

<> Türkiye
Açık Kaynak
Platformu</>

Turkey Open
Source Platform

ACIKHACK
Açık Kaynak Hackathon Programı



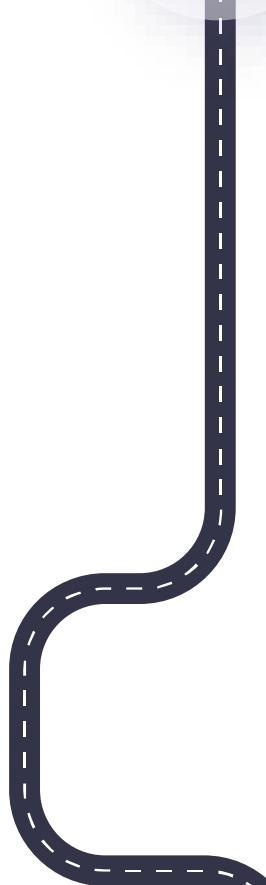
MÜZE ASİSTANI

11.08.2020

Susam SokAĞı Ekibi

İçindekiler

Problem



Biz Kimiz ?

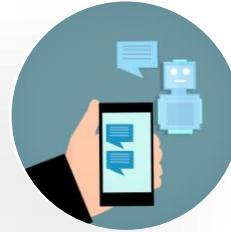
Ekip üyeleri ve görev tanımları



Sorunun Büyüklüğü



Hizmet Eksikliği



Çözüm

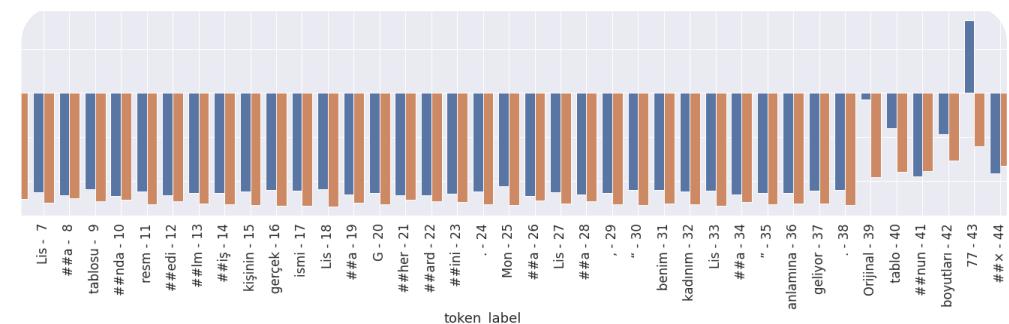
Yenilikçi bir sistem geliştirilmesi

Kullanılan Teknolojiler



Kullanılan model ve mimari yapısından bahsedilerek çalışma yapısının anlatılması

State-of-art



Arayüz ve Demo

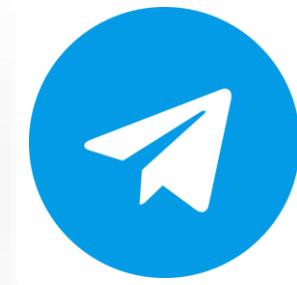


2020 - 2021

Projenin avantajlarından bahsedilerek, gelişime açık olan olası gelecek hedeflerinin bahsedilmesi.



