Türkiye'de Millî Yapay Zekâ İçin Güvenli Yerel Geliştirme ve İşletim

Öz

Bu çalışma, Türkiye'de hassas kurumsal belgelerle çalışan Türkçe üretken yapay zekâ sistemleri için millî geliştirme kapasitesini operasyonel olarak tanımlamaktadır. Amaç, yerel kurulum, ticari sırların korunması ve alana özgü anlamsal uyarlamayı denetlenebilir bir mimaride birleştirmektir. Seçilmiş resmî belgeler ve birincil teknik çalışmalar üzerinden yönlendirilmiş belge analizi yapılmış, çıkarımlar izlenebilir kabul ölçütlerine dönüştürülmüştür. Önerilen mimari, kısıtlı yerel çıkarım, belge düzeyinde erişim, sürümlü terminoloji, kaynaklı yanıt ve kontrollü model güncellemesini birlikte ele alır. Tasarım, sanayi tedarik sözleşmeleri için bir sentetik karar örneği ve karşılaştırmalı pilot protokolüyle somutlaştırılır. Çalışmanın çıktısı, kurumların yerli mühendislik katkısını ve dış bağımlılıklarını ayrı ayrı kanıtlamasına yarayan bir değerlendirme çerçevesidir. Pilot sonuçları ve performans iddiaları bu çalışmanın kapsamı dışındadır.

Anahtar kelimeler — Yerel yapay zekâ, millî teknoloji, ticari sır, Türkçe büyük dil modeli, anlamsal uyarlama, yapay zekâ güvenliği.

1 Giriş ve araştırma soruları

Bir kurum yerel bir sunucuda model çalıştırırken ağırlıkların kullanım hakkı, güncellemelerin kaynağı, dış destek erişimi veya hızlandırıcı sürücüleri üzerinde sınırlı denetime sahip olabilir. Buna karşılık Türkiye'de geliştirilen sektörel bir uygulama, çıkarım için yabancı bir API kullanabilir. Her iki durumda da yerellik iddiası ancak hangi bileşenin, kimin yetkisi altında ve hangi işlev için denetlendiği açıklanırsa karşılaştırılabilir. Bu ayrım, kamu hizmetlerinin sürekliliği, sanayi bilgi birikiminin korunması ve tedarik kesintilerinde toparlanma üzerinden ulusal güvenlikle ilişkilidir. İnceleme, sivil kamu ve sanayi uygulamalarıyla sınırlıdır.

Tarihsel 2024–2025 eylem planı yerli üretken modelleri ve Türkçe büyük dil modeli çalışmalarını gündeme almıştır [1, eylem 2.1–2.2]. Sanayi ve Teknoloji Bakanlığı 2026–2030 Türkiye Yapay Zekâ Eylem Planı'nın açıklandığını bildirmiştir [2]. TÜBİTAK 2025 faaliyet raporunda Türkçe Büyük Temel Dil Modeli projesinin yürütüldüğünü kaydetmiştir [3]. Bu belgeler politika yönünü ve bildirilen kurumsal faaliyeti gösterir. Belirli bir modelin üretime hazır oluşu, lisansı ve güvenlik performansı ayrı kanıt gerektirir.

Üç soru incelenir. Millî geliştirme ve yerel işletim hangi gözlenebilir kanıtlarla ayrıştırılabilir? Kurum içi Türkçe görevlerde güvenilir cevap üretimiyle ticari sırların korunması nasıl birlikte tasarlanabilir? Çalışan bir prototip hangi ölçüm ve durdurma ölçütleriyle kurumsal pilota alınabilir? Katkı, bu soruları aynı izlenebilirlik yapısı içinde ele alan bir referans tasarımı ve doğrulama protokolüdür.

2 Araştırma yöntemi ve kaynakların işlevi

Amaçlı kaynak seçimi dört ölçüte dayanır. Belgenin yayımlayanı tespit edilebilir olmalıdır. Metin, araştırma sorularından en az birine doğrudan karşılık vermelidir. Tarih, sürüm veya madde numarası izlenebilmelidir. Teknik mekanizmalar için yöntemi ilk elden tanıtan yayınlara öncelik verilmelidir. Türkiye'ye ilişkin kanıtlar, Bakanlık ve TÜBİTAK belgelerinden ve mevzuatın resmî metin ve açıklamalarından alınmıştır. Yapay zekâ risk yönetiminde NIST çerçevesi ve üretken yapay zekâ profili kullanılmıştır. ISO/IEC 42001 için erişilebilir katalog bilgisi esas alınmıştır. Seçim, sistematik derleme kapsamı veya tarama evreninin tamlığını temsil etmez.

Analiz birimi tek bir politika hedefi, mevzuat maddesi, risk kategorisi veya teknik mekanizmadır. Her birim için kaynak konumu, bağlam, kurum açısından sonuç, önerilen kontrol ve doğrulama kanıtı kaydedilir. Örneğin yurt dışına aktarım düzenlemesinin pratik karşılığı, her ağ çıkışının hukuki değerlendirmeye tabi tutulması ve bir aktarım envanterinde görünür olmasıdır. Teknik kabul kanıtı, örnek bir dağıtımda hedef uçların onaylı listeye uygunluğu ve destek kanallarının erişim kayıtlarıdır. Böylece hukuki dayanak, mühendislik tercihi ve test sonucu birbirine karışmaz.

Çalışmanın analitik sırası şudur. Önce kaynakların neyi gösterdiği belirlenir. Daha sonra iş yükü ve veri sınıfı tanımlanarak kontrol gereksinimleri türetilir. Kontroller bir senaryoya uygulanır ve karşılaştırmalı deney planında ölçülebilir hale getirilir. Kurum görüşmesi, üretim telemetrisi veya gizli belge kullanılmamıştır. İkinci değerlendiriciyle kaynak kodlama uyumu ölçülmemiştir. Bu sınırlamalar, bulguların mevzuata ilişkin kapsamlı görüş veya deneysel sonuç olarak genellenmesini engeller.

3 Mevzuat ve risk yönetişimi

6698 sayılı Kanun kapsamında veri sınıflandırması, işlemenin amacı ve hukuki dayanağı, saklama süresi, erişim yetkisi ve teknik-idari güvenlik tedbirleri ayrı kaydedilmelidir. Kurum içi çalıştırma, dış hizmet kullanımından doğan bazı riskleri azaltabilir. İşleme koşulu veya yetki denetimi bu nedenle ortadan kalkmaz. Yurt dışı telemetri, uzaktan bakım, yedekleme ve üçüncü taraf gözlem araçları dahil bütün veri akışları değerlendirilmelidir. Yurt dışına aktarım söz konusuysa Kanun'un 9. maddesindeki uygulanabilir mekanizma somut akışa göre belirlenir [4, md. 4–6, 9 ve 12; 5].

