow cl$ proieler $j \rightarrow i$ güvenlik $n \rightarrow i$ veliler $r + i$

Doğru Yazım Yanlış OCR Çıktısı Birleşen Harfler Ortaya Çıkan Karakter

Bu tür hatalar, kelimenin yapısını bozduğu gibi, CAT araçlarında çeviri belleği eşleşmelerini de olumsuz etkiler. Örnek olarak "talimat" yerine "tahmat" yazılmış bir içerik daha önce çevrilmiş olsa bile eşleşme sağlamaz ve yeniden çevrilmek zorunda kalır. Eğer bu hata sonrasında anlamlı bir kelime ortaya çıkmışsa çevirinin doğruluğu çevirmenin bu konuda dikkatli olup olmamasına bağlıdır. Bu tarz hatalardan dolayı da

- Eşleşen cümlelerde %100 eşleşme (match) yerine kısmi eşleşme (fuzzy match) oluşur,
- Anlamlı ama bağlam dışı kelimeler çeviri hatalarına yol açar.

Dikkat edilmesi gerekenler

• OCR çıktısını çevirmeden önce mutlaka baştan sona bir kez okuyun. Harf benzerliği nedeniyle oluşan sessiz hataları fark etmenin en etkili yolu metni bir bütün olarak gözden geçirmektir.

- Kelime işlemci (Word vb.) tarafından verilen yazım denetimi uyarılarını kontrol edin. Altı kırmızı çizgili kelimelerin çoğu OCR kaynaklı harf hatalarına dayanır. Örnek olarak samıç, belirlendl, sunuldj gibi hatalar Word'de hemen belli olur.
- Anlamlı ama yanlış dönüştürülmüş kelimelere karşı dikkatli olun. Örnek olarak "kazara" yerine "kazam"; "onaylı" yerine "onaya"; "verildi" yerine "veridi" gibi OCR çıktıları yazım hatası uyarısı vermez ama bağlamda anlam kayması yaratır.
- Tarama kalitesi düşükse karakter şekli benzeyen harf çiftlerini (rn/m, li/h, i/l, d/cl vb.) önceden belirleyin. Bu çiftleri içeren kelimeleri listeleyip metin içinde kontrol etmek hata riskini azaltır.
- Kaynak dokümanın bir kopyasını çeviri sırasında bir kenarda hazırda tutun. Görsel referans ile sağlamasını yaparak ilerleme metindeki şüpheli karakterlerde doğrulama sağlar. İmzalı, kaşeli, sayısal veya özel fontla yazılmış belgelerde bu konuya özellikle dikkat edilmelidir.
- Belge taramasında kullanılan yazı tipine dikkat edin; serif (tırnaklı) fontlar varsa hata olasılığı yüksektir. *Times New Roman, Lucida, Georgia* gibi fontlar OCR için risklidir. Bu tarz içerikleri kontrol ederken daha dikkatli olun.

5. CAT TOOL ÖNCESİ DİKKAT EDİLMESİ GEREKENLER

Bilgisayar Destekli Çeviri (CAT) araçları çeviri süreçlerini hızlandırmak, terminolojik tutarlılığı sağlamak ve çeviri belleği kullanımını kolaylaştırmak için tercih edilir. Ancak bu araçlardan verimli sonuç alınabilmesi için kaynak metnin segmentasyon kurallarına uygun biçimde hazırlanması gerekir. OCR çıktıları, kaynak dosya dizgisi ve tablo yapıları doğru şekilde yapılandırılmazsa CAT araçlarında gereksiz etiketler, segment bölünmeleri veya eksik içerikler oluşabilir. CAT Tool entegrasyonu öncesi uygulanması gereken temel adımlara dikkat edilmesi gerekir.