Avantajlar ve Gelecek Hedefler



Susam Sokağı Ekibi

 Türk Hava Kurumu Üniversitesi / Ankara

 Bilgisayar Mühendisliği





Oğuzcan Turan

Derin Öğrenme Geliştirici

Model Geliştirme

Telegram Entegrasyonu

Backend API Geliştirme



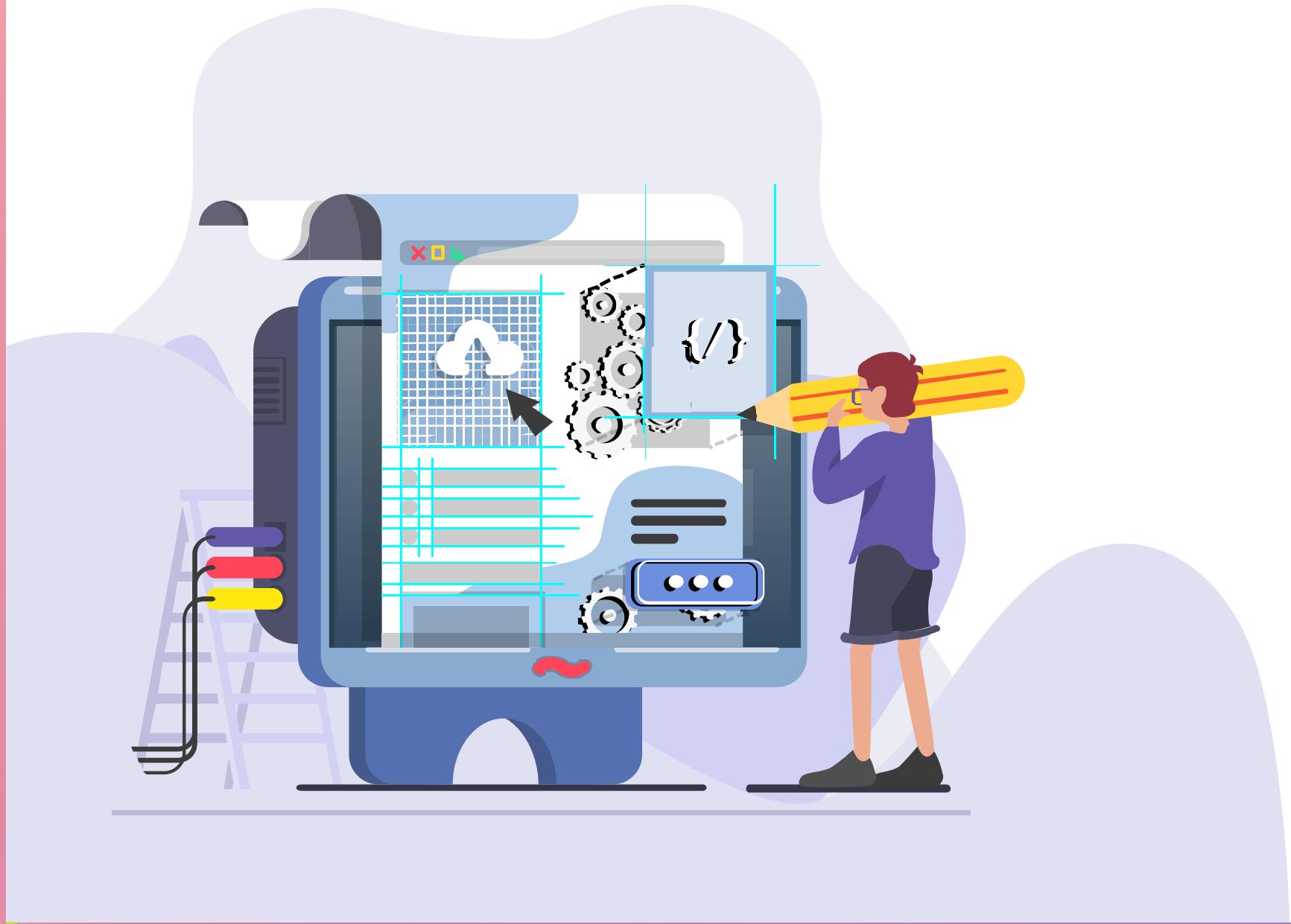
Meva Akkaya

Derin Öğrenme Geliştirici

Model Geliştirme

Telegram Entegrasyonu

Veri Tabanı Tasarımı



Yiğit Burak Akkaş

Front-end Geliştirici

Admin Dashboard Geliştirme

Backend API Geliştirme

Chatbot Arayüz Geliştirme



PROBLEM

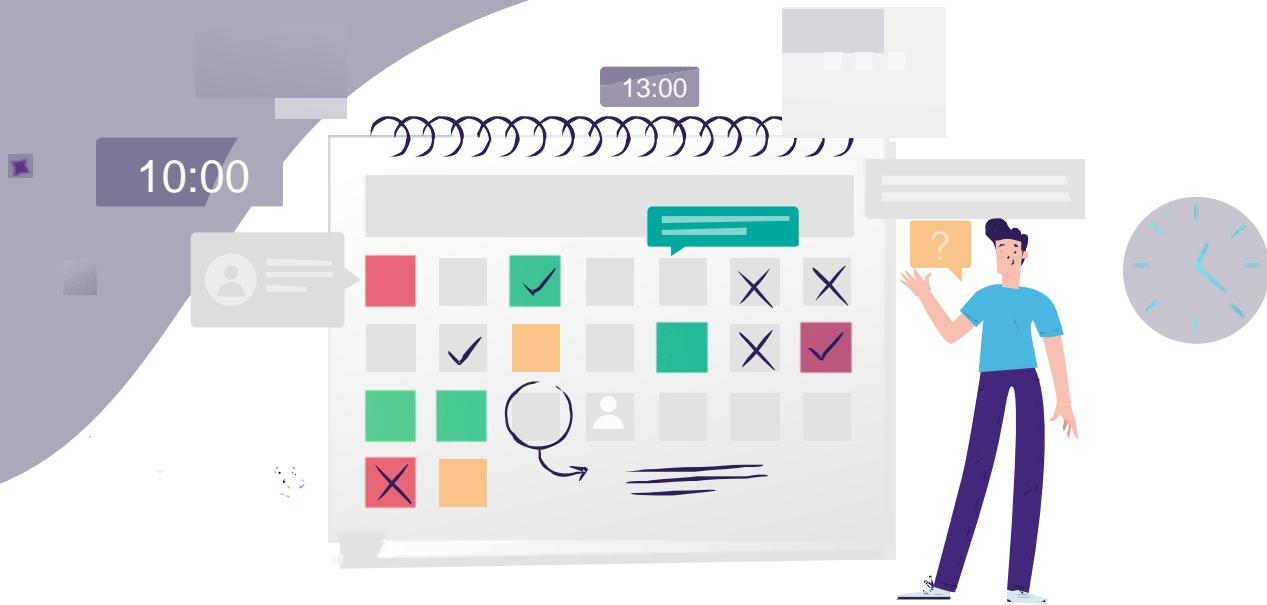
Müzelerde ziyaretçilerin, eserler ve tarihi olaylar hakkında bilgi edinmede sıkıntı çekmesi, bilgi edinmek için kişi veya kaynak bulmakta zorlanması.

Bu durum müze memnuniyetinin düşmesine ve kötü görüşler oluşmasına sebep olmaktadır.



Daha Çok Nitelikli Personel

Özellikle müzelerin en yoğun olduğu gün ve saat aralıklarında, ihtiyacı gidermek için gereken nitelikli personel sayısının fazla olması.



NE YAZIK Kİ...

Günümüzde bir çok kişinin uzun açıklama yazılarını okuyacak,
sesli metinleri dinleyecek yada sorusunu araştırıp bulacak zamanı
olamıyor



