

**T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

WIZARD of OZ TEKNİĞİYLE CHATBOT TASARIMI

**OSMAN UTAR
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

**GEBZE
2023**

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

WIZARD of OZ TEKNİĞİYLE
CHATBOT TASARIMI

OSMAN UTAR
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMANI
DOÇ. DR. MEHMET GÖKTÜRK

GEBZE
2023

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

CHATBOT DESIGN WITH THE WIZARD
of OZ TECHNIQUE

OSMAN UTAR

A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE
DEPARTMENT OF COMPUTER ENGINEERING

THESIS SUPERVISOR
ASSOC. PROF. DR. MEHMET GÖKTÜRK

GEBZE
2023



YÜKSEK LİSANS JÜRİ ONAY FORMU

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 09/02/2023 tarih ve 2023/13 sayılı kararıyla oluşturulan jüri tarafından 01/03/2023 tarihinde tez savunma sınavı yapılan Osman UTAR'ın tez çalışması Bilgisayar Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE
(TEZ DANIŞMANI) : Doç.Dr. Mehmet GÖKTÜRK

ÜYE : Doç. Dr. Pınar ONAY DURDU

ÜYE : Dr. Öğr. Üyesi Burcu YILMAZ

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
...../...../..... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

ÖZET

Sohbet robotları, şirketler ve kullanıcılar için değer katmak amacıyla çeşitli iş süreçlerini otomatikleştirmek için kullanılabilir. Sohbet Robotu geliştirme süreci boyunca, kullanıcı geri bildirimlerinin kullanıcı merkezli bir konuşma tasarımı süreci içinde bütünleştirilmesi esastır. Çalışmamızda bilgisayar mühendisliği öğrencilerinin yüksek oranda tekrarlayan ve standartlaştırılmış görevlerle karakterize edilen sorularında sohbet robotları araştırılmıştır. Bu çalışma, bir insan operatör tarafından sohbet robotu işlevselliğini simüle ederek, projenin çok erken bir aşamasında temel bir diyalog konseptinin test edildiği bir Oz Büyücüsü yaklaşımı uygulamaktadır. Bu şekilde, sohbet botu işlevlerini tamamen kodlamadan diyalog tasarımının genel uygunluğu hakkında değerli kullanıcı geri bildirimleri toplanabilmektedir.

Fikrimizi kavramsal olarak doğrulamak ve simüle edilmiş Sık Sorulan Sorular (SSS) sohbet robotunu test etmek için 15'er dakikalık iki denememize toplamda 62 kullanıcı katılmış olup 914 dakika boyunca test etmişlerdir. Sonuç olarak, test edilen konseptin kullanıcılar için uygun ve değerli olduğu kanıtlanmıştır. Önceden tanımlanmış cevapların eksiksizliği ve simüle edilmiş sohbet robotunun yetkinlik kurulumu ile ilgili kullanıcı memnuniyetine ilişkin geri bildirim oldukça makul görülmüştür. Bununla birlikte, katılımcıların çoğu, test edilen senaryoyu alakalı olarak görmüş ve önerilen sohbet robotunun soruları cevaplama sürecinde uygulanması için yüksek bir kullanıcı değeri belirtmişlerdir.

Çalışmamızda Wizard Of Oz metodunun hem tasarlanırken hem de geliştirilirken kullanılabilirliğini, bu şekilde bir yöntemle daha hızlı, etkin ve kullanıcı dostu chatbotlar tasarlanabileceği sonucuna elde edilen kullanıcı memnuniyet anketleri sonuçlarıyla varılmıştır.

Anahtar Kelimeler: Chatbots, Wizard of Oz, AI Röportajcısı, OZ Büyücüsü Çalışması, Chatbot Dizaynı.

SUMMARY

Chatbots can be used to automate various business processes to add value for companies and users. Throughout the Chatbot development process, it is essential to integrate user feedback into a user-centered conversational design process. In our study, chatbots were investigated in computer engineering students' questions characterized by highly repetitive and standardized tasks. This work applies a Wizard of Oz approach where a basic dialog concept is tested at a very early stage of the project by simulating chatbot functionality by a human operator. In this way, valuable user feedback on the overall relevance of the dialog design can be gathered without completely coding the chatbot functions.

To conceptually validate our idea and test the simulated FAQ chatbot, a total of 62 users participated in our two 15-minute trials and tested it for 914 minutes. As a result, the tested concept has proven to be convenient and valuable for users. The feedback regarding the completeness of the predefined answers and the user satisfaction regarding the competency setup of the simulated chatbot was considered quite reasonable. However, most of the respondents saw the tested scenario as relevant and noted a high user value for the proposed chatbot to be implemented in the process of answering questions.

In our study, it was concluded with the results of user satisfaction surveys that the Wizard of Oz method can be used both during design and development, and that faster, more effective and user-friendly chatbots can be designed with this method.

Key Words: Chatbots, Wizard of Oz, Turkish Chatbots, Chatbots Design.

TEŐEKKÜR

Yüksek lisans eğitimimde ve akademik hayatımda desteęini ve yardımlarını esirgemeyip bilgisi ile bu çalışmanın oluşmasının yolunu açan danışmanım Doç. Dr. Mehmet GÖKTÜRK'e ve çalışma sürecimde bana daima destek olan değerli aileme en içten teşekkürlerimi sunarım.



İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
SUMMARY	vi
TEŞEKKÜR	vii
İÇİNDEKİLER	viii
SİMGELER ve KISALTMALAR DİZİNİ	xi
ŞEKİLLER DİZİNİ	xii
TABLolar DİZİNİ	xiii
1. GİRİŞ	1
1.1. Tezin Amacı, Katkısı ve İçeriği	2
2. LİTERATÜR TARAMASI	4
2.1. WOZ deneyi Kavramı	4
2.2. İlgili Çalışmalar	3
2.3. Chatbotların Geçmişi	6
2.3.1. Alan Turing ve Turing Testi	7
2.3.2. Yıllık Loebner Ödül Yarışması	8
2.3.3. Eliza	9
2.3.4. Parry	10
2.3.5. Alice	10
2.3.6. Smarter Child	11
2.4. Modern Zamanlar	11
2.4.1. IBM Watson	11
2.4.2. Siri	11
2.4.3. Alexa	12
2.4.4. Chat -GPT	12
2.4.5. Google Bard	15
3. CHATBOT TASARIMI	16
3.1. Chatbot Mimarisi	16
3.2. IBM Watson Mimarisi	18
3.3. Microsoft Bot Bilişsel Hizmetleri	19

3.4. Chatbotların Deęerlendirme Metrikleri	20
3.5. Chatbotların Sınıflandırılması	22
3.5.1. Kelime/Menü Tabanlı Botlar	23
3.5.2. İfade Tabanlı Chatbotları	23
3.5.3. Diyalog Tabanlı Chatbotlar	23
3.6. Chatbotların Kullanım Amaçları	24
4. WOZ DENEYLERİ	28
4.1. OZ Sihizbazı (WOZ) Röportaj Sistemi	28
4.2. Teorik Arka Plan	31
4.2.1. Chatbot Prototipleme ve Geliştirme	31
4.2.2. Teknolojik Yenilikler için Oz Büyücüsü Deneyleri	32
4.2.3. Chatbot Geliştirme için Oz Büyücüsü Deneyleri	32
4.3. Yöntemsel Yaklaşım	33
4.3.1 Oz Büyücüsü Çalışmasının Hedefleri	33
4.4. Chatbot Oluşturma ve Yapılandırma	35
4.4.1. Çalışma Tasarımı	35
4.4.2. Oz Büyücüsü Deneyinin Kurulumu ve Teknik Altyapı	37
5. BULGULAR ve SONUÇLAR	41
5.1. Elde Edilen Bulgular	41
5.2. Sonuç ve Tartışma	46
KAYNAKLAR	48
ÖZGEÇMİŞ	52
EKLER	53

SİMGELER ve KISALTMALAR DİZİNİ

<u>Simgeler ve</u>	<u>Açıklamalar</u>
<u>Kısaltmalar</u>	
ARPANET	: Advanced Research Projects Authority Net.
AIML	: Anlık Mesajlaşma Programları.
ALICE	: Artificial Linguistic Internet Computer Entity.
AI	: Artificial Intelligence.
BT	: Bilişim Teknolojileri.
API	: Application Programming Interface.
dk	: Dakika.
DM	: Diyalog Manager.
GPT	: Generative Pre-trained Transformer.
HCI	: Human Computer Interaction.
IBM	: International Business Machines.
LUIS	: Dil Anlama Akıllı Hizmetleri.
GTÜ	: Gebze Teknik Üniversitesi.
MIT	: Massachusetts Teknoloji Enstitüsü.
NLG	: Doğal Dil Üretimi.
NLP	: Doğal Dil İşleme.
NLU	: Doğal Dil Anlama.
RLHF	: Reinforcement Learning from Human Feedback.

ŞEKİLLER DİZİNİ

<u>Sekil No:</u>	<u>Sayfa</u>
2.1 : Chatbotların Tarihçe Çizelgesi.	7
2.2 : Turing Testi İllüstrasyonu.	8
2.3 : ELIZA Sohbet Robotunda Tipik Bir Konuşma.	9
2.4 : ALICE Sohbet Robotuyla İle Kısa Bir Konuşma Örneği.	10
2.5 : Chat GPT Kullanıcıları Aralık 2022-Şubat 2023.	13
3.1 : Chatbot Mimarisinin Tipik Bileşenleri.	17
3.2 : IBM Watson Mimarisi.	19
3.3 : Microsoft Bot Framework.	20
3.4 : Dahi.ai 's Chatbot Analizi.	21
3.5 : Dahi.ai Platform Konuşma/Kullanıcı sayısına Bir Örnek.	22
3.6 : Chatbotların Sınıflandırma Şablonu	22
3.7 : Kullanıcı Deneyimi Kalitesi Bakımından Chatbot'lar.	24
4.1 : WOZ Çalışma Yaklaşımına İlişkin Akış Şeması.	36
4.2 : Oz Sihirbazı Kullanıcı Ön Uç Yapılandırması.	38
4.3 : Oz Sihirbazı Sihirbazı Ön Uç Yapılandırması.	39
4.4 : Rocket.Chat Microservis Sistem Mimarisi.	39
4.5 : Çalışmanın Oz Büyücüsü Deneyi Çerçevesi.	40
4.6 : Chatbot Arayüzüne İlişkin Bir Kesit.	40

TABLÖLAR DİZİNİ

<u>Tablo No:</u>	<u>Sayfa</u>
5.1: Woz Deneyi Metrikleri	41
5.2: Chatbot Prototipinin Kullanıcı Deneyimi Deęerlendirmesi (%)	44
5.3: SSS Tercih, Chatbot Prototipinin Kullanıcı Deneyimi Memnuniyeti Derecelendirmesi	45



1. GİRİŞ

Yapay zekâ kavramıyla insanların makinelerle etkileşmesi fikri aynı zamanın ürünleridir. 1950 yılında Alan Mathison Turing, aklındaki “ Acaba makineler düşünebilir mi?” sorusuna cevap bulabilmek amacıyla Imitation game olarak zikredilen oyunu kurgulamıştır[1]. Bu kurguda iki taraf mevcut olup soruyu soran Sorgulayan Taraf ve Soruya günümüzdeki makinelerin (Bilgisayarların) ya da onları taklit eden insan taraftan oluşmaktadır.

