



UTAR ENTERPRISE AI PLATFORM MİMARİSİ

Retrieval-Augmented Generation (RAG) Tabanlı
Kurumsal Bilgi Yönetim Sistemi

TEKNİK WHITEPAPER

Hazırlayan: Utar Teknoloji Ar-Ge Ekibi

Tarih: Kasım 2025

Gizlilik Derecesi: Public (Halka Açık)

Konu: Veri Egemenliği, Vektör Uzayı Modellemesi ve Güvenlik
Protokolleri

Özet

Kurumların sahip olduğu yapılandırılmamış veri (unstructured data) hacmi her yıl %50'den fazla artmaktadır. Geleneksel arama yöntemleri ve statik veritabanları, bu verinin içindeki "anlamı" çıkarmakta yetersiz kalmaktadır.

Utar.ai, kurumların 10 yıllık arşivlerini (PDF, DOCX, TXT) saniyeler içinde sorgulanabilir, analiz edilebilir ve özetlenebilir bir "Kurumsal Beyin"e dönüştürür. Bu doküman, Utar platformunun; **Zero-Trust Security (Sıfır Güven Mimarisi)**, **Hibrit Arama (Hybrid Search)** ve **On-Premise/Private Cloud** dağıtım modellerini içeren derinlemesine teknik analizini sunar.

Sistemimiz, KVKK (Kişisel Verilerin Korunması Kanunu) ve GDPR uyumluluğunu, veriyi fiziksel olarak yerel sunucularda tutarak (Data Residency) garanti altına alır.

1. Problem Tanımı ve Çözüm Yaklaşımı

1.1. Kurumsal Hafıza Kaybı ve Veri Siloları

Modern işletmelerde bilgi; e-postalar, sunucular, kişisel bilgisayarlar ve bulut depolama alanları arasında dağılmıştır (Data Silos). Bir personelin ayrılması veya arşivin büyümesi, "bilgiye erişim süresini" (Time-to-Information) logaritmik olarak artırır. Geleneksel **CTRL+F** veya SQL tabanlı aramalar, anlamsal bağlamı (semantic context) kaçıır.

1.2. LLM Halüsinasyonu ve Güvenlik Riski

Genel amaçlı Büyük Dil Modelleri (OpenAI, Anthropic vb.), kurum verisiyle eğitilmemiştir. Bu modellere kurum verisi göndermek, veri sızıntısı riski yaratır. Ayrıca, modellerin bağlam dışı "uydurma" (hallucination) yapma eğilimi, kritik iş süreçlerinde (Hukuk, Finans) kabul edilemez bir risktir.

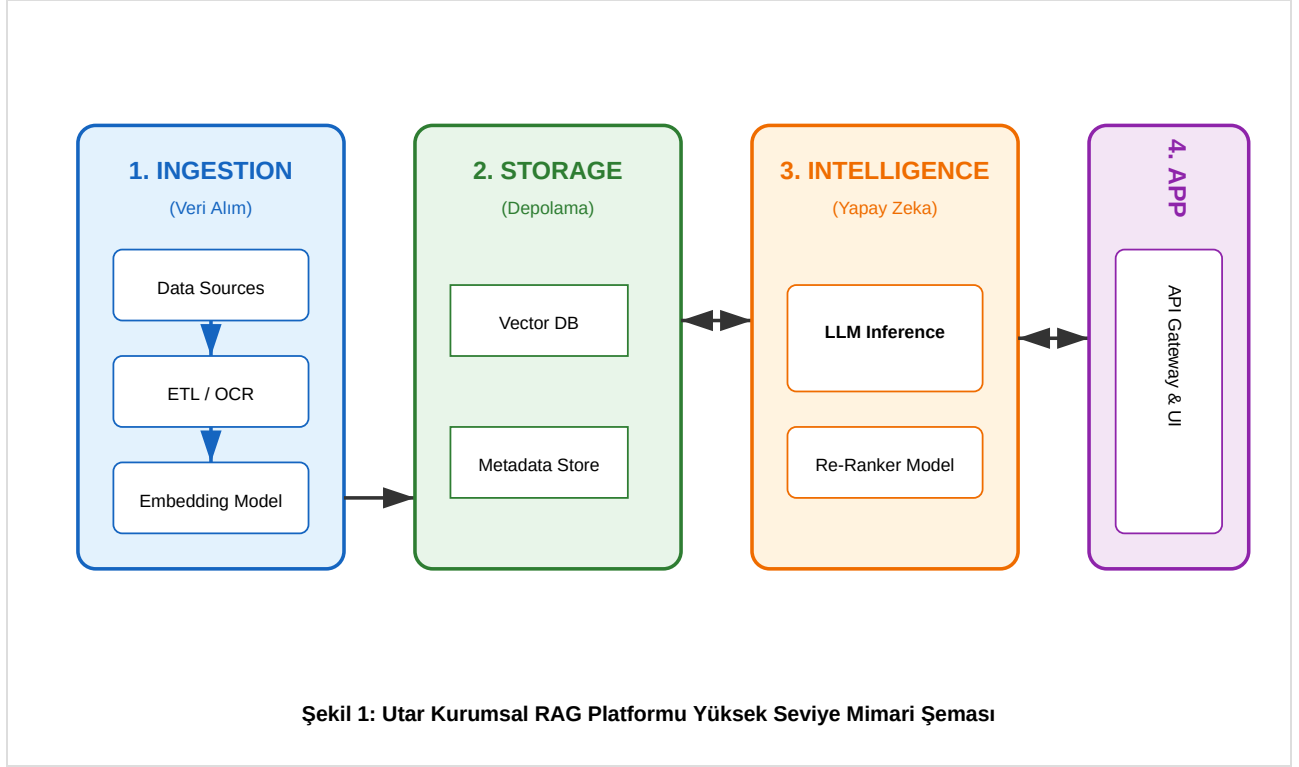
ÇÖZÜM: RAG (RETRIEVAL-AUGMENTED GENERATION)