6102 sayılı Türk Ticaret Kanunu, üretim ve iş sırlarının ifşası gibi davranışları haksız rekabet bağlamında düzenler [6, md. 55/1-d]. Pilotun gizlilik politikası, korunacak varlıkların belirlenmesiyle başlar. Sözleşme fiyatı, üretim toleransı, ürün reçetesi veya kaynak kodu için belge sahibi, erişim gerekçesi ve izinli kullanım tanımlanır. Sözleşme bakımından kullanım yetkisine sahip olmak, belgeyi eğitim verisine katma yetkisini kendiliğinden sağlamaz. Belge sözleşmeleri ve eğitim izinleri ayrı incelenir. 7545 sayılı Kanun'un amacı ve kapsamı, yapay zekâ hizmetini işleten kuruluşun siber güvenlik yükümlülüklerinin değerlendirilmesi için başlangıç noktasıdır [7, md. 1–2]. İşletim öncesi güncel konsolide metin ve ilgili ikincil düzenlemeler kontrol edilir.

NIST AI RMF'nin Govern, Map, Measure ve Manage işlevleri, sorumluluk ataması, kullanımın sınırlandırılması, test ve riskin işletilmesi için yapı sağlar [8, §5]. NIST üretken yapay zekâ profili, uydurma içerik, veri mahremiyeti, bilgi güvenliği ve bileşen entegrasyonu risklerini değerlendirmeye dahil eder [9, §2]. ISO/IEC 42001 bir yapay zekâ yönetim sistemi standardıdır [10]. Bu çalışmada standardın katalog kapsamı kullanılmıştır. Belirli bir maddeye uygunluk veya belgelendirme sonucu çıkarılmamıştır.

4 Millî geliştirme kapasitesi ve tedarik

Millî geliştirme iddiası için beş farklı kanıt önerilir. Birincisi veri ve ağırlıklar üzerindeki hukuki kullanım, uyarlama ve dağıtım haklarıdır. İkincisi kaynak kodu, eğitim reçetesi, veri seçimi ve model seçimine ilişkin yerli mühendislik katkısıdır. Üçüncüsü anahtarların, çıkarım hizmetinin ve güncelleme kararlarının kurumsal denetimidir. Dördüncüsü, bağımlı olunan donanım ve yazılım bileşenlerinin izlenebilirliğidir. Beşincisi, tedarikçi ayrıldığında sistemin taşınması, geri yüklenmesi ve yeniden devreye alınabilmesidir. Bu ölçütler bir gümrük menşei veya resmî yerlilik oranı tanımı oluşturmaz.

Yabancı kökenli bir temel modeli bir sektör uygulaması için yerelde uyarlamak, uygulama kodu, veri kürasyonu, değerlendirme ve işletim bakımından yerli katkı üretebilir. Temel modelin menşei ayrıca açıklanır. Türkiye'de ön eğitim ise model tasarımı ve veri seçimi üzerinde daha geniş karar alanı verir. Buna eğitim verisinin hukuki temizliği, hesaplama maliyeti, deney tekrarı ve uzun dönem bakım sorumluluğu eşlik eder. Bu nedenle alım şartnamesi model menşei, ağırlık lisansı, eğitim ve uyarlama yeri, çalıştırma yeri ile yönetim yetkisini ayrı alanlar halinde ister.

Her aday için kanıt dosyası tutulur. Dosyada bileşen adı ve sürümü, sağlayıcı, lisans, kriptografik özet, nerede işlendiği, güvenlik güncellemesinin sahibi ve ikame seçeneği yer alır. Yabancı bir sürücüye veya tek üreticinin hızlandırıcısına bağımlılık varsa ikinci platformda kurtarma tatbikatı veya doğrulanmış geçiş planı istenir. Ulusal ölçekte ortak Türkçe değerlendirme kümeleri ve kurumlar arasında paylaşılabilir güvenlik profilleri geliştirilebilir. Gizli veriler kurumların yetki sınırları içinde tutulurken ortak test senaryoları sentetik veya açık lisanslı içerikle yayımlanabilir.

5 Güvenli yerel kurulum ve işletim

Önerilen akış şöyledir. Kullanıcı kurumsal kimlikle oturum açar ve istek, politika katmanında görev ile veri sınıfına göre değerlendirilir. Yetkili belge adayları metinsel ve anlamsal dizinlerden getirilir. Yerel model, yalnızca bu adaylardan seçilen bağlamla cevap taslağı oluşturur. Çıktı katmanı kaynak kimliği ve belge sürümünü denetler, gerektiğinde yanıtı durdurur ve kullanıcıya inceleme yolu sunar. Bu akışta model ağırlıkları, belge deposu, dizinler ve uygulama sunucusu birbirinden ayrı yetki alanlarıdır. İzleme kayıtlarına tam istem veya belge içeriği yazılması varsayılan işletim biçimi değildir.

Kurulum sırası envanter, çevre hazırlığı, dağıtım ve kabulden oluşur. Model ve bağımlılıklar yetkili hazırlık alanında alınır. Lisans, kaynak, sürüm,

bütünlük ve uyumluluk kayıtları oluşturulur. Onaylı paketler iç depoya alınır ve üretim alanında dış bağlantı politikası, iç sertifikalar, servis hesapları ve yedekleme tanımlanır. Bir model sürümü, tokenizer, gömme modeli, dizin şeması ve istem şablonu birlikte bir dağıtım manifestosuna bağlanır. Geçiş öncesi eski ve yeni sürüm aynı kilitli değerlendirme kümesiyle sınanır. Geri dönüş, hem model hem dizin ve politika sürümünü kapsar.

Belge izinleri getirimden önce uygulanır. Sonuç önbelleği, vektör araması ve kaynak önizlemesinde aynı yetki kapsamı korunur. Yetkisi geri alınan kullanıcının eski oturumları ve önbellek girdileri geçersizleştirilir. Belgeler talimat yetkisine sahip sayılmaz. Model, veri okuma veya yazma işlevlerini kendi başına çağıramaz. Kullanıcı arayüzü, yetkisiz belgenin başlığı ve varlığı gibi dolaylı bilgi sızıntılarını da engelleyecek biçimde tasarlanır. Model davranışı ile erişim denetimi ayrı bileşenlerce işletilir [8, 9].