Sorun Ortak



55.000
den Fazla Müze

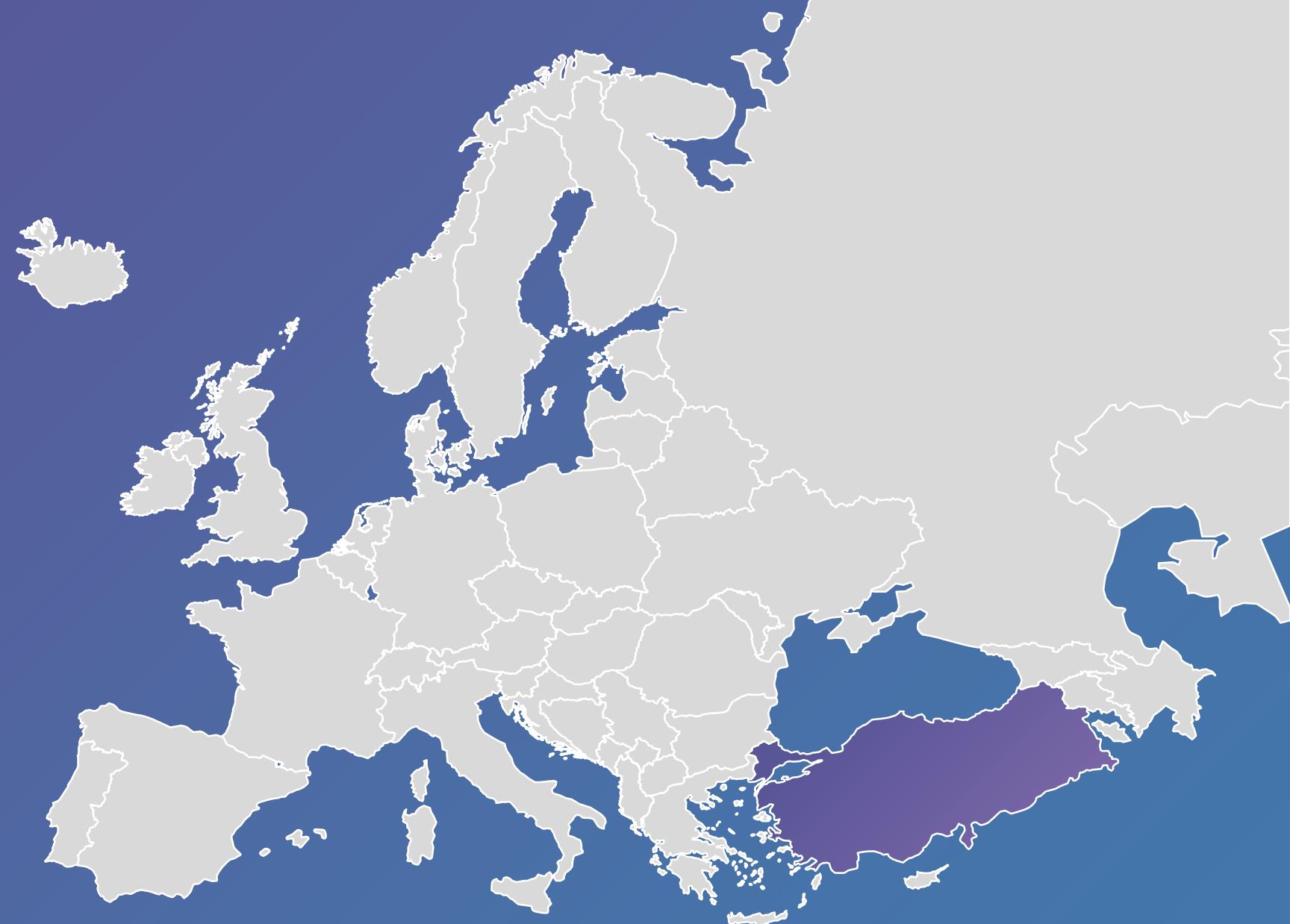


9,600,000
Yıllık Ziyaretçi Sayısı
(Louvre Müzesi, Paris)

202
Ülke

Sadece Türkiye'de
450 den
fazla müze mevcut.

Yılda **30 Milyon** ziyaretçi





“

Neden
sorularımıza
cevap verecek
bir sistem
geliştirmeyelim
ki?



Kodu Tarat

Sorunu Sor

Ziyaretçilerin, eserler hakkında sorularına
anında cevap verecek bir sistem.

i Kolay Erişim

i Hızlı Cevap Alma

i Gelişime Açık

BERT

Müzede !

Google tarafından geliştirilen BERT modelinin soru-cevap özelliğini sistemimizde kullanıyoruz.



Siz Sorun
BERT
Cevaplasın

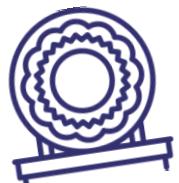
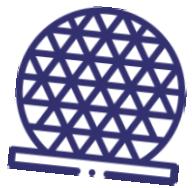




ÜZE



SISTAN



BERT NEDİR?



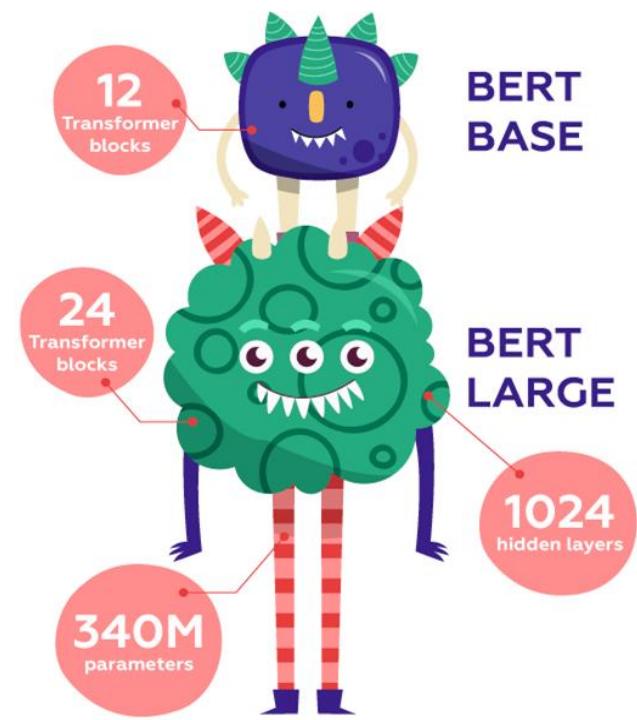
Bidirectional Encoder Representations from Transformers (BERT)