Uyar, LLM'i bir "bilgi deposu" olarak değil, bir "muhakeme motoru" (Reasoning Engine) olarak kullanır. Bilgi, vektör veritabanından dinamik olarak çekilir ve modele bağlam (context) olarak verilir. Bu, halüsinasyonu minimize eder ve verinin modelin içinde değil, güvenli veritabanında kalmasını sağlar.

2. Sistem Mimarisi ve Bileşenleri

2.1. Yüksek Seviye Mimari

Utar platformu, mikroservis mimarisi üzerine inşa edilmiştir ve Kubernetes (K8s) üzerinde orkestre edilir. Sistem dört ana katmandan oluşur:



2.2. Veri İşleme Hattı (Data Pipeline)

Verinin ham halden vektöre dönüşümü aşağıdaki adımları izler:

- Preprocessing & OCR**: Taranmış belgeler OCR motorları ile işlenir. Gürültü giderme (Noise Reduction) algoritmaları uygulanır.
- Recursive Chunking**: Metinler, anlamsal bütünlüğü korumak için 512-1024 token boyutunda parçalara ayrılır. Parçalar arası %15 örtüşme (overlap) bırakılarak bilgi kaybı önlenir.
- Embedding (Vektörizasyon)**: Çok dilli (Multilingual) modeller kullanılarak metinler 1024 boyutlu vektör uzayına izdüşürülür.

3. Arama Motoru Teknolojisi

3.1. Hibrit Arama (Hybrid Search)

Utar, sektör standartlarının ötesine geçerek tek bir arama yöntemi yerine "Ensemble" (Topluluk) yaklaşımını kullanır.

Yöntem	Teknoloji	Kullanım Amacı
Dense Retrieval	HNSW (Hierarchical Navigable Small World)	Anlamsal benzerlikleri yakalar (örn: "Sözleşme Feshi" $\approx$ "Kontrat İptali").
Sparse Retrieval	BM25 / SPLADE	Kesin anahtar kelime eşleşmelerini yakalar (örn: "TCK Madde 142").

4. Güvenlik ve Şifreleme Protokolleri

4.1. Veri Şifreleme (Encryption)

- At Rest (Durağan Veri):** Disk üzerindeki tüm veriler (Veritabanı, Loglar, Dosyalar) endüstri standardı **AES-256-GCM** (Advanced Encryption Standard) algoritması ile şifrelenir.
- In Transit (Hareketli Veri):** İstemci ve sunucu arasındaki tüm trafik, **TLS 1.3** (Transport Layer Security) protokolü ile şifrelenir.

6. Sonuç ve Stratejik Vizyon

Utar Enterprise AI Platformu, kurumların atıl durumdaki devasa veri yığınlarını aktif bir istihbarat kaynağına dönüştürerek operasyonel bir paradigma değişimi sunmaktadır. Sunduğumuz mimari, bir yazılım ürününden öte, kurumun karar alma mekanizmalarını güçlendiren stratejik bir katmandır.

Platformumuzun sağladığı temel kazanımlar şunlardır:

- Güvenlik ve Uyum:** Veri egemenliği ilkelerine (Data Sovereignty) tam uyum, AES-256 şifreleme ve izole VPC mimarisi ile "Sıfır Risk" ortamı.
- Operasyonel Mükemmellik:** Hibrit Arama ve Re-Ranking teknolojileri sayesinde %95'in üzerinde bilgi erişim doğruluğu (Recall/Precision).
- Geleceğe Hazırlık (Future-Proofing):** Model-agnostik yapısı sayesinde, sistem gelecekte çıkacak daha gelişmiş LLM teknolojilerine mimari değişikliği gerektirmeden adapte olabilir.

Utar, statik bir arşiv yönetiminden "Yapay Zeka Destekli Aktif Muhakeme" çağına geçişin anahtarıdır. Sektörel regülasyonların getirdiği kısıtlamaları bir engele değil, güvenli bir teknolojik altyapı ile rekabet avantajına dönüştürmektedir.