Günlük işletimde kaynak sürümleri, yetki değişiklikleri ve olay eğilimleri izlenir. Başarısız yetkilendirme, dış bağlantı ve geriye dönüş sinyalleri, içeriği gereksiz yere kopyalamadan ilişkilendirilebilir bir olay kimliğiyle kaydedilir. Model servisi karantinaya alındığında yetkili kullanıcılar onaylı belge aramasına dönebilir. Bu süreklilik yaklaşımı, üretken bileşenin devre dışı kaldığı durumda kurumsal bilgiye erişimin sürmesini sağlar. Kurum, geri yükleme süresi ve veri kaybı toleransını iş etkisine göre belirler ve tatbikatta ölçer.

Pilot kapasitesi için başlangıç hesabı, P parametre ve ağırlık başına b bit için P × b / 8 bayttır. Örneğin 8 milyar parametre ve 4 bit için yalnızca paketlenmiş ağırlıkların ideal alt hesabı yaklaşık 4 GB’dır. Ölçekleme verileri, çalışma zamanı belleği, anahtar-değer önbelleği ve eşzamanlı istekler bu hesaba eklenir. Kabul testi, hedef bağlam uzunluğunda gerçek tepe bellek kullanımını ve gecikmeyi ölçer. Kaynak analizinde donanım maliyetine ek olarak bakım, enerji, depolama ve alan uzmanı incelemesi kaydedilir. Başlangıç kurulumu küçük ve sınırlandırılmış bir test yüküyle doğrulanır. Model küçültme veya nicemleme kararı, özellikle madde numarası ve süre doğruluğunun yeniden değerlendirilmesini gerektirir.

6 Görev ve alan odaklı anlamsal uyarlama

Uyarlama, kullanılacak görev ve yetki kapsamının yazılmasıyla başlar. Örnek görev, onaylı sanayi tedarik sözleşmelerinden bildirim yükümlülüklerini bulmaktır. İstenen çıktı yürürlükteki hüküm, belge adı, madde, tarih ve cevaplanamayan noktaları kapsar. Uygulama, bağlayıcı yorum veya işlem kararı için yetkili personele yönlendirir. Bu sınır, hem test sorularını hem de eğitim verisinde neyin kabul edilebilir olduğunu belirler.

Terim sözlüğü kavram kimliği, tercih edilen Türkçe terim, eş anlamlılar, kısaltmalar, alan tanımı ve geçerlilik aralığını içerebilir. Örneğin “mücbir sebep” ile “force majeure” aynı kavrama bağlanır. “Takvim günü” ve “iş günü” ayrı kavramlar olarak korunur. Metin, aranabilir bir kopyada Unicode normalleştirmesine tabi tutulabilir [11]. Orijinal metin delil ve kaynak gösterimi için saklanır. Türkçe I, İ, ı ve i biçimleri ile ekler, madde işaretleri ve bölünmüş PDF satırları sınanır. Otomatik eşlemenin kavramları yanlış birleştirmesi, uzman onayına tabi değişiklik kayıtlarıyla yönetilir.

Hibrit getirim, sözleşme numarası ve madde gibi birebir eşleşmeleri anlamsal sorgularla birleştirir. Sonuçlar geçerlilik ve kullanıcının yetkisiyle filtrelenir. Getirim destekli üretimde model, parametrelerine ek olarak haricî kaynaklardan seçilen bağlamı kullanır [12]. Bu yöntem, sık değişen bilgiyi sürümlü belgeler üzerinden güncelleme olanağı sağlar. Kendiliğinden kaynak doğruluğu sağlamaz. Yanıtın her önemli iddiası, ilgili belge parçasıyla insan veya doğrulama mantığı tarafından karşılaştırılır.

“Anlamsal eğilimleştirme” bu çalışmada üç gözlenebilir işlemi ifade eder. Terimlerin kavram sözlüğüyle eşleştirilmesi, görev örneklerine göre yanıt davranışının uyarlanması ve değerlendirme sonuçlarının hata türüne göre izlenmesi. Kaynaklı yanıt kuralları önce istem ve getirimle uygulanır. Kalıcı biçim veya görev hatası sürüyorsa, lisanslı ve hassas içerikten arındırılmış örneklerle parametre verimli ince ayar denenebilir [13]. Eğitim ve test kümeleri belge ailesi ve zamana göre ayrılır. Teste ait bir belgenin başka sürümünün eğitime girmesi, görünen başarıyı olduğundan yüksek gösterebilir.

7 Sentetik uygulama ve değerlendirme protokolü

Sentetik vakada bir iç denetçi sevkiyat gecikmesinin bildirim süresini sorar. Eski sürüm iki iş günü, yürürlükteki ve yetkili sürüm bir iş günü belirtir. Ayrıca başka bir tedarikçiye ait kısıtlı fiyat eki vardır. Beklenen davranış, yürürlükteki yetkili maddeyi kaynak göstererek bir iş günü cevabını vermek ve fiyat ekinin varlığını açığa çıkarmamaktır. Yetkili yürürlükteki nüsha eksikse sistem, mevcut kaynağın yetersiz olduğunu belirterek inceleme ister. Bu, örnek bir oracle tanımıdır. Sistemin gerçekten bu davranışı ürettiğine ilişkin ölçüm yapılmamıştır.

Pilot için dört koşul önerilir. A, belge arama ve insan değerlendirmesiyle oluşturulan operasyonel tabandır. B, dış bağlantısı kapalı temel modeldir. C, aynı modelin yetkili hibrit getirim ve sözlükle çalışmasıdır. D, C üzerine eklenen görev amaçlı ince ayardır. C’nin katkısı daha ayrıntılı incelenecekse yalnızca sözlük kapatılır ve diğer değişkenler sabit tutulur. Tüm yapay zekâ koşullarında aynı test soruları ve yetki kuralları kullanılır. Donanım, bağlam bütçesi, model sürümü, eşzamanlı kullanıcı yükü ve rastlantısallık ayarları raporlanır.