2018 - Google

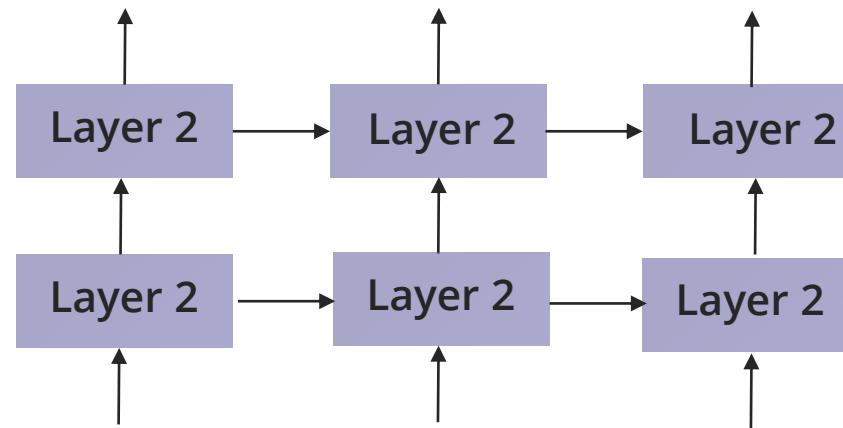
Doğal dil işleme (NLP) için yeni bir yöntemdir.

Bert_large ve bert_base adı verilen 2 temel modelde sahiptir.

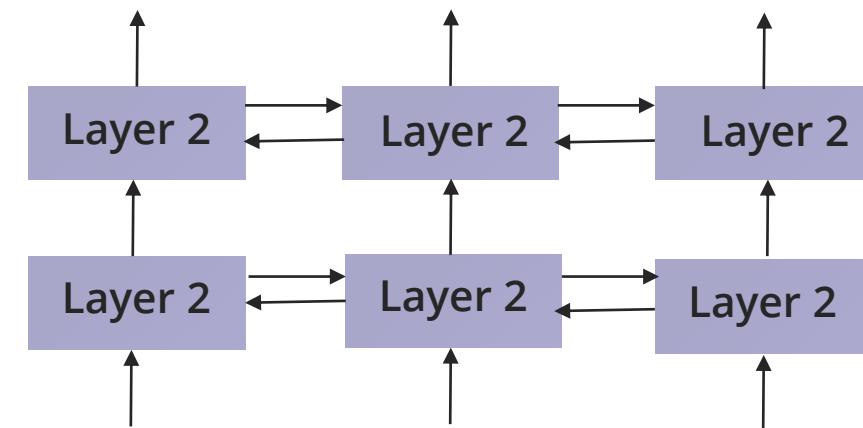
BooksCorpus (800 milyon kelime) ve Wikipedia (2.5 milyar kelime) dahil, toplam 3.3 milyar kelimelik bir korpus üzerinden 40 epoch ile önceden eğitilmiştir.



Tek Yönlü (Unidirectional)



Çift Yönlü (Bidirectional)



- Usain Bolt, çok fazla rekor kırmış bir atlettir. (Sporcu)

- Annesi atlet giymezse çok kızardı. (İç çamaşırı)



PERFORMANS

SQuAD1.1 Leaderboard

Rank	Model	EM	F1
	Human Performance <i>Stanford University</i> (Rajpurkar et al. '16)	82.304	91.221
1 Oct 05, 2018	BERT (ensemble) <i>Google AI Language</i> https://arxiv.org/abs/1810.04805	87.433	93.160
2 Sep 09, 2018	nlnet (ensemble) <i>Microsoft Research Asia</i>	85.356	91.202
3 Jul 11, 2018	QANet (ensemble) <i>Google Brain & CMU</i>	84.454	90.490



Bert performansını diğer son teknoloji NLP sistemleriyle karşılaştırdığımızda, soru cevaplamaya dayalı puanlamada insan performansını aşarak en yüksek skoru aldığı görülmektedir

MODELİN EĞİTİLDİĞİ VERİ

BERT-base (Turkish):

<https://huggingface.co/dbmdz/bert-base-turkish-uncased>

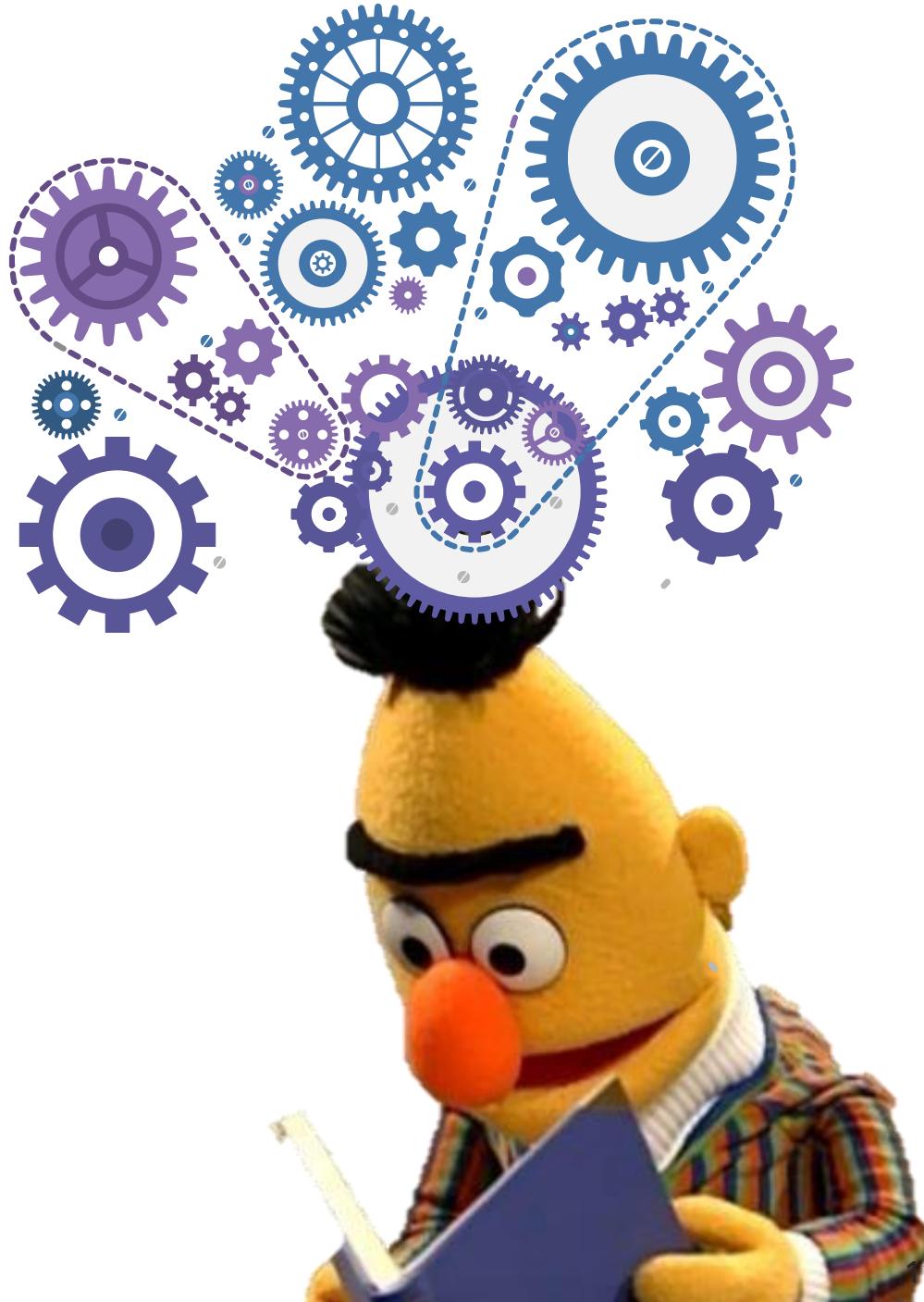
BERT-QA (Turkish):