5.1. Biçimlendirme İşaretlerine Dayalı Segmentasyon

CAT araçları, çevrilecek metni segment adı verilen parçalara ayırarak işler. Bu segmentler genellikle cümle düzeyinde olur ancak segmentlerin sınırlarını belirleyen asıl unsur biçimlendirme işaretleridir. Her dosya formatı, paragraf veya satır sonunu farklı bir şekilde tanımlar:

- Word belgelerinde: Paragraf sonu (¶) ve manuel satır sonu (¬)
- HTML formatında: ,
> gibi etiketler
- XML/JSON formatlarında: öğe içi boşluklar, satır yapısı ve veri kapsayıcıları
- TXT veya CSV dosyalarında: satır sonu karakterleri (LF, CRLF)

Buradaki paragraf kavramı anlam bütünlüğü olan metin bloğu değil, yazılım açısından tanımlanmış teknik bir sınırdır. Word belgesinde her Enter tuşu paragraf (¶) oluşturur ve bu da CAT aracında yeni bir segment anlamına gelir.

Cümle Ortasında Paragraf Sonu

CAT aracı, paragraf sonunu segment sonu olarak işler. Bu nedenle, anlam olarak devam eden bir cümle içinde kullanılan gereksiz Enter (yani paragraf sonu), segmentin ikiye bölünmesine neden olur. Aynı şekilde sayfa sonunda "Sayfa Sonu" veya "Bölüm Sonu" ile bölünen cümleler de ayrı segmentler olarak yer alır.

• Yanlış:

Başvurunuz incelenmiş,¶ değerlendirmeye alınmıştır.¶

• Doğru:

Başvurunuz incelenmiş, değerlendirmeye alınmıştır.¶ veya satır sonu ile cümle bütünlüğü koruyarak ilerleyebilirsiniz; Başvurunuz incelenmiş,J değerlendirmeye alınmıştır.¶

Bu gibi bölünmeler:

- Segment sayısını gereksiz yere artırır
- Çeviri belleğinde parçalı kayıt oluşmasına neden olur
- Tekrar analizlerini bozar
- Tutarlılık sorunları yaratır

Yanlış Satır Sonları

Tersine, anlam bütünlüğü olan listeler tek paragrafa sıkıştırılmışsa CAT aracı bu yapıyı tek segment olarak algılar ve çevirmen açısından segment içinde karışıklığa neden olabilir. Özellikle uzun listeler veya çoklu cümle yapıları segment içinde boğulabilir.

 Örnek: Lütfen aşağıdaki belgeleri ibraz edin: J Kimlik fotokopisi, J ikamet belgesi, J banka dekontuJ
 2 adet vesikalık fotoğraf.¶

Yukarıdaki örnek satır sonu sebebiyle tek cümle olarak algılanarak tek segment olarak eklenir ama her satır arasında etiket olacak şekilde görünür. OCR araçlarını düzeni korumak için satır sonlarını tutmak amacıyla bu işareti aktif şekilde kullanır.

 Doğru kullanım: Lütfen aşağıdaki belgeleri ibraz edin:¶ Kimlik fotokopisi¶ İkamet belgesi¶ Banka dekontu¶ 2 adet vesikalık fotoğraf¶

Bu şekilde segmentler hem okunabilir hem çeviri belleği açısından faydalı hale gelir.

Bu aşamada Word'de dizgi ile segmentasyonu temizlemek için:

- Giriş \rightarrow ¶ (Biçimlendirme İşaretlerini Göster) seçeneğini açın.
- Ctrl + H (Bul ve Değiştir) ile ^p (paragraf sonu) ve ^l (satır sonu) karakterlerini bulun.
- Cümle ortasında paragraf sonu varsa kaldırın.
- Liste yapıları segment olacak şekilde bölünmüş mü kontrol edin.

5.2. Metne Dayalı Segmentasyon Kuralları

CAT araçları segment sınırlarını yalnızca biçimsel işaretlere göre değil, dilsel yapıya, noktalama işaretlerine ve cümle bütünlüğüne göre de belirler. Segmentasyonun doğru yapılabilmesi için çeviri öncesi metnin dilbilgisi yapısına uygun şekilde düzenlenmiş olması gerekir.