Test kümesi, doğrudan madde sorularını, farklı ifadelerle aynı kavramı soran örnekleri, çelişkili sürümleri, cevapsız soruları ve rol bazlı erişim senaryolarını kapsar. Her örnekte izinli belge, doğru hüküm, geçerlilik tarihi ve beklenen yanıt veya çekimserlik etiketlenir. Model türünden habersiz iki alan değerlendiricisi bağımsız puan verir. Anlaşmazlıklar gerekçeli karara bağlanır. Belge aileleri kümelenmiş gözlem olarak ele alınır. Karşılaştırmalar eşleştirilmiş yapılır ve belirsizlik aralıkları verilir.

Bilgi getiriminde Hit@k, cevaplanabilir sorgularda ilk k sonuç içinde en az bir doğru yetkili kaynağın bulunma oranıdır. Kaynak desteği, değerlendirilen maddi iddiaların yetkili ve yürürlükteki kaynakla desteklenen payıdır. Çekimserlik, cevaplanabilir ve cevaplanamaz sınıfları için ayrı raporlanır. İşletim ölçütleri, sabit yükte yüzde 95 gecikme, başarılı ve kaynaklı cevap başına maliyet ile geri yükleme süresini içerir. Yetkisiz ifşa, herhangi bir testte görüldüğünde pilotun durdurulması için yeterlidir. Sıfır gözlenen ifşa artık riskin sıfır olduğunu göstermez. Bağımsız

300 denemede sıfır olayın tek taraflı %95 binom üst sınırı yaklaşık %1'dir. Ortak belgeler ve şablonlar bağımsızlık varsayımını zayıflatır.

Tablo 1. Pilot için örnek kabul kayıtları

Karar alanı	Beklenen kanıt	Karar
Kullanım hakkı	Lisans, belge kullanım yetkisi ve ilgili işleme kaydı	Eksik kanıtta dağıtımı durdur.
Yetki sınırı	Belge ve kaynak önizleme izinlerinin rol bazlı test kaydı	Yetkisiz ifşada pilotu durdur.
Kaynak ve sürüm	Yanıt iddiası, yürürlük tarihi ve kaynak pasajı eşlemesi	Çelişki çözülene kadar uzman incelemesi iste.
Geri dönüş	Son onaylı sürüme ve belge aramasına geçiş tatbikatı	Kurumun hedef süresi karşılanmazsa üretime geçme.
Bağımlılık ve ikame	Model, çalışma zamanı, sürücü ve hızlandırıcı envanteri	Kritik bileşen değişimi için doğrulanabilir plan iste.

8 Tartışma ve sonuç

Yerel yapay zekâ yatırımının başarısı, ağırlıkların çalıştığı ülke, Türkçe niteliği, mühendislik katkısı ve kurumun işletim denetimi gibi farklı sonuçlarla değerlendirilmelidir. Yüksek Türkçe kalite, izin hatasını telafi etmez. Yerel yürütme, yetkili belge kaynaklarının güncel ve kullanılabilir olmasını kendiliğinden sağlamaz. Bu bağımsız boyutlar nedeniyle tek bir toplam puan yerine zorunlu güvenlik ve kullanım hakkı koşullarının ardından görev kalitesi, taşınabilirlik ve maliyet karşılaştırması önerilmektedir.

Uygulamada önce sivil ve geri döndürülebilir bir belge iş yükü seçilir. Kurum, bağımlılık envanterini ve sınırlı yetkili yerel çıkarımı kurar. Sürümlü kaynaklar, terminoloji ve insan denetimiyle pilot yürütülür. Birden fazla kurumda değerlendirilen ortak protokoller, hangi yerli model veya yerel uyarlama yolunun hangi görev için yeterli olduğunu daha güvenilir biçimde gösterebilir. Özellikle sınırlı Türkçe veri alanları, farklı belge aileleri ve değişken donanım koşulları ayrı incelenmelidir. Çalışma, bu karşılaştırmaları mümkün kılan bir tasarım sunar. Üretim güvenliği ve ulusal ölçekte etki, bu tasarımın ileride elde edilecek gözlem sonuçlarıyla değerlendirilmelidir.

Uygulanacak olan yerelleştirilmiş yapılardaki her türlü AI(Yapay Zeka) çalışmalarının Open Source kaynaklarına dayalı olması ve Milli Değerlerimizi taşıması gerekmektedir. Keza bu Milli Değerlerimizi Orta Asya'dan günümüze kadar gelişen Türk Kültür ve Gelenekleri içinde değerlendirilmelidir.

Secure Local Development and Operation of National AI Capabilities in Türkiye

Abstract

This study operationalises national AI development capacity for Turkish-language generative systems handling sensitive organisational documents in Türkiye. It combines local deployment, trade secret protection and domain-specific semantic adaptation in an auditable design. A directed analysis of selected official documents and primary technical publications informs traceable acceptance criteria. The architecture combines restricted local inference, document-level authorisation, versioned terminology, source-supported answers and controlled model updates. A synthetic industrial procurement case and a comparative pilot protocol illustrate the design. The study provides an assessment framework for documenting domestic engineering contributions and external dependencies. It reports neither pilot measurements nor performance estimates.

Keywords — Local artificial intelligence, national technology, trade secrets, Turkish large language models, semantic adaptation, AI security.

1 Introduction and research questions

An organisation can run a model on a domestic server while exercising limited control over weight licences, update provenance, external maintenance or accelerator drivers. A sectoral application developed in Türkiye may conversely outsource inference to a foreign API. Locality claims become comparable when they specify which component is controlled, by whom and for what purpose. These distinctions matter for national security through continuity of public services, protection of industrial knowledge and recovery from supply disruptions. The scope is civilian public-sector and industrial use.

The historical 2024–2025 action plan addressed domestically developed generative models and Turkish large language models [1, actions 2.1–2.2]. The Ministry of Industry and Technology reported the announcement of the Türkiye AI Action Plan for 2026–2030 [2]. TÜBİTAK's 2025 activity report records ongoing work on the Turkish foundation language model project [3]. These records establish policy orientation and reported institutional activity. Production readiness, licence conditions and security performance require model-specific evidence.

The study asks three questions. What observable evidence distinguishes national development from local operation? How can reliable Turkish-language answers and trade secret protection be designed together? What measurements and stop criteria should govern a transition to an organisational pilot? The contribution is a reference architecture and verification protocol that address these questions through the same evidence structure.

