

## Dil Kodlamasında Tek Bir Standarda Doğru; UNICODE

Sezar krallığını tüm Dünya'ya yayabilseydi, Osmanlı'nın yerküre üzerindeki hakimiyeti hala devam edebilseydi, belki de şu anda Unicode hakkında bir yazı yazmaya gerek kalmazdı. Aslında Unicode gibi bir standarda ihtiyaç da kalmazdı. Bilim dili olma yolunda önemli mesafeler kaydeden İngilizce'nin Dünya üzerindeki yaygınlığı bile dil kodlamasındaki karmaşayı engelleyemedi.

Özellikle İnternet'in yaşamın tüm alanlarındaki kullanımı ve mesafelerin ortadan kalkması, ulusal dillerinde yayın yapan siteler arasında dil kodlamalarında çakışmalara sebep oldu.

İnternet'in ilk yıllarında, içeriğin bu günkü kadar iştah kabartmadığı dönemlerde, dil kodlaması bir sorun gibi gözükmemekteydi. Ulusal dilde yayın yapan sitelerin azlığı, uluslararası firmaların genelde yayın dili olarak İngilizce'yi tercih etmeleri, dil kodlaması ile ilgili problemleri o günlerde gözardı etmemize neden oldu.

İçeriğin kontrol edilemez bir hızla büyümesi, ortamlar arası değişik formatta verinin kullanım ihtiyacı, her dil yapısı ile ilgili o yapıya özgü bir karakter tanımlama dizisinin ortaya çıkmasına neden oldu. Tarayıcılarımızın dil kodlama ayarlarından aşına olduğumuz, Turkish(Windows), Turkish(ISO), Western European (Windows), Greek(ISO), Chinese Traditional (Big5) ve daha ismini sayamadığımız onlarca farklı kodlama, web üzerinde sayfaların doğru görüntülenebilmesi için kullanılmaya başlandı. Her sayfanın kod bölümüne tarayıcının uygun kodlamaya göre gösterim yapabilmesini sağlayan etiketler yerleştirildi.

Her dilin farklı kodlamaya sahip olması, aynı sayfada birden fazla dil içeren sitelerde ve siteden siteye geçişlerde kod çakışmalarına neden olmaya başladı. İki farklı dil kodlamasında kullanılan aynı numara, farklı karakterleri tanımlayabiliyordu. "ı" yerine "y" işaretinin çıkması, "ğ" yerine "?" "ş" yerine de "?" işaretinin çıkması hep bu kod çakışmalarından kaynaklanmaktaydı. Western European'da "y" karakterini tanımlayan nümerik değer ile Türkçe'deki "ı" karakteri tanımlayan nümerik değer aynı olması yaşanan problemlerin nedenini ortaya koymaktaydı.

Tüm bu sorunlara "her karakter için sabit (unique) bir numara" yaklaşımıyla çözüm bulundu ve Unicode ortaya çıktı. Tüm dil kodlamaları Unicode çatısı altında toplanmaya başlandı ve karmaşık kodlama denizinde tek bir fener tüm dillere yol gösterir oldu.

Unicode ile birlikte, platformdan bağımsız, programdan bağımsız, dilden bağımsız bir yapı kurulmuş oldu. Bu yapıdan veri sistemleri de nasibini aldı. Veritabanı üreticileri sistemlerini Unicode'u destekler hale getirmeye başladılar. MS SQL ve Oracle bu desteği vermeye başlarken, My SQL gibi açık kodlu sistemlerin de yeni versiyonlarında Unicode'u destekleyecekleri açıklandı.

Yazılım devleri de Unicode'a hızlı ve sert bir geçiş yaptılar. Bu geçişten Macromedia kullanıcıları da nasiplerini aldılar. ColdFusion, Flash ve Dreamweaver üçgeni arasında XML ile başlayan Standartlaşma rüzgarı Unicode ile devam etti. Flash içerisinde dışarıdan metin okuma fonksiyonundan ColdFusion'daki SQL veritabanı bağlantısına kadar her alandaki metin yapısı Unicode üzerine kuruldu.

Ortak bir dile geçemeyen insanoğlu, ortak dil kodlamasıyla en azından bilişim teknolojilerinde belirli bir dil standarda geçmeyi başardı.

R. Orçun Madran