Segment ayrımları genellikle şu işaretler üzerinden yapılır:

- Nokta (.)
- Soru işareti (?)
- Ünlem (!)
- İki nokta (:) (kısıtlı olarak)
- Satır sonu karakterleri
- Parantez, tırnak, tire

gibi bağlayıcı işaretler (doğru kapatılmadığında segmenti bölebilir)

5.2.1. Uzun Cümlelerin Bölünmemesi

Dil yapısı gereği uzun cümleler tek segment hâlinde kalabilir. Ancak bu segment çevirmen açısından ağır, karmaşık ve düzensiz olabilir. Özellikle mahkeme kararları, sözleşmeler vb. hukuki metinler uzun cümlelerden oluşur, bu tarz yapılar da CAT aracında yoğun segmentlere neden olur.

• Örnek:

Başvuru sahibi tarafından doldurulmuş olan şikayet formu ilgili müdürlüğe teslim edildikten sonra başvuru evrakı eksiksiz olduğu takdirde işlem süreci belirlenen itiraz sürelerinin ardından başlatılarak gerekli süreç müdürlükçe takip edilerek idari soruşturma yürütülecektir.¶

Bu tür cümleleri bölmek için CAT aracında segment kesme kuralları (örnek olarak "ve", "çünkü", "ancak" gibi bağlaçlardan sonra) özelleştirilebilir veya çevirmen segmenti manuel olarak bölebilir. Ayrıca OCR sonrası teknik düzenlemeler yapılırken anlamı bozmayacak şekilde kaynak cümlelere müdahale edilebilir, bu şekilde de cümle parçalama yoluyla segmentler daha basit cümlelere ayrılabilir. CAT sonrası çıkan çeviri kaynak ile anlam ve yapı bütünlüğünü koruduğu sürece çeviri sürecindeki vakit kaybını önleyecektir.

5.2.2. Noktalamadan Sonra Boşluk Olmamasından Dolayı Segment Hataları

Bazı CAT araçları cümle sonu işareti + boşluk kombinasyonunu segment ayrımı olarak kabul eder.

• Yanlış:

İmzanız eksiktir.Belge geçersiz sayılacaktır.¶

 Doğru: İmzanız eksiktir. Belge geçersiz sayılacaktır.¶

Bu gibi durumlarda segment bölünmeyebilir ve tüm cümleler bir araya toplanır. Noktalama işaretlerinden sonra boşluk bırakıldığından emin olun. OCR sürecinde bu tarz metinler sayfa görsel bozukluklarından dolayı ortaya çıkabilir, sayfa içindeki lekeler fazladan noktalama işareti olarak algılanmış olabilir.

5.2.3. Tırnaklar, Parantezler ve Düzgün Kapanmayan Yapılar

Açılıp kapanmayan tırnak işaretleri (" "), tek tırnak (' '), parantezler (), [] veya { } segmentasyon algoritmasını şaşırtabilir. CAT araçları, kapanış eksikse bu karakteri segmentin devamı sanabilir veya segmenti sonlandırmayabilir ya da kodlama olarak algılayabilir. Günümüzdeki araçlar bunların önüne büyük oranda geçmiş durumda ama OCR kalitesinin kötü olması durumunda bu işaretlerin diğeriyle karışması sık yaşanan bir durumdur. Mesela bir parantez işaretinin üzerinki bir leke, silinme veya kırışıklık ")" işaretinin "}" şeklinde algılanmasına neden olabilir.

• Örnek (Yanlış):

Başvuranın onay alanına elle "Okudum, anladım, onaylıyorum. yazması gerekmektedir.¶

• Doğru:

Başvuranın onay alanına elle "Okudum, anladım, onaylıyorum." yazması gerekmektedir.¶

Bu örnekte yanlış olan versiyonda CAT aracı nokta (.) işaretinden dolayı 2 farklı segment oluşturacaktır. Ayrıca çeviride düzeltilse bile fazladan "işareti QA hatası olarak görünecektir.