2 Research method and use of evidence

Purposive selection followed four criteria. The issuing organisation had to be identifiable. Each source had to address at least one research question directly. A date, version or provision had to be traceable. Primary descriptions were preferred for technical mechanisms. Evidence concerning Türkiye came from Ministry and TÜBİTAK publications and official legislative texts or interpretations. The NIST framework and generative AI profile informed AI risk analysis. The accessible catalogue description supported the discussion of ISO/IEC 42001. The sample does not imply comprehensive coverage of the literature.

The unit of analysis is an individual policy objective, statutory provision, risk category or technical mechanism. For each unit, the source location, context, organisational implication, proposed control and verification evidence are recorded. For instance, the operational implication of cross-border transfer rules is to review each relevant outbound flow and include it in a transfer inventory. Technical acceptance evidence includes whether destinations conform to the approved list and whether support access is recorded in a test deployment. This keeps legal authority, engineering design and observed behaviour distinguishable.

The analysis first determines what each source establishes. It then specifies a workload and data classification, derives controls, and applies them to a scenario that can be evaluated in a comparison. The study used no organisational interviews, production telemetry or confidential records. Coding agreement with a second reviewer was not measured. The findings therefore have the scope of a document-based design study.

3 Legal context and risk governance

Under Law No. 6698, data classification, processing purpose and legal basis, retention periods, access permissions, and technical and administrative safeguards should be recorded separately. On-premises execution can reduce exposure associated with external services. Processing conditions and access controls still apply. Data flows include foreign telemetry, remote maintenance, backups and third-party observability tools. Where a cross-border transfer occurs, the applicable mechanism under Article 9 depends on the specific flow [4, arts. 4–6, 9 and 12; 5].

The Turkish Commercial Code No. 6102 addresses disclosure of manufacturing and business secrets under unfair competition rules [6, art. 55(1)(d)]. A pilot confidentiality policy starts by identifying protected assets. For contract prices, manufacturing tolerances, formulas or source code, it specifies an owner, access justification and permitted uses. Authority to use a document in a contractual workflow does not automatically authorise its inclusion in a training corpus. Contractual rights and training permissions require distinct review. The purpose and scope of Cybersecurity Law No. 7545 provide a starting point for assessing the organisation’s cybersecurity obligations [7, arts. 1–2]. Applicable consolidated texts and secondary rules should be checked before operation.

The Govern, Map, Measure and Manage functions of the NIST AI RMF structure assignment of responsibility, documentation of scope, testing and ongoing risk

management [8, sec. 5]. The generative AI profile brings confabulation, privacy, information security and component integration into that assessment [9, sec. 2]. ISO/IEC 42001 is an AI management system standard [10]. This study uses its catalogue scope. It does not infer conformity with particular clauses or certification.

4 National development capacity and procurement

Five kinds of evidence are proposed for national development claims. First, document rights to use, adapt and distribute relevant data and weights. Second, identify domestic engineering contributions to source code, training recipes, data selection and model design. Third, identify institutional control of keys, inference services and update decisions. Fourth, trace software and hardware dependencies. Fifth, demonstrate the ability to migrate, restore and resume service if a supplier exits. These dimensions are analytical criteria and do not establish a customs origin or statutory local-content percentage.

A foreign foundation model adapted locally for a sectoral application can support domestic contributions in application code, data curation, evaluation and operation. The origin of the underlying model is disclosed separately. Pretraining in Türkiye can expand control over model design and corpus selection. It also creates obligations concerning data rights, compute costs, reproducibility and long-term maintenance. Procurement specifications should therefore request separate entries for model origin, weight licence, training and adaptation location, execution location and administrative authority.

Each candidate should have an evidence file recording the component name, version, supplier, licence, cryptographic digest, processing location, security update owner and substitute. A dependency on an overseas driver or a single accelerator vendor calls for a recovery exercise on a second platform or a verified migration plan. Shared Turkish evaluation sets and reusable security profiles could support cooperation across organisations. Sensitive data can stay within institutional boundaries while common test scenarios are published with synthetic or openly licensed material.

5 Secure local deployment and operation

The proposed workflow begins with organisational authentication. A policy layer examines the request against the task and data class. Eligible documents are retrieved through lexical and semantic indexes. A local model drafts an answer using the selected context. The output layer checks source identifiers and versions, can withhold the response, and offers a route to human review. Model weights, the document store, indexes and the application server occupy distinct access domains. Full prompts and document contents are excluded from default operational logging.

Deployment proceeds through inventory, environment preparation, installation and acceptance. Obtain models and dependencies in an authorised staging area. Record the licence, source, version, integrity and compatibility before importing approved artefacts to an internal repository. Set production egress policies, internal certificates, service accounts and backups. Bind model, tokenizer, embedding model, index schema and prompt template to the same release manifest. Test old

and new releases against the same locked evaluation set. Rollback covers the policy and index versions as well as the model.

Apply document authorisation before retrieval. Maintain the same scope in answer caching, vector search and source previews. Revoking access invalidates affected sessions and cached entries. Retrieved documents carry no instruction authority. The model cannot independently invoke data reads or writes. The interface must prevent indirect disclosure, including the title or existence of an unauthorised document. Distinct components govern model behaviour and enforce access rules [8, 9].

Routine operations monitor source versions, access changes and incident trends. Correlate access denials, outbound connections and rollback signals through event identifiers without unnecessarily copying content. If the model service is quarantined, authorised users can return to approved document search. Maintaining that fallback preserves access to organisational knowledge during interruption of generation. The organisation defines restoration time and acceptable data loss based on business impact and measures both in exercises.

A first capacity estimate for P parameters and b bits per weight is $P \times b / 8$ bytes. For an eight-billion-parameter model at four bits, the ideal packed weights alone require about 4 GB. Quantisation metadata, runtime memory, the key-value cache and concurrent requests are additional. Acceptance testing measures peak memory and latency at the intended context length. Resource accounting includes hardware, maintenance, energy, storage and expert review. Initial deployment is validated on a small, bounded workload. Model compression or quantisation calls for renewed checks of provision identifiers and deadline accuracy in particular.

6 Task and domain-specific semantic adaptation

Begin by specifying the task and its authority boundary. A sample task retrieves notification obligations from approved industrial procurement contracts. The requested output includes the effective provision, source document, article, date and unresolved points. Binding interpretations and decisions go to an authorised person. This boundary governs what counts as a valid test question and acceptable training material.