Sorgulayan taraf aldığı cevaplara göre cevap verenin bir makine mi yoksa bir insan olup olmadığını tahmin etmeye çalışır ve eğer makine yani bilgisayar verdiği tutarlı cevaplarla sorgulayan tarafı insan olduğuna ikna ederse bu tasarlanan makine ya da bilgisayarın Turing Testi’ni geçtiğini gösterir. Bir başka söyleyişle, bilgisayarın düşünebilme yeteneği olan bir yapay zekâ olduğunu kanıtlamış olur.

İlgili makalesinde Alan M. Turing, testi kurgularken iki taraf arasında gerçekleşen etkileşimin yazılı kurgulamış olup, Turing’in ortaya attığı bu yeni kavram olan yapay zekâ kavramı, bilgisayar bilimine yeni ufuklar açmış ve tarihteki ilk sohbet robotu diye tanımlanan chatbot’un literatüre girmesine ön ayak olmuştur.

Chatbot, örneğin kullanıcıların soru yazmasına ve karşılığında bu sorulara anlamlı cevaplar üretmesine izin vererek, insan benzeri konuşmayı simüle eden yapay olarak akıllı bir sohbet aracıdır[2].

Sohbet botları fikri , Eliza sohbet robotunun ("Doktor" olarak da bilinir) Rogerian bir psikoterapisti taklit etmek için inşa edildiği Massachusetts Teknoloji Enstitüsü'nden kaynaklanmıştır[3,4].

Sohbet diyalogları şeklinde tekrarlayan paydaş (yani müşteriler, ya da potansiyel müşteriler) sorularını otomatikleştirmenin bir yolu olarak sohbet robotları, iç ve dış iletişim süreçlerinde giderek daha fazla uygulanmaktadırlar[4,5]. Bu tür süreçlerin verimliliği artırma potansiyellerini ortaya çıkarmak için, bu otomasyon teknolojisiyle ilgili öngörülen kullanıcıların gereksinimleri ve beklentilerini göz önünde bulundurarak uygun bir konuşma tasarımı konsepti oluşturmak zorunludur[6,7]. Erken kullanıcı geri bildiriminin entegrasyonu, geliştirme sürecinde çok önemlidir[4]. Bu da onu teknoloji geliştirme süreçlerinde yaygın bir uygulama haline getirir[8,9]. Paydaş geri bildirimi ve dolayısıyla işlevsel bir sohbet robotu sistemine sahip olmadan erken bir aşamada bir sohbet robotunun oluşturulması ve geliştirilmesi için gerekli girdiyi

sağlamının bir yolu, bir Oz Büyücüsü (WOz) deneyi yapmaktır[10]. Bir WOz testinde, uygulayıcılar test deneklerini tamamen gelişmiş bir teknolojik sistemle etkileşime girdiklerine inanmaya yönlendirirken, bu şekilde hareket eden test operatörlerinin kendileridir, bu durumda insan formunu gizleyen sohbet robotu görevi görür[11,12]. Çalışmamızda, tercihler ve genel kullanım için bir SSS sohbet robotunun pratik örneğinde, aşağıdakileri yapmak için daha geniş bir sohbet botu kullanıcı testi senaryosunda bir WOz deneyi yapılmıştır:

- Geliştirilen tercih SSS sohbet robotu prototipinin niyet veri tabanını alaka düzeyi ve cevap uygunluğu açısından değerlendirmek,
- Konuşma tasarımı ve özellikle ön içerik, algılanan kullanıcı memnuniyeti, kullanıcının kabul düzeyi ve kullanım sınırlamaları hakkında geri bildirim toplamak,
- Yeni sohbet robotu amaçları biçiminde alakalı içerik ve ayrıca sohbet robotu için potansiyel eğitim verileri sağlamaktadır.

1.1. Tezin Amacı, Katkısı ve İçeriği

Çalışmamızda Bilgisayar Mühendisliği adayı öğrencilerinin yüksek oranda tekrarlayan ve standartlaştırılmış görevlerle karakterize edilen sorularında sohbet robotları araştırılmıştır.

Bu çalışma, bir insan operatör tarafından sohbet robotu işlevselliğini simüle ederek, projenin çok erken bir aşamasında temel bir diyalog konseptinin test edildiği bir Oz Büyücüsü yaklaşımı uygulamaktadır.

Bu şekilde, sohbet botu işlevlerini tamamen kodlamadan diyalog tasarımının genel uygunluğu hakkında değerli kullanıcı geri bildirimleri toplanabilmektedir.

Ayrıca bu deneylerin üç kapsayıcı hedefi bulunmaktadır.

- Geliştirilen tercih SSS sohbet robotu prototipinin niyet veri tabanını alaka düzeyi ve cevap uygunluğu açısından değerlendirmek,
- Konuşma tasarımı ve özellikle ön içerik, algılanan kullanıcı memnuniyeti, kullanıcının kabul düzeyi ve kullanım sınırlamaları hakkında geri bildirim toplamak ,

- Yeni sohbet robotu amaları biiminde alakalı ierik ve ayrıca sohbet robotu iin potansiyel eėitim verileri saėlamaktadır.



2. LİTERATÜR TARAMASI

Bu bölüm, chatbot tasarımı konusundaki mevcut araştırmaları derleyip değerlendirir. Chatbot'un kullanımlarına ilişkin kısa bir açıklama yapılmaktadır. Literatürdeki mevcut chatbot tasarım yaklaşımları tanıtılmaktadır.

2.1. WOZ Deneyi Kavramı

Maulsby ve diğerleri. Bir otomasyon ajanı ile WOz testi üzerine bir çalışma yürüten, sihirbaz için katı bir davranış planının önemini vurgulamaktadır (hatta bir algoritma uygulanmasını tavsiye etmektedirler)[14]. Bu, tutarlı davranışı ve dolayısıyla deneysel güvenilirliği sürdürmek için önemlidir[14]. Genel olarak, aşağıdaki roller atanmıştır:

-Katılımcı:

Başvuru sürecinde chatbot ile sohbet eden, potansiyel adayların hedef grubuna ait, işe alınan chatbot kullanıcıları.

-Sihirbaz

Deneysel çalışma çerçevesine dayalı olarak ayrı bir odadan/uzaktan uzaktan yapay zeka tabanlı gibi görünen otomatik yanıtlar olarak önceden formüle edilmiş mesajlar göndererek veya geçici yanıtlar oluşturarak WOz çerçevesini çalıştıran araştırma projesine ait bir araştırmacı.

-Moderatör

Deney boyunca katılımcıya eşlik eden ve süreç boyunca bir giriş, yönergeler ve rehberlik sağlayan, araştırma projesine ait başka bir araştırmacı.

2.2. İlgili Çalışmalar

2017'de Zamora, bir hafta boyunca altı sohbet botundan birini kullanan 54 kişinin çalışmasını yayınlamıştır[15]. Kullanıcıların bazılarının telefonlarında sahip oldukları Siri'yi kullandığı; diğerlerinin ekonomi haberleri ve sosyal medyayı kontrol etmek, yemek sipariş etmek ve sinema bileti rezervasyonu yapmak için halk tarafından

erişilebilen, görev odaklı sohbet robotları kullandılar. Makale başlığının da belirttiği gibi - "Üzgünüm Dave, korkarım bunu yapamam: Chatbot algısı ve beklentileri" beklentileri karşılanmadı. Kullanıcılar yavaş yanıtlar, istenmeyen bildirimler bildirdiler ve hız veya verimlilikte kazanç elde etmemişlerdir. Siri kullanıcılarının yanı sıra, birkaç kişi sohbet botunu kullanmaya devam edebileceklerini söylediler.

Jain vd. 8 Facebook Messenger sohbet botunun her birini 3 gün boyunca her gün denemek için nispeten teknoloji meraklısı ve ilk kez kullanan kullanıcıların kaydını yaptırdı[16]. Yazarlar, Chatbottle'ın haber, seyahat, alışveriş, sosyal, oyun, yardımcı program, sohbet ve eğlence olmak üzere 8 alanın her birinde en yüksek puan alan Messenger sohbet robotunu seçti. Her sohbet botunun 1000'den fazla beğeni almıştı. Botlar ve katılımcılar 9968 mesaj alışverişinde bulundu. Temel bulgu: "kullanıcıların beklentileri karşılanmadığı." oldu. Kısa da olsa aktif bir denemeydi. Bir sonraki çalışma, en azından akıllı asistanlar için ilk deneyimlerin genellikle belirleyici olduğunu göstermektedir.

Luger ve Sellen, "The gulf between user expectation and experience" alt başlıklı bir çalışmada, Siri, Cortana ve Google Now'ın mevcut kullanıcılarını inceledi[17]. Akıllı asistanların genellikle beklentileri karşılayamadıkları ve kullanıcıların bunları düşük, düşük seviyeli görevler için kullanmaya karar verdikleri sonucuna varmışlardır. Kullanıcılar birkaç kez yeni görevleri denediklerini ve başarısız olduğunda vazgeçtiklerini bildirdi.

Luger ve Sellen bir ikileme karşılaştı: Kullanıcılardan biri, başlangıçta yazılımı eğlenceli bir egzersiz olarak denediğini bildirdi. Ara sıra komik programlanmış bot yanıtlarını beğendiler. Ancak yazarlar, mizahın akıllı asistanların sunamayacağı insan benzeri nitelik beklentileri yarattığı sonucuna vardı. Tasarımcılara eğlenceli cevaplar vermemelerini tavsiye ettiler.

2018 yılında yapılan bir başka çalışmada 337 yeni çalışana birkaç ay boyunca dağıtılan bir insan kaynakları SSS chatbotu üzerine çalışmasında bunun tersi sonuca varıldı[18]. Yazarlar mizah içeren dikkatli bir tasarım önermişlerdi, önceki bir chatbot kullanıcılarının konjenial konuşmadaki mizahsız çabalarını görmezden geldiğini ve SSS cevaplarını aramak için anahtar kelimeler yazmaya başvurduğunu belirtti. Mizah da bazı insanları sanal yoldaşlara çekmektedir.

Sekiz kişilik bir ekip 18 aylık bir toplantı planlama, sohbet botu, calendar.help oluşturmak, test etmek ve düzenlemek için 18 ay harcadı. Talepler üç aşamada işlendi:

-Seviye 1

Tamamen yazılım tarafından ele alınanlar,

-Seviye 2

Sistemin kapasitesini aşan rutin sorunlar ('mikro görevler') bunları gerçekleştirmek için basit komut dosyaları kullanan kişiler tarafından ele alınmıştır,

-Seviye 3

Zor vakalar uzmanlara aktarıldı

Ekip, mümkün olduğunda mikro görevleri otomatikleştirmek ve zor görevleri daha kolay işlenen mikro görev setlerine ayırmak için yükselmeleri analiz etti.