5.2.4. Noktalama İşaretinin Yanlış Tanınmasından Kaynaklanan OCR Hataları

OCR işlemlerinde karşılaşılan yaygın sorunlardan biri de noktalama işaretlerinin yanlış tanınması, atlanması ya da yanlış karakterlerle karıştırılmasıdır. Özellikle tarama kalitesi düşük olan belgelerde nokta (.), virgül (,), iki nokta (:), noktalı virgül (;) ve ünlem işareti (!) gibi karakterler yeterince net algılanamayabilir. Bu durum, CAT araçlarının segmentasyon kurallarını yanlış uygulamasına yol açar. Örnek olarak cümle sonundaki bir nokta okunamazsa o cümle bir sonrakiyle aynı segmentte kalır; bu da hem çeviri belleği eşleşmelerini bozar hem de segmentlerin anlam bütünlüğünü zedeler. Bazı durumlarda nokta veya virgül bir önceki harfle birleşebilir ya da noktalı virgül "1, j" gibi harflerle karıştırılarak metne hatalı karakter olarak yansıtılır. Benzer olarak virgül gibi devam ettirici bir işaret nokta gibi algılanırsa segment gereksiz yere bölünebilir. Noktalama karakterlerinin bu şekilde yanlış tanınması yalnızca segmentasyonu değil, aynı zamanda çevirinin doğruluğunu ve tekrar oranlarını da ciddi biçimde etkiler. Bu nedenle OCR sonrası kontrollerde özellikle noktalama işaretlerinin eksiksiz, doğru yerde ve doğru karakterle yer alıp almadığı dikkatle gözden geçirilmelidir.

Doğru Yazım	Yanlış OCR Çıktısı	Yanlış Algılanan Karakter	Ortaya Çıkan	Açıklama
alındı.	alındıı	1+.	11	Nokta tanınmaz, segment kapanmaz
tamamlandı.	tamamlandı,	•	,	Segment bir sonraki cümleyle birleşerek bütün hale gelir
bildirildi,	bildirildi.	,	•	Virgül nokta olarak algılanıp segment erken kapanabilir
karar verildi;	karar verildiı	;	1	Noktalı virgül tanınmaz, harfe dönüşebilir, segment birleşir
belgeler sunuldu:	belgeler sunuldui	:	1	İki nokta karakteri tanınmaz, harfe dönüşebilir
açıklandı!	açıklandıı	!	1	Ünlem işareti "1" olarak okunabilir, segment kapanmaz
aktarıldı.	aktarıldu	1+.	u	"1." birleşimi "u" algılanabilir, segment kapanmaz
tamamlandı;	tamamlandu	1+;	u	Noktalı virgül algılanamaz, sesli harfe dönüşür
gönderildi.	gönderildj	i+.	j	Nokta yukarı çekilmiş, "j" gibi görünür

Dikkat edilmesi gerekenler:

- Cümle sonlarında nokta (.) karakterinin gerçekten bulunduğundan ve doğru tanındığından emin olun. Noktanın harfle birleşmesi veya harfe dönüşmesi CAT aracının segment sonunu algılayamamasına neden olur.
- Virgül (,), iki nokta (:), noktalı virgül (;) ve ünlem (!) gibi işaretlerin yanlış karakterlerle karışıp karışmadığını kontrol edin. Özellikle virgül yerine nokta, iki nokta yerine "i", ünlem yerine "l" ya da "1" gibi yanlış dönüşümlerde segment doğru şekilde bölünmez.
- Yazım denetiminde noktalama ile birleşmiş yanlış karakterler için özel dikkat gösterin. Örnek olarak "sunuldu.", "belirlendi." gibi kullanımlar "sunuldj", "belirlendl" şeklinde yanlış algılanmış olabilir.

- Cümle sonunda nokta görünmüyorsa segment kapanmamış olabilir; bu tür durumlarda ilgili satırı manuel olarak düzeltin. Bazı CAT yazılımlarında segment bölme özelliği bulunmakta, el ile ayırın.
- OCR çıktısını CAT aracına aktarmadan önce cümle sonlarının doğru işaretlendiğini test segmentasyonu ile kontrol edin. Özellikle nokta içermeyen cümle sonlarında birleşik segment olup olmayacağına göre metni düzenleyin.