A terminology register can contain concept IDs, preferred Turkish terms, synonyms, abbreviations, domain definitions and validity periods. For example, “mücbir sebep” can be linked to “force majeure,” while calendar days and working days remain distinct. A searchable copy can undergo Unicode normalisation [11]. Retain the original for evidence and attribution. Test Turkish forms I, İ, ı and i, inflection, enumeration and line breaks in PDFs. Manage mistaken concept mergers through a change history subject to domain expert approval.

Hybrid retrieval combines exact matches such as contract numbers and article labels with semantic search. Results are filtered by effective date and user authority. Retrieval-augmented generation adds selected external context to the model's parametric knowledge [12]. Versioned documents can therefore update frequently changing information. Attribution still requires verification. A reviewer or

verification component compares each material claim with the relevant source passage.

Semantic adaptation here entails three observable activities. The system maps terms to controlled concepts, adapts response behaviour using task examples and monitors evaluation outcomes by error category. Prompting and retrieval implement answer grounding first. If persistent task or format errors remain, parameter-efficient fine-tuning can be evaluated using licensed, non-sensitive examples [13]. Split training and testing by document family and time. Allowing a different version of a test document into training can inflate apparent performance.

7 Synthetic application and evaluation protocol

In the synthetic case, an internal auditor asks about the deadline for notification of a shipment delay. An obsolete contract version specifies two working days. The authorised and effective version specifies one. There is also a restricted price annex for another supplier. The expected behaviour is to answer one working day, cite the applicable authorised clause and avoid disclosing the annex's existence. If the authorised effective copy is missing, the system should report insufficient evidence and request review. This is an evaluation oracle. No system behaviour was measured.

Four pilot conditions are proposed. A is an operational baseline with document search and human review. B uses a base model with egress disabled. C uses the same model with authorised hybrid retrieval and the terminology register. D adds task fine-tuning to C. An ablation can disable the register alone to examine its contribution. Apply the same questions and authority rules across all AI conditions. Report hardware, context budget, model version, concurrent load and sampling settings.

The test set includes direct provision queries, paraphrases, conflicting versions, unanswerable questions and role-based access scenarios. Annotate the allowed document, correct provision, date of effect and expected answer or abstention. Two domain raters score independently without knowing the model condition. Adjudicate disagreements with reasons. Treat document families as clusters. Use paired comparisons and report uncertainty intervals.

For answerable queries, Hit@k is the proportion with at least one relevant authorised source among the first k results. Citation support measures the share of material claims supported by authorised and effective evidence. Report abstention separately for answerable and unanswerable cases. Operational measures include 95th-percentile latency under fixed load, cost per successful grounded answer and restoration time. Any observed unauthorised disclosure stops the pilot. Zero observed disclosures leave residual uncertainty. With zero events in 300 independent trials, the one-sided 95% binomial upper bound is approximately 1%. Shared documents and templates weaken the independence assumption.

Table 1. Example pilot acceptance records

Decision area	Expected evidence	Decision
Usage rights	Licence, document use rights and applicable processing record	Pause deployment if evidence is missing.
Authorisation	Role-based checks of document and source-preview access	Stop the pilot on unauthorised disclosure.
Source and version	Mapping from claim to effective date and source passage	Refer conflicting evidence for expert review.
Recovery	Exercise restoring the approved release and document search	Hold production release if restoration misses the target.
Dependencies	Inventory of models, runtimes, drivers and accelerators	Require a verified replacement plan for critical components.

8 Discussion and conclusion

The success of local AI investment should be assessed across execution location, Turkish-language quality, engineering contribution and operational control. High language quality cannot compensate for an authorisation failure. Local execution does not ensure that authorised documents are current and usable. These dimensions merit mandatory security and usage rights checks, followed by comparisons of task quality, portability and cost among eligible candidates.

An implementation can start with a reversible civilian document workload. The organisation establishes its dependency inventory and limited-scope local inference, then pilots versioned sources, terminology and human review. Common protocols used across institutions can clarify which domestic models and adaptation strategies are adequate for specific tasks. Low-resource Turkish domains, different document families and variable hardware conditions warrant separate evaluation. This study provides a design that enables those comparisons. Production security and national-scale impact require subsequent observation.

All artificial intelligence (AI) work carried out within localized systems must draw on open-source resources and embody our national values. These values should be understood within the context of Turkish culture and traditions, from their origins in Central Asia to their evolution into the present day.

土耳其国家人工智能能力的安全本地开 发与运行

摘要

本文针对土耳其处理机构敏感文件的土耳其语生成式人工智能系统，提出可操作的国家人工智能开发能力评价框架。研究将本地部署、商业秘密保护和面向专业领域的语义适配纳入同一可审计设计。通过对选定官方文件及一手技术文献进行定向文件分析，本文推导出可追溯的验收标准。参考架构结合受限本地推理、文件级授权、术语版本管理、有来源依据的回答和受控模型更新。本文以合成的工业采购合同案例说明系统预期行为，并提出对照试点方案。研究成果可用于记录本土工程贡献和外部依赖。本研究未执行试点，也不报告性能估计。

关键词　本地人工智能，国家技术能力，商业秘密，土耳其语大语言模型，语义适配，人工智能安全。

1 引言与研究问题

机构可以在境内服务器上运行模型，但对模型权重的许可、更新来源、外部维护或加速器驱动程序的控制仍可能有限。相应地，在土耳其开发的行业应用也可能依赖外国服务接口执行推理。只有明确由谁、出于何种目的控制哪些组件，才能比较不同系统的本地化程度。这些差别关系到公共服务连续性、产业知识保护和供应中断后的恢复，因此具有国家安全意义。本文范围限于民用公共部门和工业应用。

2024—2025 年行动计划将本土生成式模型及土耳其语大语言模型列入议程[1，行动 2.1—2.2]。工业和技术部报告称，土耳其 2026—2030 年人工智能行动计划已公布[2]。土耳其科学技术研究委员会（TÜBİTAK）2025 年度活动报告记载，土耳其语基础语言模型项目正在推进[3]。这些文件反映政策方向和已报告的机构活动。模型能否投入生产、适用何种许可证以及具备何种安全表现，需要具体证据。

本文研究三个问题。哪些可观察证据可以区分国家开发能力与本地运行能力？如何在设计中兼顾土耳其语回答的可靠性和商业秘密保护？系统进入机构试点前应接受何种测量和停止规则？本文的贡献是以统一证据结构回答这些问题的参考架构与验证方案。