<https://huggingface.co/savasy/bert-base-turkish-squad>

TQuAD dataset:

<https://github.com/TQuad/turkish-nlp-qa-dataset>

Modelin eğitildiği veri, Teknofest 2018 Yapay Zeka yarışması kapsamında Türk & İslam Bilim Tarihi üzerine oluşturulan Türkçe Soru-Cevap verisidir.



	Başlık / Title	Paragraf / Paragraph	Soru-Cevap / Q&A
Geliştirme / Development	72	275	892
Eğitim / Train	681	2232	8308



Tablonun
boyutları
nedir?

Peki BERT nasıl yardımcı oluyor?

Uzun ve kompleks cümleler içinden doğru cevabı nasıl
bulmaktadır.



Metin:

Sanat tarihçileri için devasa bir konu olan Leonardo da Vinci'nin tablosu Mona Lisa, medeniyetinin sahip olduğu en özel parçalardan biri.

Mona Lisa tablosunda resmedilmiş kişinin gerçek ismi Lisa Gherardini. Mona Lisa, "benim kadınım Lisa" anlamına geliyor. [Orijinal tablonun boyutları](#)

77x53 cm. Mona Lisa tablosunun başka bir adı daha var: La Gioconda. Bu isim ise Mona Lisa'nın "Wife of Francesco del Giocondo" yani Frances del Giocondo'nun Karısı unvanına sahip olması nedeniyle verilmiş. Kesin olmamakla birlikte Mona Lisa'nın gerçekte Lisa del Giocondo olduğu düşünülüyor. Mona Lisa tablosunun herhangi bir sigortası yok. Bunun nedeni de sigortalanamayacak kadar değerli görülmesi. Hiçbir sigorta şirketi bu riske girmek istemiyor yani. Yüz tanımda kullanılan sisteme göre Mona Lisa'nın yüzü %83 mutlu, %9 bikkin, %6 korkmuş ve %2 sinirli mimiklere sahip. Kanada Ulusal Araştırma Merkezindeki bilim adamları tarafından, renkli lazer cihazlarla yapılan bir taramaya göre Mona Lisa'nın gülüşü yeni doğum yapmış kadın gülümsemesiymiş....

..

...

Soru: Tablonun boyutları nedir?

CEVAP: Tablonun boyutları 77x53 cm.



NEDEN BERT?