Sistemin dördüncü versiyonu Seviye 1 ve 2'deki programlama taleplerinin % 39'unu işleyebilmiş ve tüm süreçlerin yaklaşık% 5'i insan yardımı olmadan tamamlanmıştır[19].

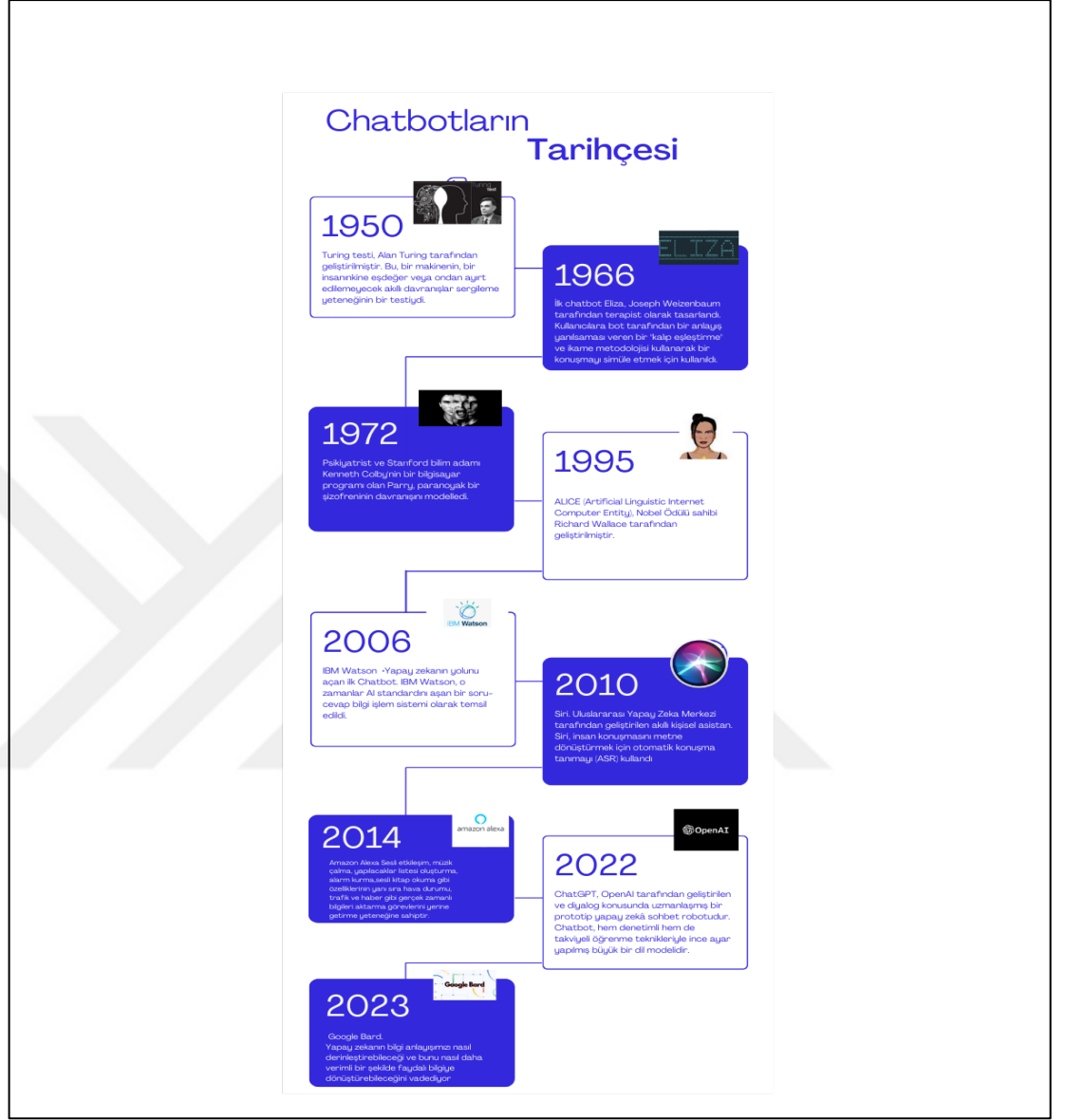
Bir insan yardımcısı sohbet botunun her biri 30 dakikalık, metin tabanlı bir konuşmada 53 gerçek iş başvurusunda bulunan kişiyle görüştüğü bir Oz Büyücüsü alan çalışmasıyla yapmışlardır.

Sohbet transkriptlerinin ve kullanıcı geri bildirimlerinin ayrıntılı bir analizi, kullanıcıların sohbet botundan hoşlandıkları ve hoşlanmadığı şeylerin yanı sıra sohbet botu ile etkileşimlerinin paternlerini bulmaya çalışıyorlar. , NLP veya konuşma yetenekleriyle bile akıllı görünen etkili, gerçek dünyadaki chatbot görüşmecilerini karşılaştırıp, gerçek dünyadaki chatbot görüşmecileri oluşturmak için bir dizi pratik tasarım önerisi formüle etmişlerdir. Bu tasarım önerilerine dayanarak, iş görüşmeleri ve müşteri görüşmeleri gibi çeşitli görüşme bağlamlarında işlev görebilen chatbot görüşmelerinin yapılabileceğini bulgularına eklemişlerdir[20].

2.3. Chatbotların Geçmişi

Chatbot tarihi, yirminci yüzyılın altmışlı yıllarının başlarında başlar (Şekil 1). Chatbot icadı, yirmi birinci yüzyılın başında bir durgunluğa tanık oldu. Bununla birlikte, son yıllarda bir konu olarak chatbot, kar amacı güden ya da kar amacı gütmeyen kurum ve kuruluşlar için en çok aranan hizmetlerden biri haline geldi.

Sonraki bölümlerde, erken çağlar ve modern zamanlar için chatbot tarihi açıklanmaktadır.

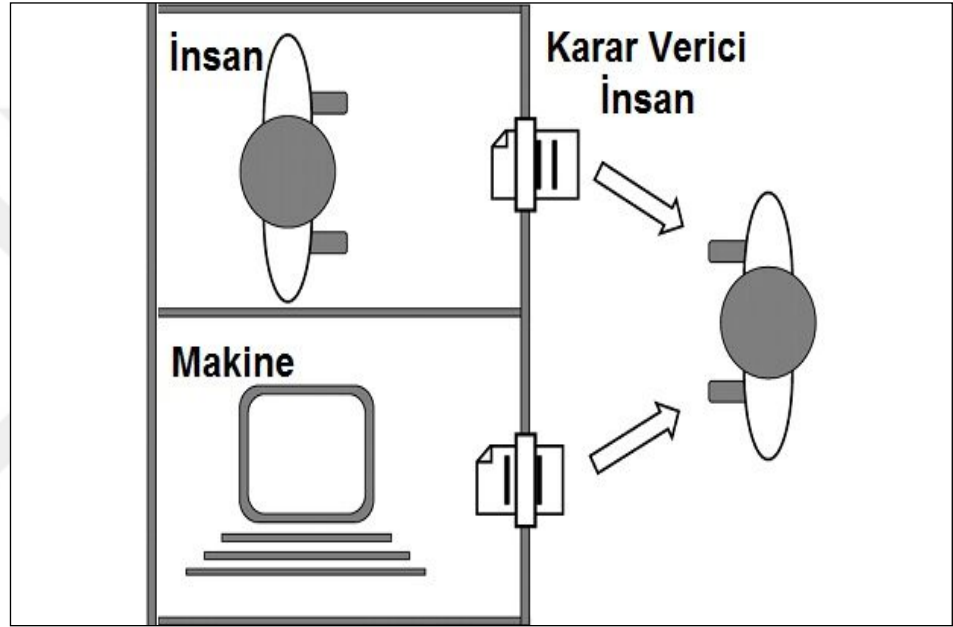


Şekil 2.1: Chatbotların Tarihçe Çizelgesi.

2.3.1. Alan Turing ve Turing Testi

1950'de İngiliz matematikçi Alan Turing, "Makineler düşünebilir mi?" adlı makale, yirminci yüzyılın çığır açan makalelerinden biriydi[1]. Gazete, şu anda "Turing testi" olarak bilinen "Taklit Oyunu" adlı bir oyun başlatmıştı. Test, bir

makinenin bir insanı taklit edip edemeyeceğini bulmayı amaçlıyordu. Test, makinenin yeterince zeki olup olmadığına karar vermek için belirli kriterler belirledi. Bu nedenle, bir sonuca varmak için test, Turing'in makalesinde tanımladığı gibi üç tarafı içeren konuşmaya dayalı bir yaklaşım getirdi: bir sorgulayıcı ve iki bilinmeyen özne: bir bilgisayar ve bir insan. Sorgulayıcının görevi, hem insanlara hem de bilgisayarlara sorular yazarak kimliklerini belirlemektir (Şekil 2). Bu, sorgulayıcının bilinmeyen iki özneyi ayırt edememesi durumunda bilgisayarın Turing testini geçmesinin tek yoludur (Turing, 1950).



Şekil 2.2 : Turing Testi İllüstrasyonu.

2.3.2. Yıllık Loebner Ödül Yarışması

1990 yılında, mucit ve iş adamı Dr. Hugh Loebner, dönüş testi kısıtlamalarını somutlaştıran Loebner ödüllü yarışmasını başlattı. Yarışmanın amacı, AI uzmanlarına ve onların sohbet robotlarının insanlara benzediğini gösterme konusundaki uzmanlıklarına meydan okumaktı. Yarışmanın en üst sıraları 100.000\$'lık büyük ödülü ve bir altın madalyayı kazanır. Loebner, yanıtları insanlardan ayırt edilemeyen ilk bilgisayar için 100.000 \$ ve bir Altın Madalya ile büyük ödülü belirledi. Bununla birlikte, birçok iddia, rekabeti dönme testi kısıtlamalarına uyma konusunda

güvenilmez olmakla suçluyor. Bu iddialar kimin katıldığı ve kimin katılmadığı arasında tartışmalıdır[21].

2.3.3. Eliza

Profesör Joseph Weizenbaum, bilimsel yaşamı boyunca, 1964'te MIT Yapay Zeka Laboratuvarı'nda Eliza adını verdiği bir yazılım programını tanıttı. Eliza, bir senaryoya dayalı Rogerian bir psikoterapisti temsil ediyordu. Rogerian psikoterapist, hastalarına akıl sağlıklarıyla ilgili sorularını yeniden ifade ederek yardımcı olmak için tasarlanmıştır[23]. Aşağıda tipik bir konuşma gösterilmektedir.

```
=====
EEEEEEEE L      IIIIII ZZZZZZZ      AAA
E         L      I      Z      A      A
E         L      I      Z      A      A
EEEEEE   L      I      Z      A      A
E         L      I      Z      A      A
E         L      I      Z      A      A
EEEEEEEE LLLLLLL IIIIII ZZZZZZ      A      A
=====
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA. WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE... !
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...
=====
```

Sekil 2.3: ELIZA Sohbet Robotunda Tipik Bir Konuşma.

