

T.C. GELİR İDARESİ BAŞKANLIĞI

YAPAY ZEKÂ DESTEKLİ E-ÖZELGE PROJESİ

Proje Amacı:

Gelir İdaresi Başkanlığı eski adıyla Gelirler Genel Müdürlüğü, vergiler ve diğer gelirleri tahsil etmekle sorumlu olan, Türkiye Cumhuriyeti Hazine ve Maliye Bakanlığına bağlı bir kurumdur.

Gelir İdaresi Başkanlığının Misyonu: Mükellef haklarını gözeterek vergide gönüllü uyumu artırmak ve kaliteli hizmet sunarak vergi ve diğer gelirleri toplamaktır.

Gelir İdaresi Başkanlığının Vizyonu: Ekonomik aktiviteleri kavrayarak vergi ve diğer gelirleri tam ve zamanında toplayan, kayıtlı ekonomiyi teşvik eden, yetkin beşerî sermayesiyle kaliteli hizmet sunan, mükellef odaklı, yenilikçi ve katılımcı örnek bir idare olmaktır.

Hazine ve Maliye Bakanlığının bağlı kuruluşu olan ve genel bütçeli idare kapsamında olan Başkanlığımız, merkez teşkilatı ve doğrudan merkeze bağlı taşra teşkilatı olarak örgütlenmiştir. Taşra teşkilatı kapsamında 81 il defterdarlığı faaliyet göstermektedir.

Mükellefler ve vergi sorumluları, vergi durumları ve vergi uygulamaları bakımından kendileri açısından açık olmayan, tereddüt ettikleri hususlar hakkında Gelir İdaresi Başkanlığı ve yetkili kıldığı makamlardan yazı ile izahat isteme hakkına sahiptirler. Yetkili makamlar da mükelleflerin yazı ile yapacakları taleplere yine yazıyla cevap vermek durumundadırlar.

Bu kapsamda Defterdarlıklarca mükelleflerden alınan özelge talepleri idarece incelenir. Eğer talep özelge kapsamındaysa öncelikle e-Özelge sistemi içerisinde yer alan Özelge Havuzundan (daha önce verilmiş mevcut özelgelerin olduğu havuz) vergi, kanun, kanun maddesi vb. kriterlere göre sorgulama yapılarak benzer özelge varsa bunlar emsal alınarak özelge hazırlanır ve e-imzalı olarak mükellefe tebliğ edilir.

Gelir İdaresi Başkanlığı'nın e-Özelge süreçlerini hızlandırmak, insan müdahalesine olan bağımlılığı azaltmak ve doğru, ihtiyaca uygun özelge hazırlanmasını sağlamak için **Yapay Zekâ Destekli** teknolojilerin entegrasyonunun sağlanması hedeflenmektedir. Bu proje, özelge taleplerinin daha hızlı ve doğru şekilde karşılanmasını, özelge havuzunun etkin bir şekilde kullanılmasını ve vergi uygulamalarına ilişkin mükellef tereddütlerinin en aza indirilmesini amaçlamaktadır.

Proje Kapsamı:

1. Özelge havuzundaki mevcut verilerin Yapay Zekâ ile analiz edilmesi ve benzer özelgelerin tespit edilmesi.
2. Özelge taleplerine doğru ve hızlı yanıt verecek bir otomasyon altyapısının oluşturulması.
3. Defterdarlıklar tarafından yürütülen manuel sorgulama süreçlerinin otomatikleştirilmesi.
4. Gelir İdaresi Başkanlığı bünyesindeki Özelge Komisyonu ile süreçlerin dijitalleştirilmesi ve entegrasyonun sağlanması.
5. Özelge havuzundaki mevcut verilerin, talep edilen özelgelerle eşleştirilmesi için yapay zekâ tabanlı bir sistem geliştirilmesi ile bu sistemin, mükellef taleplerine uygun ve doğru özelge önerileri sunması.

6. Projenin sonunda, mükellef memnuniyetini artıracak ve vergi uyumunu teşvik edecek bir sistemin devreye alınması.

Proje Çıktıları:

- Özelge taleplerinin değerlendirilme süresinin azaltılması.
- Özelge havuzunun kullanım oranının artırılması (%85 doğruluk oranı).
- Mükellef memnuniyetinin ölçülebilir şekilde iyileştirilmesi.
- Tüm süreçlerin %100 dijitalleştirilmesi ve insan müdahalesine olan bağımlılığın azaltılması.