Pretrain -
Türkçe Dil
Desteği

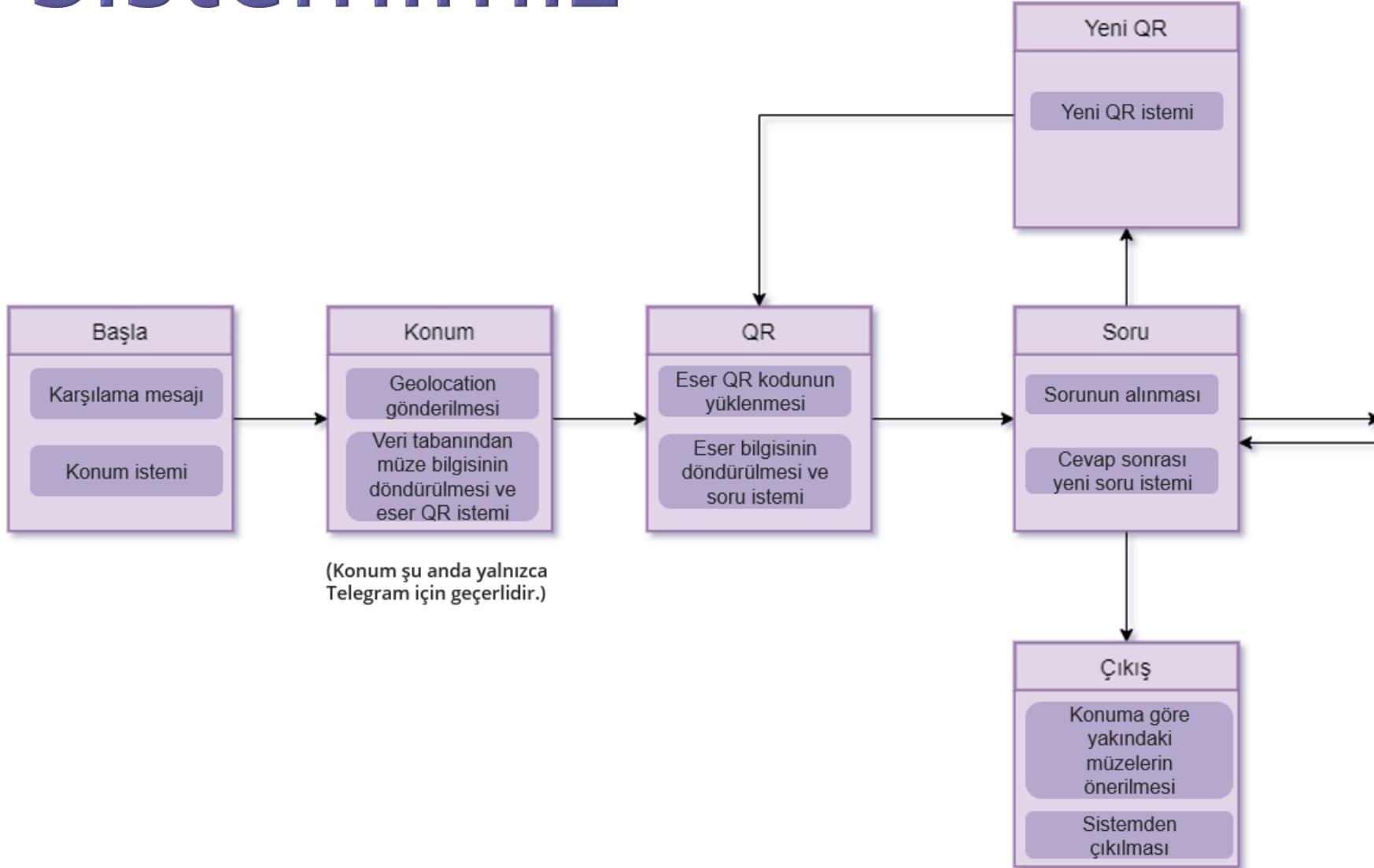
Soru ve
Cevapta
yüksek
performans

Çift Yönlü
Yaklaşım

Uzun ve
Kompleks
Cümleleri
Anlayabilme



Sistemimiz





01

Türkçe Dil Uzmanlığı
Erişim

02

Limitsiz Kullanıcı
Erişim

03

Yüksek Entegrasyon

04

Platform Bağımsız

05

Yüksek Veri Desteği

Bu proje sayesinde bilgilendirme alanında çalışan personelin oluşturduğu
parasal kaynak ihtiyacını kazanca çevirebilmek.

KULLANIM ALANLARIMIZ

1 Müzeler

Ziyaretçilere müzelerde bulunan eserler hakkında bilgi veriyoruz ve onların kafalarındaki soruları cevaplıyoruz.

3 Kütüphane

Kütüphanelerde, ziyaretçilerin aradıkları kitap ile alakalı hızlı bilgi erişimine olanak tanıyoruz.

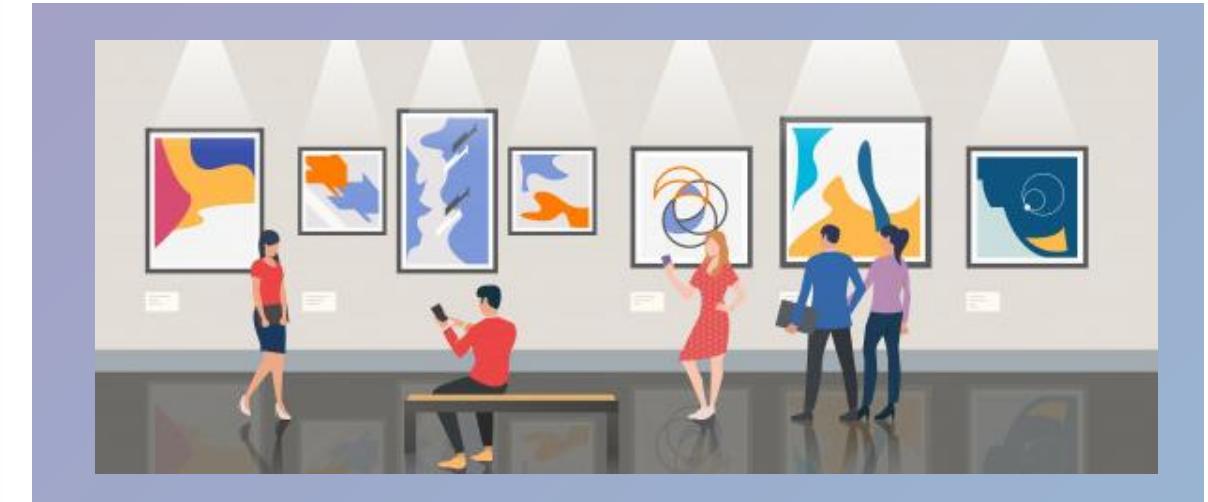
2 Sanat Galerileri

Sanat galerilerinde bulunan sanat eserleri ile alakalı bilgilendirme yapıyoruz.

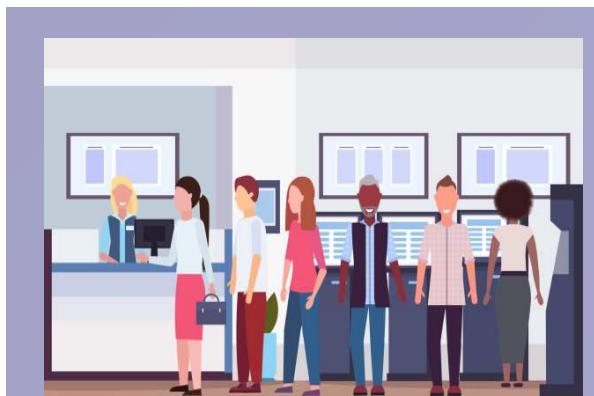
4 Destek Masaları

Yoğun çalışan destek masalarında, sıkıntı çeken müşterilere ilk destek konusunda çalışıyorız.

Nerelerde Kullanılabilir ?