Eliza, kullanıcı girişi, sohbet botuna sağlanan önceden tanımlanmış anahtar kelimeleri karşılıyorsa verimli bir şekilde tepki verdi. Eliza kullanımı anlamadığında, bilgiler önceden programlanmış verilerle eşleşene kadar kullanıcıdan daha fazla bilgi almaya çalıştı. Eliza'nın sorduğu sorular sohbeti uzatmayı amaçlıyordu; böylece kullanıcı diyalogu sürdürmek ve daha uzun süre etkileşimde bulunmak için motive olur. kullanıcının daha fazla etkileşimde bulunmasını sağlar.

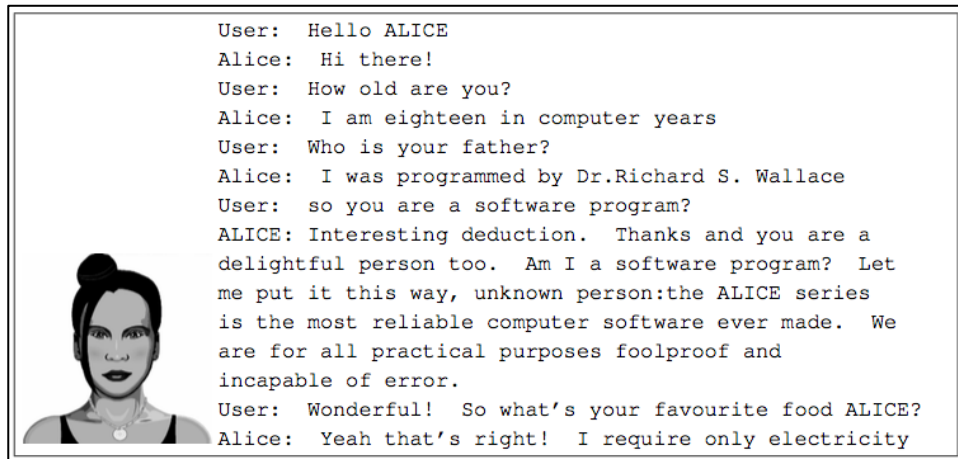
Prof Joseph, insan ve makineler arasındaki ilişkinin yüzeysel olduğuna içtenlikle inandığı için insan taklidi yaratmayı amaçladı. Beklenmedik bir şekilde Eliza ile etkileşime giren insanlar öznel olarak duyguları makineye atfeder.

2.3.4. Parry

Eliza'nın yanında bir başka klasik chatbot daha var: Parry. Psikiyatrist Kenneth Colby onu Stanford Üniversitesi'nde yarattı. PARRY'nin amacı paranoid şizofrenik konuşmayı taklit etmektir[24]. Aynı yıl, Washington D.C.'deki Uluslararası Bilgisayar İletişimi Konferansı, Eliza ve Parry'nin sohbet ettiği bir etkinlik düzenledi. Programlanma biçimleri nedeniyle çıktının tahmin edilebilir olmasına rağmen, Eliza ve Parry toplantılarına farklı vesilelerle devam ettiler.

2.3.5. Alice

1995 yılında yapay dilbilimsel internet bilgisayar varlığı anlamına gelen ALICE'in yaratılması, chatbot pazarında devrim yarattı. ALICE'in başarısı, AIML'nin kullanılmasının bir sonucuydu. Alice kalıpları, önceki sohbet robotlarında önemli bir yükseltme olduğunu kanıtladı ve insana en çok benzeyen sistem ödülünün yanı sıra birkaç Loebner yarışması kazandı. Bununla birlikte, AIML'nin sınırlaması, chatbot'un uzun bir süre sohbeti sürdürme kapasitesinin olmaması nedeniyle nihai dönüş testini geçememesiyle sonuçlandı[25]. Alice konuşmasının bir örneği Şekil 4'te sunulmuştur.



Şekil 2.4: ALICE Sohbet Robotuyla ile Kısa Bir Konuşma Örneği.

2.3.6. Smarter Child

2000 yılına gelindiğinde Activebuddy Inc., AIM ve MSN Messenger'da yayınlanan SmartChild adlı bir bot geliştirdi. Başlangıcından bu yana SmarterChild, doğal dil ile giderek daha popüler hale gelen anlık mesajlaşma ve SMS ağları arasındaki entegrasyonu gerçekleştirmeyi amaçlayan 30 milyondan fazla mesajlaşma kullanıcılarını kendine çekti. Bot, hava durumu tahminleri, stok güncellemeleri, film saatleri veya diğer yararlı bilgiler gibi birçok hizmeti sağlama yeteneğine sahipti. Bot, kişiliğine sahip olarak farklılaştırıldı. Ancak SmartChild, müşterinin teknoloji büyümesiyle uyumlu hale getirilmesi gereken ihtiyaçlarını karşılayamadı.

2.4. Modern Zamanlar

2.4.1. IBM Watson

Modern çağda IBM Watson, yapay zekanın yolunu açtı. IBM Watson, o zamanlar AI standardını aşan bir soru-cevap bilgi işlem sistemi olarak temsil edildi. Bu başarı, IBM Watson tarafından kullanılan teknolojinin bir sonucuydu. 2011 yılında 'Jeopardy' dizisinin en başarılı adayları, 1 milyon dolarlık ödülü kazanan Watson'a yenildi. Watson, son yıllarda genişlemeye ve büyümeye devam etti[26].

2.4.2. Siri

Yoğun bir rekabet çağında, Apple Inc. 2010'da Siri'yi tanıttı. Siri, Uluslararası Yapay Zeka Merkezi tarafından geliştirilen akıllı kişisel asistan. Siri, insan konuşmasını metne dönüştürmek için otomatik konuşma tanımayı (ASR) kullandı, ayrıca komutları ve kısa dikteleri dönüştürme becerisine ek olarak: "sabah için alarmı ayarla". Siri, yazıya dökülmüş metni kullandığında doğal dil işleme standartlarını aşmayı başardı[27].

2.4.3. Alexa

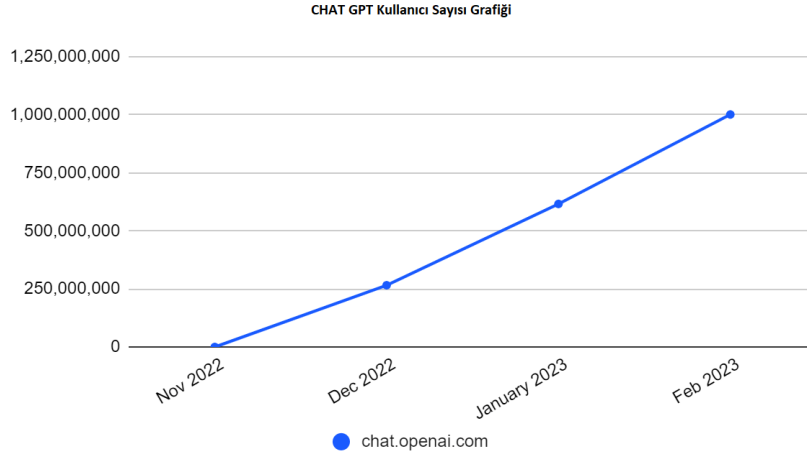
Piyasada iyi bilinen bir diđer chatbot ise Amazon'un kişisel asistanı Alexa'ydı. Alexa'nın yapılacaklar listesi, hava durumu raporu verme ve alarm kurma, evdeki diđer cihazları kontrol etme, "Alexa, oturma odası fanını aç" gibi birçok işlevi vardır. Yıllar geçtikçe Alexa, yeteneklerini geliştirmeyi başardı ve Hunt of Yeti of Alexa oyunları gibi farklı kullanıcı deneyimleri oluşturdu.

Her geçen yıl pazarda birkaç akıllı sohbet robotu yerini aldı. Bu chatbot'lar, müşteri deneyimini iyileştirmek ve yüksek beklentileri karşılamak için kullanıcıya sunduđu faydalı hizmetler yelpazesıyla fark yaratıyor. Microsoft'tan Cortana, Google Asistan, yaygın örneklerden bazılarıdır. Farklı dedektörlerden konum, dokunma, mimik vb. bilgileri toplayabilir, sistemde saklayabilir ve gerektiğinde kullanabilirler.

2.4.4. Chat-GPT

2015 yılında bir araştırma laboratuvarı olarak birkaç araştırmacı ile kurulan OpenAI, GPT1, 2 ve 3 sonra 30 Kasım 2022 tarihinde ChatGPT'yi dünyaya tanıtmıştır. Kullanıcıların tüm sorgularına ayrıntılı cevap vermek ve pek çok farklı görevi yerine getirmek için tasarlanan ChatGPT, insan dilini öğrenmek ve çok yönlü yapay zekâ asistanları gibi etkileşimlerde bulunmak amacıyla kullanılmaktadır.

ChatGPT; bir dil modelinin yapısına dayanan, verilen bir cümleyi okuyarak ve dilin yapısını anlayarak cevap veren bir sistemdir. Kendisine yöneltilen soruları insan gibi yanıtlayan ve kısa süre içinde anlayıp anlamlı cümlelere dönüştüren ChatGPT, başlangıçta yalnızca İngilizce hizmet verse de günümüzde Türkçe dil seçeneğini de desteklemektedir. Özelliklerinin bazıları programlama yapabilmesi, özgün şiir/roman/öykü ve deneme yazabilmesi, gelişmiş soru-cevap sistematiğine sahip olması, önerilerde bulunabilmesi, problemlere uygun bilgisayar kodu yazabilmesi ve kullanıcıları yönlendirebilmesidir.



Şekil 2.5: Chat GPT Kullanıcıları Aralık 2022-Şubat 2023.

Araştırmalara göre, kullanıma sunulduğu ilk ay içerisinde 57 milyon aktif kullanıcıya ulaşan ChatGPT, günde yaklaşık 13 milyon bireysel kullanıcı tarafından ziyaret edilmektedir. Bunun yanı sıra, 100 milyon kullanıcıya en hızlı ulaşan uygulama olarak rekor bir başarıya imza atmış ve bu noktaya ulaşması 1-2 ay süren Google+'ı geride bırakmıştır[28].

ChatGPT, InstructGPT ile aynı yaklaşımları benimsiyen, ancak veri toplama konusunda birkaç farklılık içeren İnsan Geri Bildiriminden Pekiştirmeli Öğrenme (Reinforcement Learning from Human Feedback-RLHF) kullanmaktadır. RLHF, makine öğrenmesi algoritmalarını geliştirmek için insan geri bildirimlerini kullanmaya odaklanan bir yapay zekâ alt alanıdır. RLHF'de insan, makine öğrenmesi algoritmasına geri bildirim sağlar ve bu geri bildirim modelin davranışını ayarlamak için kullanılır. İlk modelin eğitilmesi için "Denetimli İnce Ayar" kullanılmış ve insan yapay zekâ eğitmenleri, her iki tarafı da (kullanıcı ve bir yapay zekâ asistanı) temsil eden diyaloglar sunmuşlardır. Eğitmenlerin yanıtları oluşturmaları için modele bağlı fikirler iletilmiş ve geliştirilen yeni konuşma veri seti, diyalog formatına dönüştürülerek InstructGPT veri seti ile birleştirilmiştir.