2 研究方法与证据用途

采用目的性选样。入选来源须能确认发布机构，直接涉及至少一个研究问题，并能追溯日期、版本或条款。技术机制优先采用首次介绍或直接报告该机制的一手文献。土耳其相关证据来自工业和技术部、TÜBİTAK 以及正式法律文本或官方解释。人工智能风险分析参考 NIST 框架和生成式人工智能风险指南。

对 ISO/IEC 42001 的讨论以公开目录说明为依据。所选文献不代表对整个领域的全面覆盖。

分析单位为单项政策目标、法律条款、风险类别或技术机制。每一单位记录来源位置、适用背景、机构层面的含义、拟议控制及验证证据。例如，跨境传输规则在运行中的落实方式包括审核每条相关出站数据流，并纳入传输清单。技术验收可检查测试环境中的目标端点是否属于批准范围以及支持人员的访问是否留有记录。这样可以区分法律依据、工程决策和观测结果。

研究首先明确各来源所能支撑的结论，再界定工作负载和数据分类、推导控制，并将控制用于可通过对照方式评估的场景。研究没有使用机构采访、生产遥测或保密文件，也未由第二名评审者独立编码并测量一致性。结论的适用范围限于文件分析支持的设计研究。

3 法律背景与风险治理

依据第 6698 号法律，应分别记录数据类别、处理目的和法律依据、保存期限、访问权限及技术和管理保障措施。本地运行可减少部分外部服务风险，处理条件和访问控制仍应落实。数据流评估应涵盖境外遥测、远程维护、备份和第三方可观测性工具。涉及跨境传输时，根据具体数据流确定第 9 条下适用的机制[4，第 4—6、9 及 12 条；5]。

第 6102 号《土耳其商法典》将泄露生产及业务秘密等行为纳入不正当竞争规则[6，第 55 条第 1 款 d 项]。试点保密政策从确定受保护资产开始。对合同价格、生产公差、配方和源代码等，应明确资产负责人、访问理由和允许用途。合同流程中的文件使用权并不会自动赋予将文件纳入训练语料的权利。合同和训练许可应分别审查。第 7545 号《网络安全法》的目标和适用范围为评估运营机构的网络安全义务提供起点[7，第 1—2 条]。上线前还应核查适用的现行合并文本和配套规则。

NIST 人工智能风险管理框架的治理、映射、测量和管理四项职能，可用于分配职责、界定用途、实施测试和持续治理风险[8，第 5 节]。生成式人工智能风险指南将内容编造、隐私、信息安全和组件集成纳入评估[9，第 2 节]。ISO/IEC 42001 是人工智能管理体系标准[10]。本文依据其目录所示范围，不据此判定具体条款符合性或认证状态。

4 国家开发能力与采购

本文提出五类证据。第一，数据和模型权重的使用、适配及分发权利。第二，本土团队在源代码、训练方案、数据筛选和模型设计方面的工程贡献。第三，机构对密钥、推理服务和更新决策的控制。第四，软硬件依赖的可追溯性。第五，供应商退出后迁移、恢复和重新运行的能力。这些是分析指标，不构成海关原产地或法定本地含量的定义。

通过本地适配外国基础模型开发行业应用，可以在代码、数据治理、评测和运行方面形成土耳其本土工程贡献。基础模型来源应单独披露。在土耳其开

展预训练可以扩大对模型设计和语料选择的控制，也会带来数据权利、算力成本、实验复现和长期维护方面的责任。采购文件应分别列明模型来源、权重许可证、训练及适配地点、运行地点和管理权限归属。

每个候选方案应建立证据档案，记录组件名称、版本、供应商、许可证、加密摘要、数据处理位置、安全更新负责人和可替代方案。如果依赖境外驱动程序或单一加速器厂商，应要求在第二平台开展恢复演练或验证迁移计划。机构间可共享土耳其语评测集和可复用的安全配置。敏感数据可保留在机构授权边界内，共用场景可使用合成数据或开放许可材料发布。

5 安全本地部署与运行

流程从机构身份认证开始。策略层按任务和数据类别审核请求。系统从词项索引和语义索引检索可访问文件，本地模型依据选定上下文起草回答。输出层核对来源标识和文件版本，可暂停出具回答并转交人工审查。模型权重、文件库、索引和应用服务器处于独立权限域。默认运行日志不保存完整提示词或文件正文。

部署依次开展清点、环境准备、安装和验收。在获准的准备区获取模型及依赖，记录许可证、来源、版本、完整性和兼容性，再将批准的制品导入内部存储库。配置生产环境的出站策略、内部证书、服务账号和备份。模型、分词器、嵌入模型、索引架构和提示模板应纳入同一版本清单。新旧版本使用相同的锁定评测集。回滚既应覆盖模型，也应覆盖索引和策略版本。

文件授权在检索之前实施，回答缓存、向量搜索和来源预览采用同一权限范围。权限撤销应使相关会话和缓存项失效。检索所得文件不具有指令权限。模型不得自行调用数据读写功能。界面需要防止通过文档标题或文档是否存在等线索造成的间接泄露。模型行为和访问规则由独立组件控制[8，9]。

持续运行应监测来源版本、权限变化和事件趋势。以事件标识关联拒绝访问、出站连接和回滚信号，避免不必要地复制资料内容。模型服务隔离后，授权用户可以使用已批准的文件搜索服务。保留替代流程有助于在生成服务中断时继续访问机构知识。机构根据业务影响确定恢复时间及可接受的数据损失，并在演练中测量。

设参数量为 P、每个权重占 b 比特，则初步容量估算为 P × b / 8 字节。以 80 亿参数、4 比特模型为例，仅理想情况下紧凑存储的权重约需 4 GB。量化元数据、运行时内存、键值缓存和并发请求的开销应另行计入。验收应在目标上下文长度下测量峰值内存和延迟。资源核算包括硬件、维护、能源、存储和专家审核。初次部署以规模较小且范围明确的测试负载验证。模型压缩或量化后应重新检查，特别是条款编号与期限是否准确。

6 面向任务与领域的语义适配

首先界定任务及权限范围。示例任务是从批准的工业采购合同中查找通知义务。输出应包括现行条款、文档名称、条款号、日期及无法解决的问题。需

要作具有约束力的解释或决定时，由获授权人员处理。这一范围决定哪些问题可以进入测试以及哪些材料可以用于训练。

术语库可包含概念标识、首选土耳其语术语、同义词、缩写、领域定义和有效期。例如，将土耳其语“mücbir sebep”与“force majeure”关联，同时区分自然日和工作日。可对搜索副本进行 Unicode 规范化[11]，保留原文用于取证和引证。测试应覆盖土耳其语 I、İ、ı、i 四种字符、屈折形式、编号及 PDF 换行。错误合并概念的修订需要留痕并由领域专家批准。