Kütüphaneler



Destek Masaları

PAZAR ANALİZİMİZ

Sayılar İle Türkiye'de Müze Turizmi

64

FARKLI İL

50K

ORTALAMA
YILLIK MÜZE
ZİYARETİ

35+ M

ZİYARETÇİ
35.048.417
ziyaretçi/yıl

174

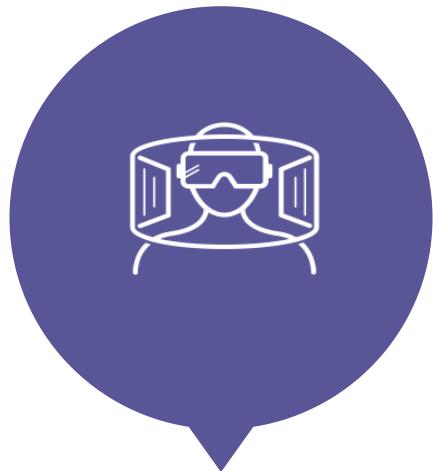
MİLYON TL
YILLIK
GELİR

278

FARKLI MÜZE

RAKİP HİZMETLER

SUSAM SOKAĞI ekibi olarak geliştirdiğimiz Bert destekli soru-cevap analizi ve chat-bot uygulaması, Türkçe dili ile ön eğitimli ve öncelikli pazar hedefimiz Türkçe dili ve Türkiye bulunan müzeler olması sebebiyle bu alanda çalışan müze hizmetleri gösterilmektedir.



360° VR / AR

Yeni ve maliyeti yüksek bir teknolojidir. Toplu kullanımlara uygun bir teknoloji değildir. Yoğun ziyaretlerin olduğu müzelerde kullanımı imkansızdır.



Sesli Dinleti

Müzelerin genelinde en çok tercih edilen, sesli kayıt ile eser hakkında bilgi anlatımı üzerine bir teknolojidir.



Hologram

Yüksek maliyetli bir teknoloji olması ile beraber her eser için kullanılamayan bir hizmettir.

AVANTAJLARIMIZ

Başarılı Olduğumuz Alanlar

Müze asistanı projesi günümüz Covid-19 ve pandemi sürecinde yeni normal standartlarına uygun şekilde ziyaretçilerin temasını en aza indirgeyen kişisel sanal asistan hizmeti vermektedir.

Yüksek Başarı Oranı

Sistemimiz Türkçe ön-eğitimli bir dil modeli ile Türk dil yapısına hakim olduğu için, spesifik alanlarda verilen veriler ile **%90+** bir başarı elde etmektedir.

Yüksek Entegrasyon İmkanı

Telegram Messenger ve Web Arayüzü desteğimiz ile günlük hayatı son kullanıcıya hitap eden her türlü akıllı cihaza entegre edilebilir ve erişim sağlanabilir bir sisteme sahibiz.

Ölçeklenebilir Sistem

Sistemimiz yazılımsal ve tasarımsal yapısındaki başarı ile, yüksek veri boyutları ve yüksek anlık kullanıcı sayılarına sorunsuz hizmet verebilir.

Yakında Gelecek Özelliklerimiz...

Projenin İleriye Dönük Gelişim Planı



Görüntü İşleme

Müzelerdeki eserleri görüntü işleme kullanarak otomatik tanıyan bir sistem güncellemesi.



Veri Toplama Otomasyonu

Otomatize olarak sanat, tarih ve eserler hakkında bilgi barındıran web arşivlerin analiz edilmesi ve veri seti oluşturulması.



Diyalog Bazlı Chatbot

Sistemimizde soru cevap yapımız ile beraber çalışacak diyalog tipi bir chat bot yapısına güncellemek.



Sesli Konuşma Analizi

Entegre ettiğimiz platformlarda çeşitliliğe giderek, sesli komut ve konuşma analizi ile iletişim geliştirmek.



MUSEI VATICANI



55,000
Müze/Dünya



104 Farklı Dilde

BERT dil modelimizin desteklediği 104 farklı dil ile dünyadaki her müzede hizmet verme hedeflerimizden birisi.



202 Farklı Ülkede

202 farklı ülkede, farklı kültürlerde insanların bilgiye erişimini kolaylaştırmak hedefimiz.