İnsan geri bildirim, modelin eylemlerini ödüllendirmek, cezalandırmak veya model parametrelerini ayarlamak gibi çeşitli şekillerde sağlanabilmektedir. RLHF'nin amacı, performansları ve karmaşık görevleri çözme becerilerini geliştirmek için insan uzmanlığını ve bilgisini makine öğrenmesi algoritmalarına dahil etmektir.²⁷ Bu nedenle GPT-3.5 serisi olarak adlandırılan modeller, 2021'in 4.çeyreğinden önce metin ve kod karışımı üzerinde eğitilmiş bir model serisidir. GPT-3.5, GPT-3'ü temel

alır. 'İnsan Geri Bildiriminden Pekiştirmeli Öğrenme' adlı bir kavram kazandırılan GPT-3.5, Denetimli İnce Ayar süreci ile eğitilmiş InstructGPT olarak da adlandırılmaktadır.

Pekiştirmeli Öğrenme davranışçılıktan ilham almış, öznelerin belirli bir ortamda maksimum ödül alması için hangi eylemleri yapması gerektiğiyle ilgili olan bir makine öğrenmesi yaklaşımının adıdır. Güçlendirmeli Öğrenme, belirli bir durum için mümkün olan en iyi yolu veya davranışı seçme olasılığını en üst düzeye çıkarmak için yazılım temsilcilerinin nasıl harekete geçmesi gerektiğini anlatan bir makine öğrenimi alanıdır. Yeni Bir Dönem Bot sanatının durumu, 1960'ların başlarında JCR Licklider tarafından henüz var olmayan bir alanı, insan-bilgisayar etkileşimini özetleyen ileri görüşlü makalelerde ortaya konan bir yolu izliyor. Licklider, MIT'de ARPA yöneticisi olarak internete dönüşen ARPANET (Advanced Research Projects Authority Net)'i öngören ve finanse eden bir psikolog ve mühendisti. İnsanların ve dijital teknolojinin etkileşiminde üç dönem öngördü. İlki, konuşma tanıma ve dil anlama dahil olmak üzere giriş cihazlarının ve ekranların iyileştirilmesine ayrılacak bir HCI planıydı. Son dönem, yapay zeka meslektaşlarının 1980 yılına kadar geleceğine dair güvence verdiği ultra akıllı bilgisayarlar tarafından yönetilecek. Arada, insanların ve bilgisayarların ortak olarak çalışacağı dönem olacak. 1960 yılında, "birçok insan-makine sistemi var. Ancak şu anda insan-bilgisayar ortak yaşamı yoktur." İnsan-bilgisayar simbiyozu vardır[17]. Bugün insan-bilgisayar ortakyaşamı var. Bu hiçbir şekilde eşit bir ortaklık değildir, ancak yazılım bizim adımıza günün her saati özerk olarak çalışır. Etkileşimli botlar bunun bir parçasıdır. Bunları inşa ederken kendimiz ve nasıl çalıştığımız, bilgi ve merakımızın kapsamı, dilin incelik ve rutin olarak üstlendiğimiz görevlerin sayısı ve karmaşıklığı hakkında çok şey keşfederiz.

Yarım asırdan fazla bir süre önce tasarlanan yetenekli makineler, sanal arkadaşların, akıllı asistanların ve göreve yönelik sohbet robotlarının unsurlarını sentezlemelidir. Bu girişimin kapsamını keşfediyoruz. Licklider'in ilk dönemindeki çalışmaları tamamlamak 50 yıl sürdü: insan-bilgisayar etkileşiminin temelini oluşturmak. Licklider, yeni adım attığımız insan-bilgisayar etkileşiminin ikinci çağının 10 ila 500 yıl sürebileceği ve "insanlık tarihindeki entelektüel açıdan en yaratıcı ve heyecan verici dönem olması gerektiği" sonucuna vardı. AI tabanlı Sohbet botları inşa ederken potansiyel olarak ilginç bir sorun da kişiselleştirme ile gizlilik arasındadır. Kullanıcılar, botun onlar hakkında önemli şeyleri hatırlaması ve onları tanıması gerektiğini hissetti. Ancak, özel görüşmeler de muhtemelen gizli kalmalıdır.

Ortaya çıkan bir soru, kullanıcı özel bir şey olduğunda botu nasıl belirtmelidir? Örneğin, tarayıcılar için olduğu gibi botlar için de "gizli" mod olmalı mı? Ya da bir botun geçmişini, tıpkı bir tarayıcı gibi temizleyebilmeli mi - örneğin, "son 30 dakika içinde söylediğim her şeyi unut" mu? Bir başka ilginç soru, bir botun sadece tek bir kişiliğe sahip olup olmamasıdır.

Duygu analizi, müşterilerin nasıl hissettiği ve şirketin pozitifliği iyileştirmek veya artırmak için neler yapabileceği konusunda bir işletmeye rehberlik etmek için bireysel sohbetleri ve tüm sohbetlerin genel temasını izleyebilir.

2.4.5. Google Bard

Google Bard, ChatGPT gibi rakip yapay zekanın (AI) ardından başlatılan yeni bir sohbet robotu aracıdır. Bir insanla yapılan konuşmaları simüle etmek için tasarlanmıştır ve sorabileceğiniz sorulara gerçekçi ve yararlı yanıtlar sağlamak için doğal dil işleme ile makine öğreniminin bir kombinasyonunu kullanır. Bu tür araçlar, büyük destek personeli ekiplerini işe almadan müşterilerine doğal dil desteği sunmak isteyen veya Google'ın kendi arama araçlarını genişletmek isteyen küçük şirketler için özellikle yararlı olabilir. Bard'ın amacı, kullanıcılarla doğal bir şekilde etkileşim kurarak onlara yardımcı olmak ve onları eğlendirmektir.

Bard ayrıca, kullanıcılarla ilginç konular hakkında sohbet edebilir ve hatta şaka yapabilir. Bard'ın, kullanıcılara nasıl bir deneyim sunacağı, kullanıcının sorularına verdiği yanıtların doğruluğuna ve kalitesine bağlıdır.

Bard'ın geliştirilmesi, Google'ın derin öğrenme teknikleriyle geliştirdiği NLP teknolojisi sayesinde mümkün oldu. Bu teknoloji, milyonlarca kelime ve cümle örneği üzerinde çalışarak, Bard'ın doğal bir dille konuşmasına ve anlamasına yardımcı olur.

Sonuç olarak, Bard Google'ın geliştirdiği bir sohbet robotudur ve doğal dil işleme teknolojisi sayesinde kullanıcılara doğal bir şekilde yardımcı olur. Bard, yapay zeka teknolojilerinin müşteri hizmetleri alanındaki kullanımı hakkında önemli bir örnek teşkil etmektedir.

3. CHATBOT TASARIMI

Bu bölümde Chatbot mimarisi, Chatbot Metrikleri ve Chatbotların kullanım amaçları işlenmiştir.

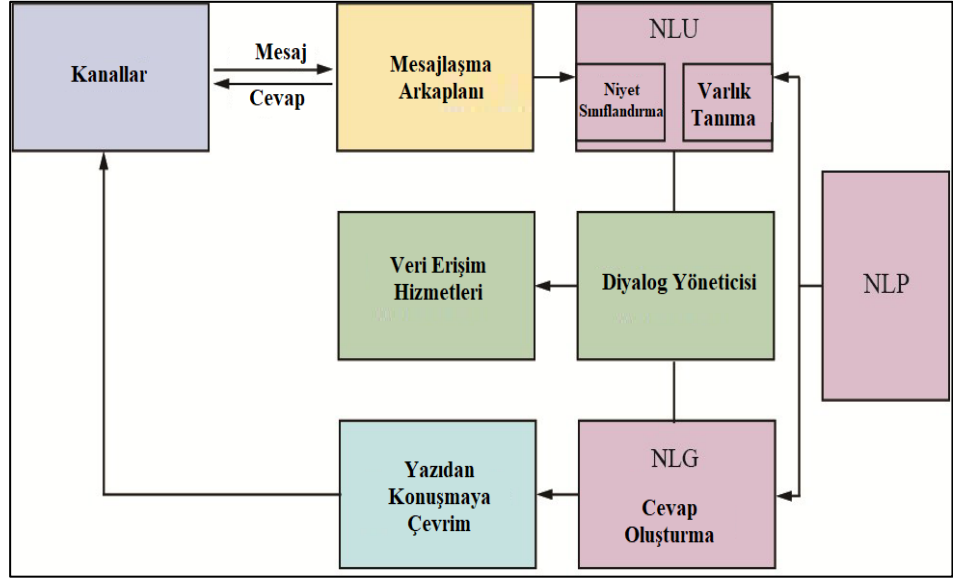
3.1. Chatbot Mimarisi

Firmalar ve organizasyonlar için tasarımın önemi hiçbir zaman bugün olduğu kadar kabul görmemiştir. Ulrich ve Eppinger'e göre tasarım, bir ürün veya hizmetin fiziksel formunu tanımlarken en önemli unsurdur[29].

Chatbot tasarımı, başka bir ürün tasarlamak kadar önemlidir. Mevcut araştırma, çok sayıda araştırmacının teknik bakış açısına ve bilgisayar bilimine dayalı tasarım modelleri önerdiğini fark etti[30]. Bunun bir sonucu olarak, bu bölüm chatbot tasarımı ile ilgili bazı kavramları ortaya koymaktadır.

Tipik bir sohbet robotu mimarisi üç ana bileşenden oluşur: Şekil 6'da gösterildiği gibi NLP (Doğal dil işleme), NLU (Doğal Dil Anlama), Diyalog Yöneticisi (DM), Doğal Dil Üretimi (NLG). bilgisayarların büyük miktarda veriyi işlemesini sağlayan önceden programlanmış bir dizi algoritma. NLP, sohbet robotları ve otomatik olarak oluşturulan anahtar kelime sekmeleri gibi hizmetlerde bulunur. Chatbot'lar için NLP, doğal dil konuşmasını anlamak ve onu yapılandırılmış verilere dönüştürmek için kullanılır. NLU ve NLG, NLP'nin küçük bir parçasıdır. NLU'nun NLP'den daha dar bir amacı vardır. NLU, kullanıcı mesajını anlamaya ve onu amaçlara ve varlıklara göre sınıflandırmaya odaklanır. NLG, kullanıcıya insan benzeri tepkiler üretir.

Bu bileşenler, bir chatbot oluşturmak için zorunlu kabul edilir; bunun yanı sıra, sohbet robotu mimarisinde ASR gibi birkaç isteğe bağlı öge vardır[31]. NLU motoruna sağlanan girdi metin veya ses olabileceğinden, ASR sesi işler ve konuşmayı metne dönüştürür. NLU motoru, amaç sınıflandırıcı ve varlık tanımadan oluşur. Niyet sınıflandırıcısı, sohbet robotları için NLU sistemlerinde çok önemli bir unsurdur ve kullanıcının niyetini derinden anlamak ve doğru cevabı sağlamak kâlfetli bir iştir.