Mevcut Durum

Gelir İdaresi Başkanlığı taşra birimi olan defterdarlıklar tarafından özelge talepleri alındıktan sonra, öncelikle bu konuda daha önce verilmiş bir özelgenin olup olmadığı havuzdan manuel olarak sorgulanmaktadır. Sorgulama sonucunda ihtiyacı karşılayacak bir örnek bulunmaz ise gelen talep yeni bir özelge düzenlenmesi için Gelir İdaresi Başkanlığı bünyesinde bulunan Özelge Komisyonuna gönderilmektedir.

Mevcut durumda Özelge Havuzunda çekilen veriler memur eliyle yapıldığından talep kapsamındaki özelgelere emsal teşkil edecek verilerin tamamının bulunması mümkün olmamaktadır. Bu işlemlerin insan eli değmeden yapay zekâ ile yapılması sistem taramasının ihtiyacı karşılar nitelikte ve doğru olmasını sağlayacaktır.

Teknik İsterler:

1. Geliştirilecek çözüm, veri tabanlarından (Postgresql ve Sybase) gerekli verileri sağlayabilecek ve İDARE'nin uygun gördüğü veri tabanlarına sonuçları aktarabilecek şekilde tasarlanacaktır.
2. İDARE'ye ait 17.534 ana özelge ve bunlara bağlı 141.721 üzerinde alt özelge metni kullanılarak doğal dil işleme algoritmalarının geliştirilmesi gerekmektedir. Özelge talebi mevcut havuzdaki özelge talebi ile eşleşiyor ve özelge günceliğini kaybetmemiş ise yapay zekâ çözümü mevcut özelgeyi emsal olarak özelge metni geliştirecektir.
3. 2009 yılından bu yana yıllık olarak cevaplanan özelge sayısında her yıl ortalama %30-%40 oranında artışlar gözlemlenmiştir. Toplam veri boyutu 2 GB'tır. Proje kapsamında kullanılacak veriler, gerekli güvenlik sözleşmeleri imzalandıktan sonra İDARE tarafından sağlanacaktır.
4. Proje süresince geliştirme ortamı İDARE tarafından sağlanan ortamlarda ve İDARE lokasyonunda gerçekleştirilecektir.
5. Canlı ortama geçildiğinde verilerin aktarımı ve nerde tutulacağı İDARE ve YÜKLENİCİ (Proje Yürütücüsü Kuruluş) tarafından ortak çalışmalarda belirlenecektir.
6. Geliştirilecek çözüm kanun/konu bazında (VUK, KVK, GVK, KDV, ÖTV, HARÇ, DAMGA, Hazine Payı (406 S.K.), DHV, Diğer) uzmanlığa sahip olacak, birbirine benzer özelge konularının ayrımını yapabilecektir.