55,000 Müzede

55 Bin farklı müze ile milyonlarca eserin incelemesini tutarak global bir arşiv olmak hedefimiz.



中国国家博物馆
NATIONAL MUSEUM OF CHINA

ENTEGRASYON



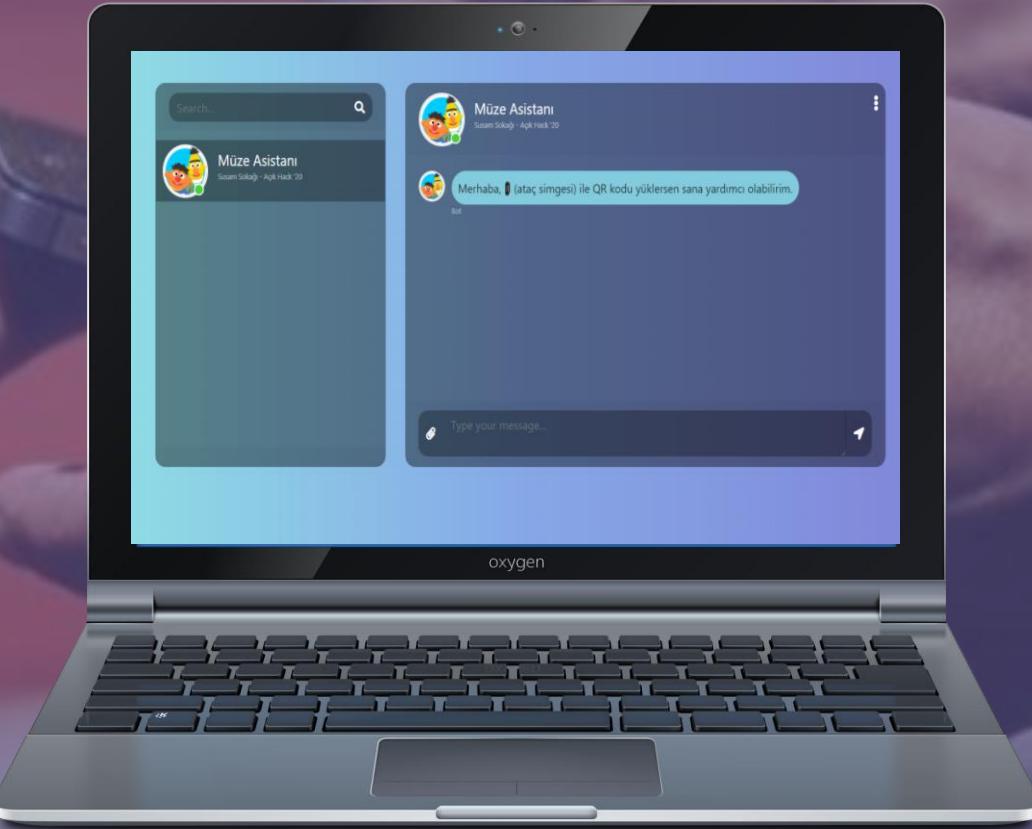
ARAYÜZÜMÜZ

Hızlı Erişim

Tek mesaj ile anında erişim sağlayın ve sınırsız sayıda soru yöneltin.

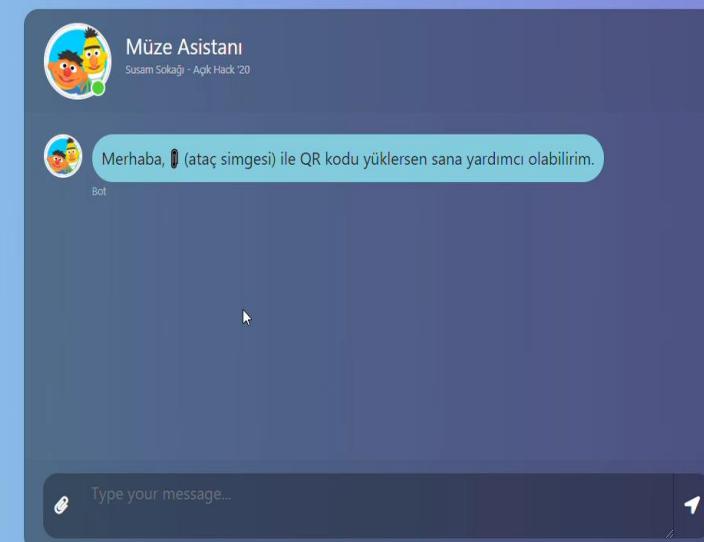
Çoklu Platform Desteği

Telefon, Tablet, Akıllı saat ve her türlü akıllı cihazınızdan erişim imkanı.



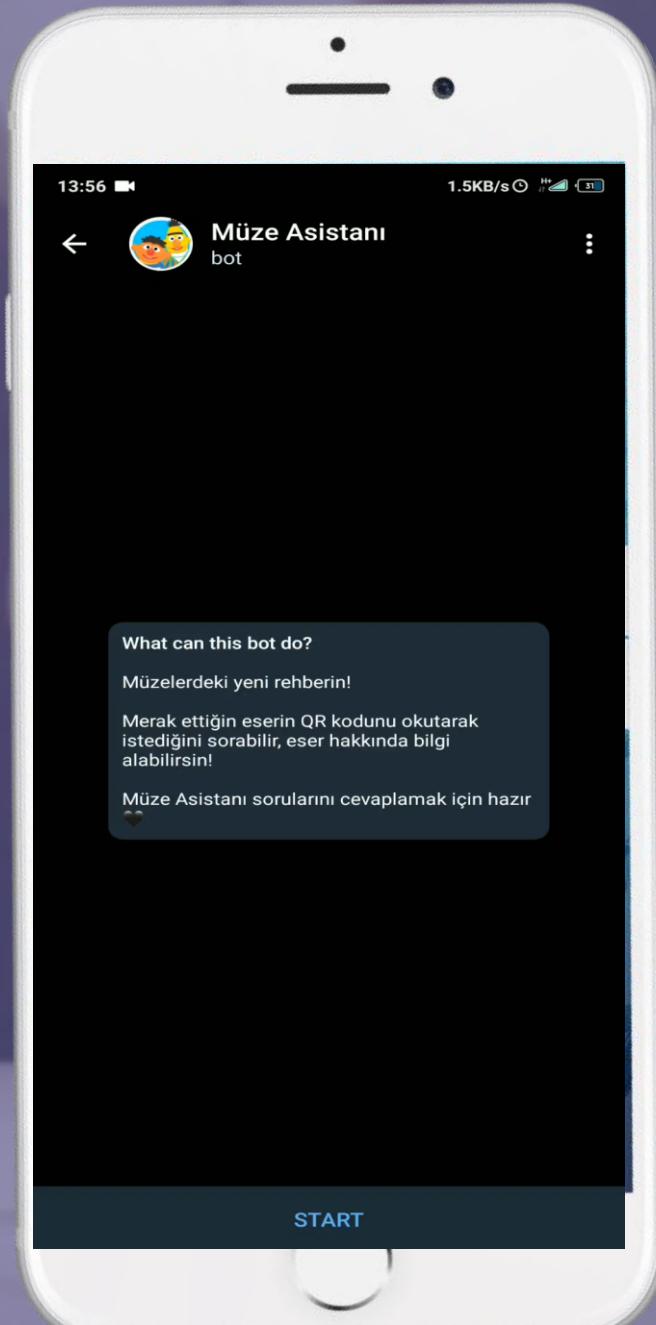
Demo

Web sayfasından
asistan ile
konuşabilirsiniz!



Demo

Telegram bot ile konuşabilirsiniz!



Kodu Tarat !

Mona Lisa Hakkında Merak Ettiklerini Sor

Hemen Dene



t.me/SusamBot





TEŞEKKÜRLER!