Şekil 3.1: Chatbot Mimarisinin Tipik Bileşenleri.

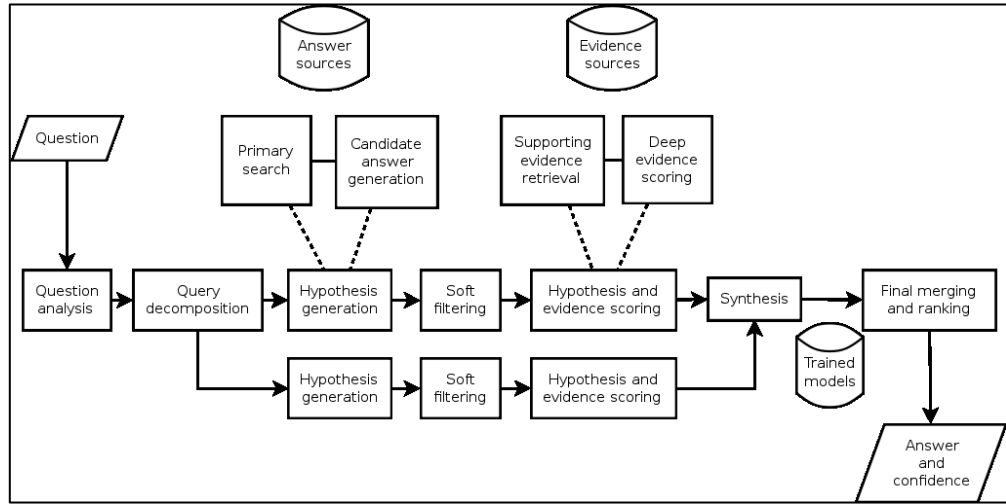
Daha sonra DM, veri bilgi tabanı, uygulama sistemleri veya API gibi veri erişim hizmetlerine dayalı olarak kullanıcı girişinin gerçek bağlamını işler. DM işlevi, insan benzeri yanıtlar sağlamak amacıyla konuşmayı yönetmektir. DM'nin karşılaştığı, konuşma bağlamı, alan doldurma, içerik değiştirme ve inisiyatif gibi farklı zorluklar vardır. Switch deyimi, kullanıcı girişimlerini içerir; bu tür DM, kullanıcı çabası üzerine kuruludur. Switch deyimi, normal bir arama motorundan ziyade kişiselleştirilmiş bir konuşmayı temsil ettiği için SSS için uygulanabilecek verimli bir yöntemdir. Slot doldurma yöntemi, bot kullanıcının varlığını tanıyamadığında kullanılır.

Doğal dil üretimi (NLG), istatistiksel verileri insanların DM'ye dayalı olarak anlayabileceği metne dönüştüren bir yazılım sürecidir. Programlaması en uygun yanıt belirler ve onu tekrar doğal dile çevirir. Response Generator, kural tabanlı sistem, Geri Alma Tabanlı sistem, Model tabanlı buluşsal yöntemler, üretken sistem gibi farklı modellere dayalı yanıtlar sağlar. Kural tabanlı sistem, insan tanımlı kurallara göre yanıt verir, bot, eşleşen giriş kalıplarını veya anahtar sözcükleri kullanarak yanıt verir. Alice, kullanıcılara yanıtları iletmek için AIML etiketlerini kullanan kural tabanlı bir sohbet robotu olarak[32]. Alma tabanlı sistem, önceden tanımlanmış yanıtlardan oluşan bir havuz kullanır. Sistem, önceden tanımlanmış bir konuşmadan en uygun kanıtı seçmeyi amaçlar ve bunu kullanıcının yeni sorgusuyla eşleştirir[33]. Alma tabanlı sistemler, verilen iki mesajın benzerlik temsilini kullanan LSTM modeli gibi farklı modellerden oluşur[34]. Önceki sistemden farklı olarak üretici sistem, kullanıcı tarafından sağlanan bilgilerle sınırlı değildir. Aksine, bağlama veya konuşmanın

mevcut durumuna göre yanıtlar üretir (Shang ve diğerleri, 2015). Üretken tabanlı sisteme bir örnek, nöral makine çevirisi alanında ortaya çıkan diziden diziye modeldir. Model, metin özetleme ve soru cevaplama gibi ilgili diziden diziye tahmin problemlerinde başarılı olduğunu kanıtladı. Mevcut bot çerçevelerinin çoğu, temel olarak, mevcut uygulamalarla entegre olan doğal dil arayüzleri üzerine inşa edilmiştir ve bu da onların karşılıklı konuşmasını sağlar. Piyasada artan sayıda chatbot çözümü var. IBM Watson ve Microsoft Bot platformları en sık kullanılan botlardır. Sonraki paragraflarda, sırasıyla IBM Watson ve Microsoft Bot hakkında kısa bir açıklama sunulmaktadır.

3.2. IBM Watson Mimarisi

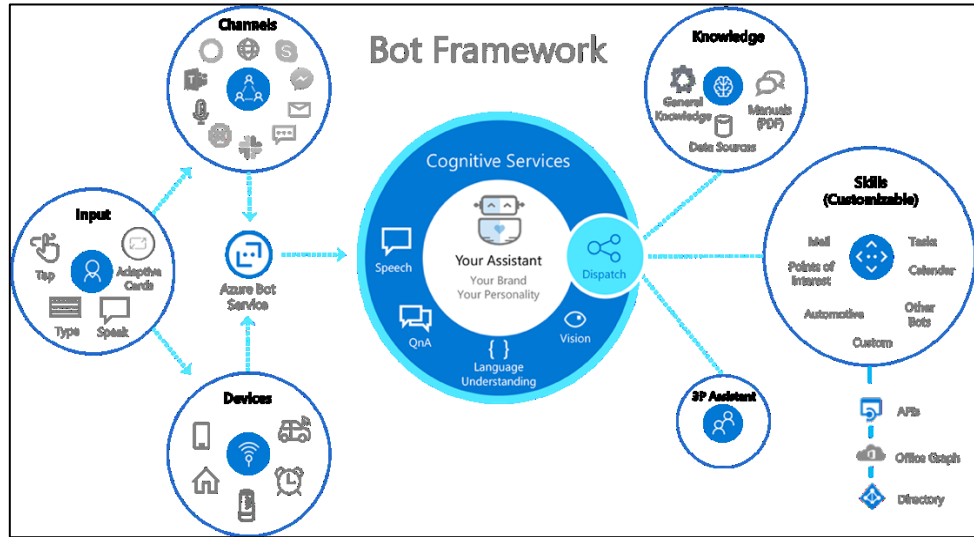
IBM, DeepQA (Derin İçerik Analizi) soru yanıtlama teknolojisi (Ferrucci, 2009) adı verilen Watson sistemini geliştirdi. DeepQA, derin ayrıştırma, anlamsal rol etiketleme ve adlandırılmış varlık tanıma gibi teknikleri içeren, NLP'yi özümseyen bir yazılım bilgi alma işlemidir. IBM, bir hipotez oluşturmak için ansiklopediler, sözlükler, taksonomiler ve ontolojiler gibi bilgi kaynaklarını kullanır. Watson, kullanıcı sorularını alır ve mevcut veri tabanında bir yanıt arar. Arama teknikleri, belge arama, pasaj arama ve metin arama motorlarını içerir. Belge aramada Watson, arama sorgusunun metnini içeren belgenin başlığını hipotez olarak seçerken, pasaj aramada arama sorgusunun metnini içeren pasajları kullanır. Bir sonraki adım, her bir hipotezin güven düzeyini belirlemek için kullanılan hipotez ve kanıt puanlamasıdır. Birleştirme ve sıralama aşaması, güven puanlarının işlendiği ve sıralı liste yanıtlarının üretildiği yerdir. IBM Watson yazılımı, hem yapılandırılmamış kaynaklarda hem de yapılandırılmış kaynaklarda olası yanıtları değerlendirir ve bu yanıtlar için puan kanıtları toplar (Şekil 7).



Şekil 3.2: IBM Watson Mimarisi.

3.3. Microsoft Bot Bilişsel Hizmetleri

Microsoft bot çerçevesi, çeşitli NLU bileşenlerinden oluşur ve alma mekanizmaları modeline dayalı olarak Microsoft bilişsel hizmetleri, bilgisayarla görme, NLP, konuşma, Bilgi çıkarma ve web aramasında AI algoritmalarını kullanır. API entegrasyonu ile Microsoft bot, sürekli öğreniyor ve gelişiyor. Bilişsel hizmetler, Doğal dili işlemek için önceden oluşturulmuş veya özel olarak eğitilmiş dil modellerini kullanan Dil Anlama Akıllı Hizmetleri (LUIS) gibi dili anlamak için API'ler sağlar. Microsoft bot tarafından kullanılan başka bir API, konuları, anahtar ifadeleri ve duyarlılığı keşfeden Metin Analizi API'sidir. İkinci hizmet, adlandırılmış varlıkları algılamak ve kişiselleştirilmiş bir SSS hizmeti gibi farklı araştırma türlerini aramak için API'leri kullanan bilgi ayıklamayı içerir. Bilgi elde etmek için başka bir araç, AI teknolojilerinden yararlanarak kullanıcı sorularını yanıtlamak için geliştirilen RESET API ve web tabanlı hizmet olan QnA Maker'dır. Üçüncü hizmet, konuşmadan metne ve metinden konuşmaya dönüştürme için algoritmalar kullanan konuşma tanıma ve konuşmadır. Bot, bazı API'ler aracılığıyla web sayfaları, resimler ve videolar gibi farklı web arama kaynaklarına erişebilir (Şekil 8).



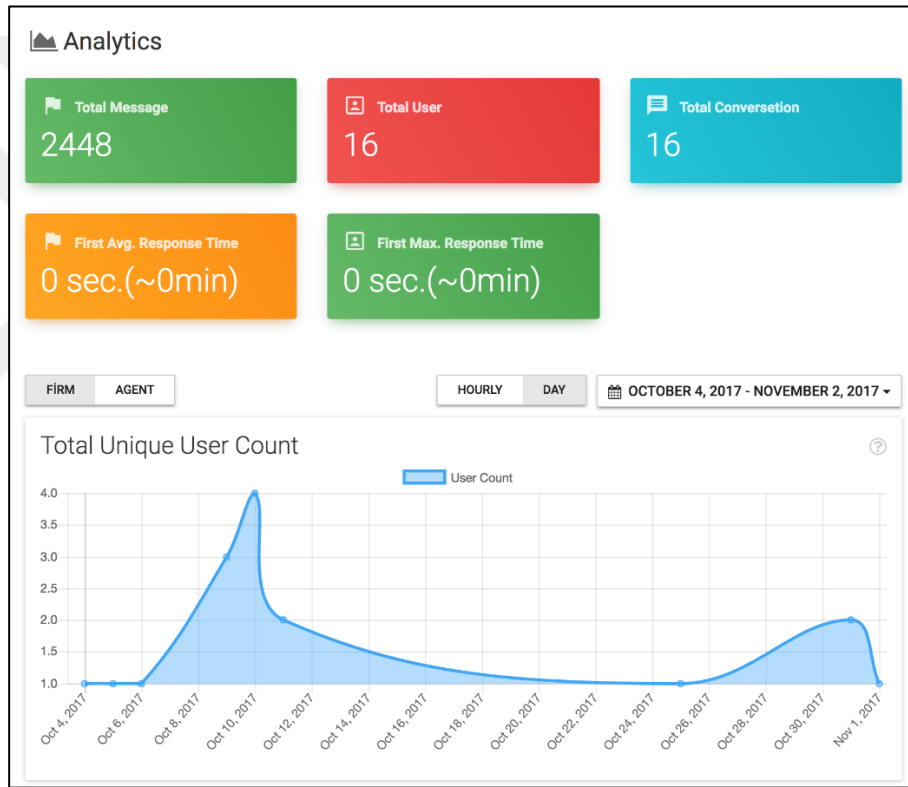
Şekil 3.3: Microsoft Bot Framework.