7. Geliştirecek çözüm, talep edilen konunun özelge havuzunda olan bir konuyla ilgili olması durumunda doğal dil işleme algoritmalarıyla kullanıcıya benzer özelgeleri sunacaktır.
8. Sisteme yeni eklenen özelgelerin özelge havuzuna eklenerek modelin geliştirilmesi sağlanacaktır.
9. Geliştirecek çözüm, talep edilen konunun özelge havuzunda olmayan bir konuyla ilgili olması durumunda kullanıcıya büyük dil modelleri (LLM) kullanılarak taslak bir öneri dokümanı hazırlayacaktır. Kullanıcılar öneri dokümanına müdahale edebileceklerdir.
10. Geliştirilecek çözüm, üretilen öneri dokümanın istenen düzeyde olmaması durumunda yeniden cevap oluşturma fonksiyonuna sahip olacaktır.
11. Model tarafından üretilen öneri özelge metinleri, cevapları kontrol eden memur tarafından belirli bir skala ile puanlanabilecektir ve daha sonrasında modelin iyileştirilmesi için kullanılabilir.
12. Kullanıcıların düzenlediği öneri dokümanları modelin geliştirilmesinde kullanılabilir.
13. Geliştirecek çözüm, talep edilen konunun özelge kapsamında olmaması durumunu tespit edebilir.
14. Geliştirilecek çözüm, özelge oluşturmada daha iyi performans gösterebilmesi için İDARE'nin sağladığı veriler kullanılarak mevcut açık kaynak modeller ince ayar (fine-tune) ile tekrar eğitilecektir.
15. Türkçe dilinde en iyi yanıtları veren modeli belirlemek için karşılaştırmalı performans testleri (benchmark) yapılacak ve raporlanacaktır.
16. Model seçimine İDARE ile birlikte karar verilecektir. Test sonuçlarına göre en uygun model seçilerek sisteme entegre edilecektir.
17. Geliştirecek çözüm Türkçe dilini anlama ve Türkçe dilinde içerik üretme yeteneği en yüksek olan açık kaynak büyük dil modelleri (LLM) kullanılacaktır.
18. Geliştirilecek çözümde, Bilgi Getirme Destekli Üretim (Retrieval-Augmented Generation) yaklaşımı kullanılacak ve özelgeyle ilgili içeriklerin doğru, güncel ve güvenilir biçimde sunulması sağlanacaktır.
19. YÜKLENİCİ, Bilgi Getirme Destekli Üretim (Retrieval-Augmented Generation - RAG) mimarisinin özelge belgelerini verimli bir şekilde işleyebilmesi için içerikleri anlamlı parçalara bölecek (chunking) ve bu parçaları metin verilerinin vektör temsillerine dönüştürerek vektör deposuna aktaracaktır.
20. İDARE, proje kapsamında kullanılacak vergi mevzuatı kaynaklarını (Bakanlar Kurulu Kararı, Cumhurbaşkanlığı Kararnameleri, Genel Yazılar, Gerekçeler, İç Genelgeler, Kanunlar, Maddeler, Özelgeler, Sirküler, Tebliğ Grupları, Tebliğler, Yönetmelikler) yaklaşık 200 Megabyte bilgilere ait veriyi sağlayacaktır. YÜKLENİCİ proje süresince ortaya çıkabilecek yeni mevzuat kaynaklarını kapsama dâhil edecektir.
21. Geliştirilecek çözümde kullanılan modellerin tamamına ait sonuçlar İDARE ile paylaşılacaktır.

22. Proje kapsamında yapay zekâ için kullanılacak tüm veri setleri ve bu veri setlerinin kullanım yöntemleri ve güncellenme sıklığı İDARE tarafından yapılacak çalışmalar sonucu belirlenecektir.
23. Geliştirilecek çözümde modellerin doğru özelge belgesini getirme oranı %85 üzerinde olmalı ve yapılacak çalışmalar ile sürekli iyileştirilebilir olmalıdır.
24. Geliştirilecek çözüm, benzer özelgelerin getirilmesi için anlık 1000 kullanıcı için ortalama 5 ve en fazla 10 saniye içerisinde optimum sürelerde sonuç üretecek ve son kullanıcıya sunabilecektir.
25. Geliştirilecek çözüm, öneri özelge metni oluşturmak için asenkron bir şekilde kullanıcılara hizmet verebilecektir.
26. Geliştirilecek uygulama taşra birimleri, defterdarlık ve merkez teşkilattaki 3000 farklı kullanıcıya hizmet verebilir olmalıdır.
27. Geliştirilecek çözüm son kullanıcılara ara yüzler sunmalı ve bu ara yüzler arama, filtreleme gibi fonksiyonel özellikleri içermelidir. Ayrıca, arayüzden benzer özelgelerin karşılaştırılmasına imkan sağlanacaktır.
28. Geliştirilecek çözümün kaynak kodları, veri tabanlarına ait tüm elemanları, güvenlik şifreleri vb. İDARE'ye teslim edilecektir.
29. Geliştirilecek çözüm ölçeklendirilebilir yapıda tasarlanmalı, veri boyutunun ve kullanıcı sayısının artması durumunda isteklere cevap verebilir yapıda olmalıdır.
30. Geliştirilecek çözümün, Uluslararası Yazılım Geliştirme standartlarına uygun olması, açık kaynaklı teknolojiler kullanılarak geliştirilmesi ve herhangi bir lisans bağımlılığı içermemesi veya lisans ihtiyacı olan uygulamaların kullanımının faydaları olduğu düşünülürse, İDARE'ye bildirilip İDARE'den onay alınması ve İDARE tarafından onay verilmesi halinde kullanımının sağlanması, ISO 27001:2022 süreçlerine uygun ve on-premise çalışması gerekmektedir.
31. Bu proje kapsamında geliştirilecek çözüm, OWASP Top 10 standartlarına uygun olarak detaylı güvenlik testlerinden geçirilecek ve bu standartlara uygun bir şekilde hiçbir güvenlik açığı içermeyecektir. Özellikle her kod güncellemesi veya sürümünden sonra OWASP Top 10 güvenlik testleri yapılarak açıklık olup olmadığı kontrol edilecektir. Bu kontroller için kullanılan test araçlarına ilişkin bilgiler ile test sonuç bilgileri İDARE ile paylaşılacaktır. Yapılan kontroller sonucunda tespit edilen zafiyetler giderilmeden çözüm yayına alınmayacaktır.
32. Kullanıcılar rol bazlı yetkilendirilecektir. Rollere ait yetkiler İDARE ile belirlenecektir.
33. YÜKLENİCİ, İDARE'nin uygun gördüğü personele teorik ve uygulamalı eğitimleri verecek. Eğitimler için gerekli eğitim dokümanlarını hazırlayıp, İDARE'ye teslim edecektir.
34. Proje süresince gerçekleştirilecek tüm çalışmalar, 6698 sayılı "KİŞİSEL VERİLERİN KORUNMASI KANUNU" çerçevesinde gerçekleştirilecektir.
35. Geliştirilecek çözüm için donanım ve altyapı özellikleri YÜKLENİCİ tarafından belirlenecek ve sözleşmenin imzalanma tarihinden sonra 3 ay içinde İDARE'ye iletilecektir.