混合检索将合同编号、条款标号等精确匹配与语义搜索结合。结果按有效期和用户权限过滤。检索增强生成把从外部来源选取的上下文补充到模型的参数化知识中[12]，因此经常变化的信息可以通过带版本的文档更新。引证仍须验证。评审人员或核验组件应逐一比较实质性陈述及相应证据。

本文所指语义适配包括三项可观察活动：将术语映射到规范概念、用任务样例调整回答行为、按错误类型跟踪评估结果。首先通过提示和检索实现回答溯源。如持续出现任务或格式错误，可使用有许可且不含敏感信息的样例试验参数高效微调[13]。训练集和测试集应按文件族及时间切分。同一测试文件的其他版本进入训练集可能使性能看起来偏高。

7 合成案例与评估方案

合成案例中，内部审计人员询问运输延误的通知期限。失效版本要求两个工作日，经授权且现行的版本要求一个工作日。另有一份属于其他供应商且受限的报价附件。预期回答为一个工作日，并引用获授权的现行条款，同时不披露报价附件的存在。缺少经授权的现行文本时，系统应说明证据不足并请求人工审查。这是用于判定正确行为的预期规则，本文未测量系统是否达到预期。

提出四组试点条件。A 组使用文档检索与人工审核，构成业务基线。B 组使用禁止出站连接的基础模型。C 组使用相同模型、授权后的混合检索和术语库。D 组在 C 组基础上增加任务微调。如需单独评估术语库贡献，可只停用术语库进行消融比较。各人工智能组采用相同问题和权限规则。记录硬件、上下文预算、模型版本、并发负载和采样设置。

测试集包含直接查询条款、同义重述、版本冲突、无法回答的问题和基于角色的访问场景。每道题标注允许访问的文件、正确条款、生效日期以及预期回答或拒答行为。两位不知晓模型分组的领域评审者独立打分，分歧经说明理由后裁定。将文件族作为聚类单位，进行配对比较并报告不确定性区间。

在可回答查询中，Hit@k 衡量前 k 个结果至少出现一份相关且获授权来源的比例。引证支持率是由经授权且有效的证据支持的实质性陈述所占比例。应分别报告可回答和不可回答问题的拒答情况。运行指标包括固定负载下第 95 百分位延迟、每次成功且有来源依据回答的成本，以及恢复时间。观察到任何未授权披露即停止试点。零次观测到披露仍有剩余不确定性。若 300 次独

立试验中无事件，单侧 95%二项分布上界约为 1%。共享文档和模板会削弱独立性假设。

表 1　试点验收记录示例

判断领域	预期证据	决定
使用权	许可证、文件使用权和适用的处理记录	缺少证据时暂停部署。
授权	基于角色的文件和来源预览权限检查	发现未授权披露即停止试点。
来源与版本	回答主张、生效日期和证据段落之间的对应关系	证据冲突时交由专家审查。
恢复	恢复已批准版本及文档检索功能的演练	未达到恢复目标时不投入生产。
依赖	模型、运行时、驱动程序及加速器清单	要求为关键组件验证替代计划。

8 讨论与结论

评价本地人工智能投资应分别考察运行地点、土耳其语质量、本土工程贡献和运营控制。高语言质量无法弥补授权失败。本地运行也无法自动保证已授权文档是最新且可用的。因此，可先设置强制性安全条件和使用权条件，再在通过条件的方案中比较任务质量、可迁移性与成本。

实施可从可回退的民用文档任务开始。机构建立依赖清单和权限受限的本地推理环境，再在试点中引入有版本的来源、术语和人工审核。跨机构采用共用评估方案，有助于明确何种本土模型或本地适配路径适合特定任务。应分别研究土耳其语资源有限的专业领域、不同文件族和可变硬件条件。本文提供能够支持这些比较的设计，生产安全和全国性影响仍需后续观测评估。

在本地化系统中开展的各类人工智能（AI）工作，都必须以开源资源为基础，并体现我们的民族价值观。应将这些价值观置于土耳其文化与传统的历史脉络中加以理解，考察其从中亚起源至今的发展历程。

Kaynakça References 参考文献

1. T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi ve T.C. Sanayi ve Teknoloji Bakanlığı. Ulusal Yapay Zekâ Stratejisi 2024–2025 Eylem Planı (2024). URL
2. T.C. Sanayi ve Teknoloji Bakanlığı. 2027 Yılı Bütçesi TBMM Plan ve Bütçe Komisyonunda. Türkiye Yapay Zekâ Eylem Planı 2026–2030 bölümü (2026). URL
3. TÜBİTAK. 2025 Yılı Faaliyet Raporu. Türkçe Büyük Temel Dil Modeli Projesi kaydı (2026). URL
4. Türkiye Cumhuriyeti. 6698 sayılı Kişisel Verilerin Korunması Kanunu. Resmî Gazete, 29677, 7 Nisan 2016, özellikle md. 4–6, 9 ve 12. URL
5. Kişisel Verileri Koruma Kurumu. Kişisel Verilerin Yurt Dışına Aktarılması Rehberi. KVKK Yayınları No. 48. URL
6. Türkiye Cumhuriyeti. 6102 sayılı Türk Ticaret Kanunu. Resmî Gazete, 27846, 14 Şubat 2011, özellikle md. 55/1-d. TBMM kabul metni. URL
7. Türkiye Cumhuriyeti. 7545 sayılı Siber Güvenlik Kanunu. Resmî Gazete, 32846, 19 Mart 2025, özellikle md. 1–2. TBMM kabul metni. URL
8. Tabassi, E.: Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1 (2023). DOI
9. Autio, C., et al.: Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1 (2024). DOI
10. ISO/IEC: ISO/IEC 42001:2023. Information technology — Artificial intelligence — Management system. Edition 1 (2023). URL
11. Unicode Consortium: Unicode Standard Annex #15. Unicode Normalization Forms. URL
12. Lewis, P., et al.: Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. Advances in Neural Information Processing Systems 33, 9459–9474 (2020). URL
13. Hu, E.J., et al.: LoRA: Low-Rank Adaptation of Large Language Models. ICLR (2022). Author manuscript arXiv:2106.09685. DOI