3.4. Chatbot Değerlendirme Metrikleri

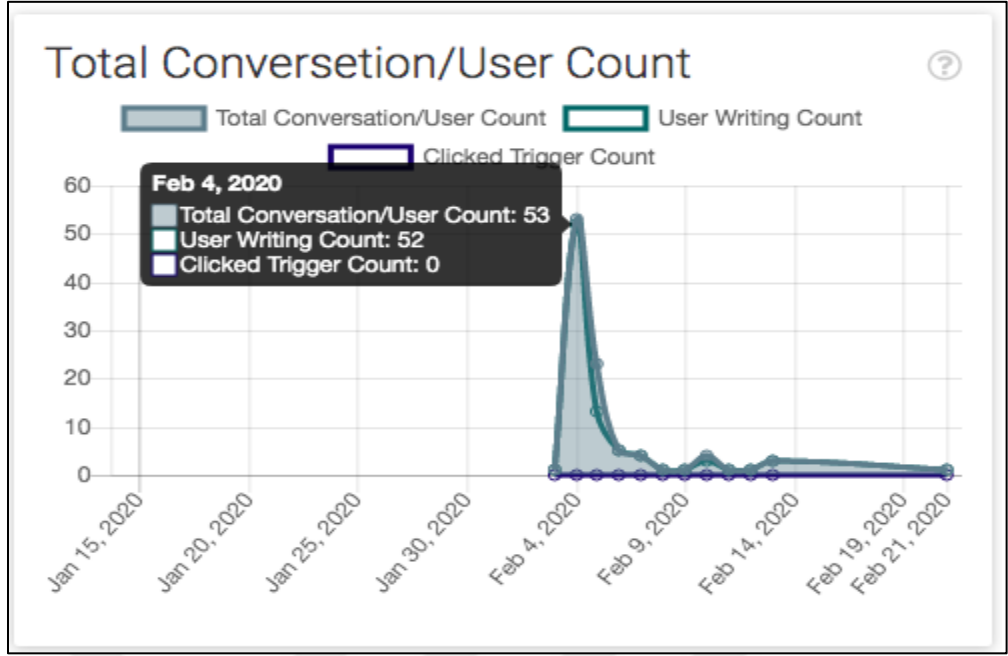
Chatbot başarısını ölçmek için literatürde çeşitli değerlendirme ölçütleri mevcuttur; bunlardan biri de içerik değerlendirmesidir. Otomatik değerlendirme ve uzman değerlendirmesi olmak üzere iki türe ayrılır[35]. Otomatik değerlendirme, makineyi geliştirir ve bir insan tarafından gerçekleştirilen uzman değerlendirmesine kıyasla uygun fiyatlı ve daha hızlıdır. Makinenin bir bot performansını değerlendirmesinin iki yolu vardır: matematiksel yöntemler kullanarak metin özetlemenin kesinliği ve hatırlanmasıdır[35]. Kesinlik, bağlamla eşleşen mevcut konuşmanın bir yüzdesini sağlar ve geri çağırma, sistemin tanımladığı yanıtların yüzde sayısını temsil eder. İnsan, uzman değerlendirmesini hesaplar ve makine değerlendirme yöntemine göre daha doğrudur. Bazı durumlarda, ideal bir çıktıya ulaşmak için iki yöntem entegre edilebilir.

Chatbot platformunda tam olarak analitik sayfasında birkaç metrik bulunur. Şekil 9 ve 10'da gösterildiği gibi, Dahi.ai gibi Chatbot platformları, kullanıcı katılımını (toplam kullanıcılar, ilgili kullanıcılar, yeni kullanıcılar) ölçen analizler sağlar. Analitik, konuşma oranlarına (toplam konuşma, yeni konuşma) dayalı olabilir. Son olarak, belirli bir zaman diliminde botu kullanmak için geri dönen kullanıcıların toplam sayı yüzdesini temsil eden elde tutma ölçümleri gibi, bot ölçümleri botun performansını ölçmenin başka bir yöntemidir. Anketler kullanılarak ölçülebilen kullanıcı memnuniyetinin yanı sıra, kullanıcılar chatbot ile deneyimlerini değerlendirebilirler. Genelde kullanılan chatbot metrikleri verilmiştir.

- Ortalama görüşme süresi
- Toplam görüşme sayısı
- Etkileşimli konuşmaların toplam sayısı
- Toplam benzersiz kullanıcı sayısı
- Cevapsız mesajlar
- İnsan yardımcısına ihtiyaç oranı
- Hedef tamamlama oranı
- Müşteri memnuniyeti puanları
- Ortalama yanıt süresi



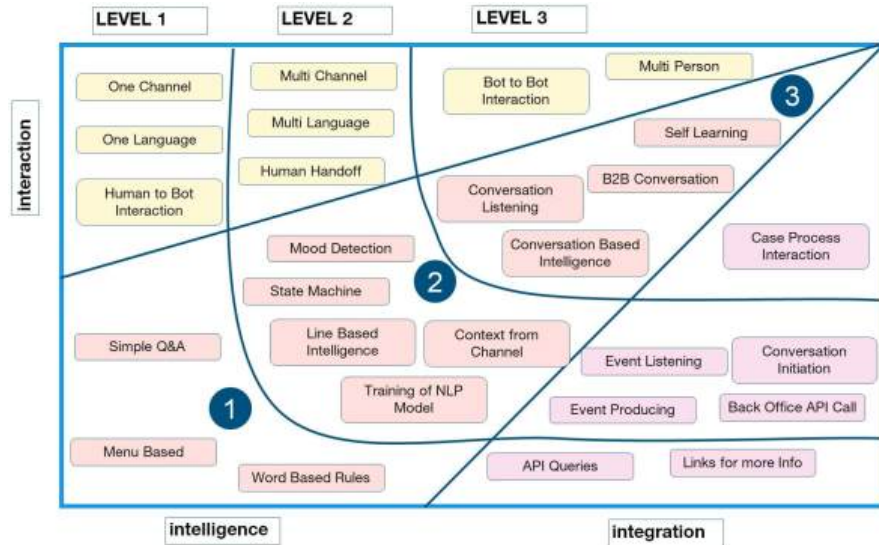
Şekil 3.4: Dahi.ai 's Chatbot Analizi Sayfası.



Şekil 3.5: Dahi.ai Platform Konuşma/Kullanıcı Sayısına Bir Örnek.

3.5. Chatbotların Sınıflandırılması

Chatbot'lar, kullanıcı arayüzünün kolaylığına, algoritmalara ve kullanılan temel teknolojilere göre de sınıflandırılmaktadır. Bu bağlamda chatbot'lar üçe ayrılmaktadır.



Source: Capgemini

Şekil 3.6: Chatbotların Sınıflandırma Şablonu.

3.5.1. Kelime/Menü Tabanlı Chatbotlar:

Bu chatbotlarda, bir yapay zeka teknolojisinden ziyade kullanıcının menü ve butonlardan seçimlerle ya da bazı ifadeleri yazarak ilerlediği görülmektedir. Ön yüzde buton ve menüler mevcut olup, arka tarafta ise daha çok IVR sistemlerinde karşılaşılan karar ağaçları kullanılmaktadır. Sistem, yazılan ifadelerdeki bazı kelimeleri tanımlayarak, bir cevap ile eşleştirmeye çalışmaktadır. Bu chatbotlar, kısıtlı tür ve sayıda soruyu cevaplayabilme yeteneğine sahiptir. Ancak Kelime / Menü tabanlı chatbotlar, yazılan reguler ifadeyi bir bütün gibi anlamlandıramamakta olup, bu konuda yetersiz kalmaktadırlar[36]. İfadelerin yanlış yazımı durumunda kullanıcıyı tam doğru yönlendirebilme yeteneği kazandırmak için olası tüm yanlış yazımların da sisteme tek tek tanıtılması gerekmektedir. Çok basit soru-yanıt chatbotlarında kullanılmalarında sakınca yoktur fakat karmaşık durumlar için ihtiyaçları tam karşılayamamaktadırlar. Bu durumu örnek vermek gerekirse, yazılımda / TL oranının tanımlanmış olduğunu varsayalım. Kelime / Menü tabanlı bir chatbota sorulan “5 TL kaç sterlin eder? Diye sorulduğunda, chatbot “TL” ifadesini tanıyarak 5 TL’nin sterlin karşılığını tam verir, ancak bu ifadenin gerçek anlamını anlayamamaktadır.

3.5.2. İfade Tabanlı Chatbotlar

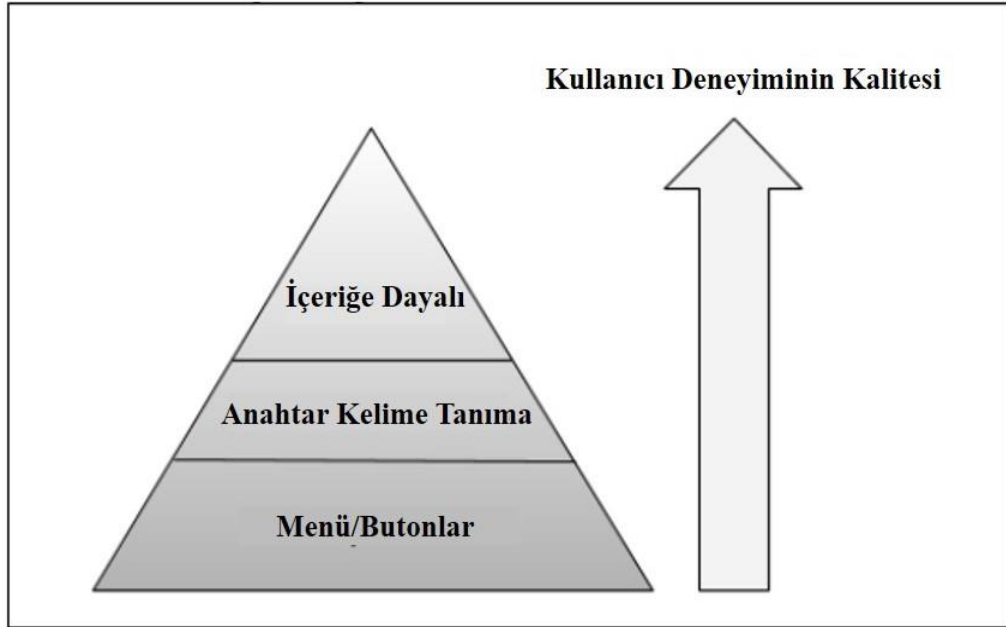
Bu chatbotlar bir Doğal dil işleme (NLP) ve makine öğrenmesi gibi teknolojilerden yararlanmaktadırlar. Bu tür bir chatbot, kullanıcının yazmış olduğu ifadeyi tam bir bütün gibi tanıyıp ve anlamlandırabilmektedir[37]. Ancak bahsedilen diyaloglar devam ettiği sırada, diyalogları en başından beri izleyerek, bağlama uygun yanıtlar verememektedirler. Yine parite örneğinden gidersek, bu türdeki bir chatbot “5 TL kaç sterlin ?” sorusunu bir ifade olarak anlamaktadır, fakat sonrasında sorulan “Peki, kaç pound?” sorusunu anlayamamaktadır.