5.2.5. Liste Biçimleri (1., a., i., *, - gibi)

"1.", "a." gibi liste başlıkları cümle sonu sanılabilir. Bu da segmenti bölmeye neden olur. Özellikle sayıyla başlayan satırlar için segmentasyon kuralına uymayan yapılar manuel kontrol gerektirir. Bu tarz listelemeler genelde Word gibi kelime işlemci yazılımlarında Numaralandırma özelliği altında yazılım kodu olduğundan CAT içinde doğrudan görünmez veya etiket olarak yer alır ancak metin içinde yer alan atıflarda segmentasyon bozulabilir. Günümüz CAT araçlarının çoğunda bu tarz durumlara karşı segment birleştirme özelliği bulunmaktadır.

Ne Yapılabilir?

- Segmentasyon kurallarını CAT aracının içinden düzenleyin (ör: MemoQ'da Regex tabanlı segmentation rules, Trados'ta SRX dosyası).
- Tüm özel karakterleri içeren segmentleri test dosyasıyla önceden işleyin.
- Gerekirse anlam bütünlüğüne göre segmentleri manuel bölün veya birleştirin.

5.3. HTML/Markdown Etiketleri veya Kod Fragmanları

,
,
, , <code> gibi HTML etiketleri, segmentasyonu doğrudan etkiler. Kapanış etiketi eksikse veya iç içe yapı varsa segmentler ya alınmaz ya da bozulur. Ayrıca bu tarz etiketlerin tanınması araca göre değişir. Günümüzdeki yazılımların çoğu bu etiketleri tanıyıp işleyebilse de eksik bir etiket segmentasyonu etkiler. Bu durum ayrıca basılı teknik doküman taramalarının bu etiketleri anlattığı örneklerde metin yerine kod olarak algılanmasıyla farklı sorunlara neden olabilir. Tabii ki bu tarz işler çok düşük ihtimalle gelir ama yine de ihtimal dahilindedir.

- Yanlış:
 - Başvuru yapılmıştır.

 - (kapanmamışsa segment kapanmayabilir.)
- Doğru:
 Başvuru yapılmıştır.

5.4. CAT Araçlarında Segmentasyon Kuralları ve Ayarlamalar

Farklı CAT araçlarında segmentasyonu belirlemek için özel kurallar belirleyebilirsiniz. Segmentasyon hatalarını en aza indirmek için aşağıdaki ayarlara dikkat etmelidir:

• **MemoQ:** Segmentation Rules ayarlarından paragraf bazlı segmentasyonu özelleştirme imkanı sunar. Kısaltmalara bağlı yanlış segmentasyonu önlemek için Language-Specific Segmentation Rules kullanılabilir.

- **SDL Trados Studio:** SRX (Segmentation Rules Exchange) dosyaları ile paragraf ve cümle bölünmelerini yönetebilir. Ayrıca, Translation Memory Settings → Segmentation Rules bölümünden özel kurallar eklenebilir.
- **Phrase:** Dil bazlı segmentasyon seçenekleri sunarak belirli diller için özelleştirilmiş ayarlar yapılmasına izin verir. Örnek olarak Japonca, Arapça ve Çincede noktalama işaretlerinin tanımlanması segmentasyon hatalarını önleyebilir.
- Özellikle cümle sonu işaretleri, özel karakterler ve kısaltmaların segmentasyona etkisi dikkate alınarak her araç için dil bazlı kurallar ayarlanmalıdır. Segmentasyon hatalarının çeviri belleğinde tekrar eden hatalara yol açabileceği unutulmamalıdır.

5.4.1. Diller Arası Farklılıklar

Çeviri sürecinde kullanılan dillerin yapısal farklılıkları, segmentasyon hatalarına neden olabilir. Özellikle farklı dil ailelerinden gelen dillerin kelime sırası, kısaltmalar ve noktalama işaretleri, CAT araçlarının segmentleme algoritmalarını doğrudan etkileyebilir.