36. Proje takvimine uygun olarak gerekli donanım ve altyapı İDARE tarafından sağlanacaktır.
37. Geliştirilecek uygulamanın loglama alt yapısı, İDARE'nin standartlarına ve alt yapısına uygun olarak Splunk (Log İzleme Aracı) üzerinde gösterime sunulacaktır.
38. Geliştirilecek çözümün, performansı ve kaynak kullanımı (CPU ve GPU kullanımı, geri bildirim süresi, token sayısı), İDARE tarafından kullanılan (Grafana, Prometheus, ELK) araçlar ile gerçek zamanlı olarak izlenebilir bir arayüz üzerinden takip edilecektir. Sistem sağlığındaki olası sorunlar ve performans problemleri bu arayüzde görüntülenebilir olacaktır.

Dokümanlar

Proje kapsamında aşağıda belirtilen raporlar hazırlanacak ve İDARE ile proje planında belirtilen tarihlerde paylaşılacaktır. İDARE gerekli görmesi durumunda ek raporlar talep edebilecektir.

- Üst Düzey Sistem Tasarımı Dokümanı

Geliştirilecek sistemin genel yapısı ve işleyişi tanımlanmalı, sistem mimarisi detaylı şekilde açıklanmalıdır. Ana bileşenlerin işlevleri ve bunlar arasındaki etkileşimler belirlenmeli, veri akışı ve süreçler netleştirilmelidir. Performans, ölçeklenebilirlik ve güvenlik gibi teknik gereksinimler ele alınmalı ve riskler ile varsayımlar ortaya konulmalıdır. Bu doküman, sistemin başarılı şekilde tasarlanıp uygulanabilmesi için yol gösterici bir rehber olmalıdır.

- Analiz ve Tasarım Dokümanı

Sistemin mevcut durumu analiz edilmeli ve ihtiyaç duyulan özellikler belirlenmelidir. Teknik gereksinimlerden tasarım ilkelerine kadar olan süreç açık bir şekilde sunulmalı, kullanıcı deneyimini iyileştirmek için arayüz tasarımı yapılmalıdır. Ayrıca, veri tabanı yapısı ve sistemin mimari çözümü detaylandırılmalıdır.

- Veri Analiz Raporu

Bu raporda, analiz edilen veri setleri ve kullanılan yöntemler açıklanmalı, veri setinin temel özellikleri tanımlanmalıdır. Bulgular, grafikler ve tablolarla görselleştirilmeli ve analizden elde edilen sonuçlar yorumlanmalıdır. Stratejik kararlar için öneriler geliştirilmeli ve sonraki adımlar belirlenmelidir. Rapor, veri temelli karar alma süreçlerine rehberlik etmelidir.

- Test Raporları

Test süreçleri ve sonuçları detaylı şekilde sunulmalıdır. Test amaçları ve kapsamı netleştirilmeli, uygulanan test senaryoları ve kullanılan yöntemler, performans ve hedeflenen çıktılarla uyum açıklanmalıdır.