Özel Dil Yapıları:

- Japonca, Çince gibi dillerde noktalama işaretleri cümle ayrımında farklı çalışabilir.
- Almanca gibi bileşik kelimeler içeren dillerde, kelime segmentasyonu dikkatle yapılmalıdır.

Kısaltmalar:

- İngilizce'de "Dr., Mr., Ms." gibi kısaltmalar yanlış segmentasyon oluşturabilir.
- Fransızca'da "Mlle., M." gibi kısaltmalar, nokta içerdiğinden yanlış ayrılabilir.

Kısaltmalar, sonundaki nokta nedeniyle CAT aracında cümle sonu gibi algılanabilir. Bu da segmentin gereksiz yere ikiye bölünmesine yol açar.

• Örnek:

Randevunuz, Dr. Ahmet Demir tarafından oluşturulmuştur.¶

Bu tür kısaltmaların ardından gelen kelimenin büyük harfle başlaması segmentasyon hatasını daha da artırabilir. Bu durumda segment kurallarına Dr., Av., Sn., Mr. gibi kısaltmaları tanıtmak gerekir.

6. ÇEVİRİ SONRASI GENEL KONTROL

Çeviri işlemi tamamlandıktan sonra yapılan nihai belge kontrolü, yalnızca dilsel doğruluğu değil, aynı zamanda belgenin dizgisel bütünlüğünü, görsel uyumunu ve kullanım kolaylığını da kapsar. Özellikle OCR'den geçmiş belgelerde, görsel bozulmalar, metin kutularında kayıplar veya format değişimleri sık yaşanır. Aşağıdaki üç başlık, bu sürecin planlı ve sistematik yürütülmesini sağlar.

6.1. Belge Dizgisinin (Sayfa Tasarımının) Gözden Geçirilmesi

6.1.1. Görsel Karşılaştırma ile Kontrol

- Kaynak dosyayı bir kenarda açık tutun ve çevrilmiş belgeyi sayfa sayfa karşılaştırarak ilerleyin.
- Tabloların, başlıkların ve sayfa düzeninin aynı sırada ve aynı hiyerarşide yer aldığından emin olun.
- Sayfa yönü (dikey/yatay), hizalama ve boşluk yapılarının kaynak belgeyle uyumlu olup olmadığını kontrol edin.

Çevirinin doğası gereği birebir yerleşmek zorunda değil ama hataların hızlı tespitine yardımcı olur.

6.1.2. Eksik veya Taşmış İçeriklerin Tespiti

- Tablo içlerinde, metin kutularında veya sınırlı alanlarda kalan metinlerin tamamının görünür olduğundan emin olun.
- Paragraf (¶) işaretlerini açarak hücre veya kutu sonlarında içerik sarkması olup olmadığını kontrol edin.
- Sayfa sonlarında taşan içerik nedeniyle oluşan fazladan boş sayfaları veya yarım kalmış segmentleri kontrol edin.

6.1.3. Biçimlendirme ve Stil Tutarlılığı

- Başlıkların aynı düzende ve formatta olup olmadığını kontrol edin.
- Alt başlıkların stil açısından tutarlı olup olmadığını kontrol edin.
- Numaralı veya madde işaretli listelerin sıralamasında kayma olup olmadığını kontrol edin.
- Tabloların hizasını, satır yüksekliklerini ve sütun genişliğinin sayfa kenarlıklarıyla uyumluluğunu kontrol edin.

6.2. Metin Hatalarının Kontrolü

6.2.1. OCR Kaynaklı Hatalardan Dolayı Oluşan Çeviri Hataları

- Anlam bütünlüğüne uymayan cümleler veya kelimeler olup olmadığını kontrol edin. OCR sırasında görünmeyen veya hatalı tanınan karakterler nedeniyle yanlış çevrilmiş kelimeler veya cümleler olabilir.
- Düz cümlelerin soru veya ünlem cümlesi olarak ya da tam tersi şekilde çevrilip çevrilmediğini kontrol edin.
- Kaynak metinde arka arkaya cümlelerin çeviride birleşip birleşmediğini veya birleşik yapıların ayrılıp ayrılmadığını kontrol edin.

- Görselden alınmış tarih, sayı, özel ad gibi ifadelerin doğru şekilde çeviride yer alıp almadığını kontrol edin. Özellikle resmi belge çevirilerinde belgede kabul edilemez hatalar arasında yer alır.
- Başlıklar, liste maddeleri veya paragrafların içerik olarak farklılaşıp farklılaşmadığını kontrol edin.
- Yanlış çevrilmiş ancak otomatik sistemlerce fark edilmeyen özel terim ya da sabit ifadeleri kontrol edin.

6.2.2. Yazım ve Dilbilgisi Denetimi

- Word uygulamasının yazım denetimi ve dilbilgisi araçlarını çalıştırın.
- Aksanlı karakterler, özel harfler (ş, ç, ğ vb.) ve dil ayarlarının doğru tanındığından emin olun.
- Gözden kaçan küçük hataları yakalamak için belgeyi son bir kere okumak faydalı olabilir.

6.3. CAT Sonrası Kontrol

Çevrilen dosyanın dışa aktarımı sırasında, özellikle Word gibi hedef formatlara dönüşüm yapılırken etiket sıralamaları, biçimlendirme yapıları, görünmeyen karakterler ve sınırları bozulabilir. Bu nedenle çeviri tamamlandıktan temel kontroller yapılmalıdır.

6.3.1. Etiket Hataları ve Sıralama Sorunları

- Yanlış yerleştirilmiş etiketleri kontrol edin. Bazı durumlarda CAT kullanım tekniklerinden dolayı etiketler doğru aktarılmayabilir, bu da yanlış şekillendirilmiş metinler olarak görünür.
- Etiket sıralamaları bozulmuşsa düzeltin. Etiketlerin iç içe geçmesi veya yanlış sırayla kapanması (örnek olarak <i>metin</i>) belge yapısını ve görünümünü bozabilir. Bu yüzden de bu tarz durumlar kullanılan yazılımda QA hatası olarak işaretlenir.

6.3.2. Kaymalar ve Format Bozulmaları

- Birleştirilmiş veya bölünmüş segmentlerden çıkan metinleri gözden geçirin. Özellikle kısa segmentlerden çıkan metinler veya liste öğeleri birbirine bağlanmış olabilir.
- Segmentin başında veya sonunda fazladan boşluk, boş satır ya da karakterler dosya çıktısında dizgiyi etkileyebilir. Çift boşlukları veya boşluksuz yerleşmiş çevirileri kontrol edin.
- Numaralandırılmış listeler bozulmuş olabilir. Çıktı alınırken otomatik liste formatı devre dışı kalmış ya da sıradan sapmalar oluşmuş olabilir.

6.3.3. Dil Kodları ve Biçimlendirme Uyumsuzlukları

- Hedef segmentte farklı dil kodu atanıp atanmadığını kontrol edin. Bazen CAT projesi açarken segmentin dili yanlış tanımlanmış olabilir, bu da yazım denetimini etkiler.
- Başlık stilleri, madde stilleri gibi biçimlendirmelerin hedef dosyada uygulanmış olup olmadığını kontrol edin.
- Stiller diğer içeriğe taşmış olabilir. Başlık veya kalın yazı stillerinin yanlış etiket kullanımından veya hiç etiket kullanılmamasından dolayı önceki veya sonraki cümleye uygulanmadığından emin olun.

6.3.4. Tablo ve Alan Bozulmaları

- Tabloların hücre içeriği CAT aracında ayrı segmentlere bölündüğü için bazen bağlamdan kaynaklı çeviri hataları olabilir, Word'e aktarıldığınızda tekrar kontrol edin.
- Özellikle, hücre içinde kaybolmuş içerik CAT aracında yine de çevrilecek öğe olarak görünür. Bu yüzden kaynak dosyada görünmeyen ama CAT üzerinde görünen içerikleri arayın.