Çünkü bu durumda bir en başından itibaren izleme ve bağlama uygun şekilde devam ettirebilme yeteneği devreye girmektedir.

3.5.3. Diyalog Tabanlı Chatbotlar

Bu chatbotlar da önceki chatbotlarda bahsedildiği gibi yapay zeka tabanlı chatbotlar ve NLP (Doğal dil işleme) ve makine öğrenmesi / derin öğrenme gibi teknolojilerden yararlanarak geliştirilmektedirler. Diyalog tabanlı chatbotlar önceki açıklamalarda olduğu gibi kelime veya ifade tabanlı bir chatbot anlayışıyla sınırlı kalmayıp, yazılan ifadeleri anlamlandırabilmektedirler. Hatta diyalogun akışını da izlemeye kalıp, buna uygun cevaplar da verebilmektedirler. Hep verilen örneği aynı zamanda diyalog tabanlı chatbot ile deneyimlemek istersek, “5 TL kaç sterlin ?” sorusunu hem ifade olarak anlayarak cevap verip, sonrasında sorulan “Peki kaç pound?” sorusu yöneltildiğinde ise konuşmacının neden bahsettiğini, nasıl bir parite dönüştürme işlemi yapması gerektiğini anlayarak ve tam ve doğru bir cevap vermektedirler.

Chatbot'un etkileşimleri tarafından oluşturulan konuşma verileridir. Chatbotlar bir kullanıcıyla her bağlantı kurduğunda bilgi toplar. Bu veriler, konuşma uzunluğu, kullanıcı memnuniyeti, kullanıcı sayısı, konuşma akışı ve daha fazlasını içerebilir.



Şekil 3.7: Kullanıcı Deneyimi Kalitesi Bakımından Chatbotlar.

3.6. Chatbotların Kullanım Amaçları

Chatbotlar, farklı iş alanlarındaki kullanım amaçları şu şekilde özetlenebilir: Finans, seyahat, turizm, perakende ve ticaret alanlarında şirket ürünleri ve hizmetleri hakkında temel bilgiler sağlamaktadır. Basit işlemlerin, rezervasyonların veya geri ödemelerin otomasyonuna yardımcı olmak amacıyla bildirimler, hatırlatıcılar ve özelleştirilmiş teklifler göndererek kullanıcıların devamlılığını sağlamaya yardımcı olmaktadır. Bu şekilde etkileşimi, marka bilinirliğini ve sosyal medya etkinliklerini artırır.

Pazarlama ve satış chatbot'ları; şirketin web sitesi veya sosyal medya kanallarıyla müşteri katılımını düzenlemekte, ziyaretçilerin proaktifliğini artırmakta ve daha fazla ürün veya hizmet satışı sağlamaktadır. Chatbot'lar ayrıca şirket ürünleri hakkında küçük anketler yapmakta, indirimler veya yeni teklifler önermekte, bülten aboneleri, satış sözleşmeleri veya müşteri geri bildirimleri almakta ve hatta potansiyel müşterilere sunulan ürün ve hizmetlerin özellikleri hakkında bilgi vermektedir. Bazı durumlarda, kurum kültürü hakkında bilgi sağlayarak şirketlerin daha geniş kitlelere ulaşmasına ve yeni takipçiler kazanmasına yardımcı olmaktadır. Tüm bunlar, müşteri deneyimini geliştirerek markayla daha güçlü ilişkiler kurulmasını sağlamaktadır.

Müşteri hizmetleri ve destek, şirketlerin müşterilerinin 7/24 hizmet beklentilerine daha iyi yanıt verebilmek için sorunlarının ve şikayetlerinin anında çözülmesine olanak tanımaktadır. Chatbotlar, kullanıcılarının basit sorunları anında çözmelerine veya sorgularını göndermelerine, sık sorulan bir soruya yanıt almalarına, bir beşeri temsilciden randevu almalarına ve hizmetlerinin düşük maliyetle ölçeklendirilmesine yardımcı olmaktadır. Bu şekilde, botlar kolay ve tekrarlayan sorunlar için ilk basamak asistanları olarak hareket etmekte ve böylece kuruluşun destek ekibinin daha karmaşık sorunlarla ilgilenmesine yardımcı olmaktadır.

Veri madenciliği ve envanter yönetimi, şirketlerin büyük miktarda veriyi işlemesine ve önceden belirlenmiş kategorilere ve kriterlere göre sıralamasına yardımcı olmaktadır. Böylelikle yöntem, uzmanların kullanıcı dostu formatlarda pay çıkarmasına zincir takibi ile trendleri kolayca analiz etmesine ve tahminler yapmasına yardımcı olmaktadır.

Hukuk hizmetleri chatbot'ları, insanlara rehberlik etmekte ve temel yasal belgeleri hazırlamalarına ve tamamlamalarına yardımcı olmaktadır.

Sağlık hizmetleri chatbot'ları, hastalıklar hakkında temel güvenilir bilgiler sağlamaktadır. Hastaların semptomlarını okudukları, sorular sordukları, tıbbi veritabanlarında depolanan verileri analiz ettikleri ve ardından onları doktor tarafından önerilen doğru tedavi aşamasına yönlendirdikleri bir hizmet sunmaktadır. Chatbot'lar ayrıca reçeteleri hatırlatmakta veya insanların ruh sağlıklarını yönetmelerine yardımcı olmaktadır.

İnsan kaynakları chatbot'ları; iş başvurusunda bulunanların ön taraması için kullanılmakta, oryantasyon eğitimi ve işe alım sırasında temel bilgiler sağlamaktadır. Ek olarak çalışanların memnuniyeti ve üretkenliklerine ait bilgileri tutmakta, insan kaynakları ile ilgili sık sorulan soruları yanıtlamakta, seyahat rezervasyonu yapmakta, bordro, yıllık izinler ve hastalık izinleri gibi idari görevlere yardımcı olmakta ve şirket içi eğitimi teşvik etmektedir.

Chatbot'ların bir yardım masası olarak hareket edebildiği ve bilgi teknolojileri (BT) ile ilgili durumları oluşturabildiği/atayabildiği BT hizmetleri, kullanıcılara önemli güncellemeleri bildirmekte veya temel soruları yanıtlamaktadır[38].

Bilgilendirici veya SSS chatbot'lar, kullanıcılara sabit bir kaynaktan edinilebilen önceden belirlenmiş bilgi sağlamak için tasarlanmıştır. Kullanıcının sorgusuna doğru yanıt vermeyi amaçlayan bu botlar ses veya metin tabanlıdır.

Veri tarama botları, günümüzde toplanan büyük miktarda veriyle başa çıkmaya yardımcı olmakta ve kullanıcıların daha ileri analizler için verilerini kolayca sınıflandırmalarına olanak tanımaktadır. Bu tür verilere örnek olarak; kelime sıklığı, kümelenmiş veriler, azalan/artan hiyerarşik sınıflandırmalar vb. verilmektedir. Chatbot'lar farklı kaynaklardan (internet siteleri, sosyal medya, metin, görüntü veya veri bankaları vb.) veri toplamakta ve bunları gözden geçirip karşılaştırarak kolay okunabilir formatlara dönüştürmektedir.

Eğlence chatbot'ları kullanıcıları eğlendirmek veya medya kullanımı amaçları doğrultusunda geliştirilmektedir. Bu tür botlara dair bazı örnekler şu şekilde sıralanabilir:

- Kullanıcıların her kategoriye veya hizmet sağlayıcısına göre kanalları takip etmesini sağlayan TV program rehberi botu,
- Bilgi yarışması sorularına kolay erişim sağlayan ve sonunda doğru cevaplara göre hızlı puan vermeyi sağlayan bilgi yarışması botu,

- Müşterileri farklı pazarlama oyunlarına çeken ve daha fazla bilgi aramak veya bir satın alma işlemi yapmak için onları şirketin internet sitesine yönlendiren bilmece botu,
- Müşterilerin film incelemeleri okumasına, farklı türlere ve kategorilere göz atmasına veya bilet rezervasyonu yapmasına olanak tanıyan sinema botu,
- Kullanıcıların çeşitli haberlere farklı alanlardan ve kaynaklardan uygulamalı olarak erişmelerine yardımcı olan haber ve medya botları,
- Kullanıcıların YouTube, Messenger, WhatsApp vb. uygulamalar ve içerikler arasında gezinmesine yardımcı olan sosyal medya botları,
- Kullanıcıların ilgilendikleri bir konu hakkında ilginç bilgileri kolayca bulmalarına veya bir podcast'e abone olmalarına ve düzenli güncellemeler almalarına yardımcı olan podcast botları,

Eğitim chatbot'ları, öğrencilerin dünya çapındaki eğitim kurumları tarafından sunulan farklı kurslara ve öğrenme içeriğine göz atmalarına imkân sağlamaktadır. Bazı chatbot'lar kurs ücretleri, süreleri, son başvuru tarihleri gibi konularda ayrıntılı bilgi sağlamakta ve böylece öğrencilerin doğru yönlendirilmesine yardımcı olmaktadır. Öte yandan diğer chatbot'lar, etkili öğretim asistanları olarak hareket etmekte ve öğrencilerin ders içeriğini daha iyi anlamalarına, ders sırasında etkileşim ve katılım düzeylerini geliştirmelerine ve verilen ödevleri tamamlamalarına destek olmaktadır. Hatta bazıları değerlendirme sürecine dâhil olmakta veya kendi kendine öğrenme aracı olarak hareket etmektedir[38].

4. WOZ DENEYLERİ

4.1. OZ Sihirbazı (WOZ) Röportaj Sistemi

WOZ sistemi, bir kullanıcının etkileşime girebileceği metin tabanlı bir sohbet arayüzü sunan web tabanlı bir sistemdir. Bir insan operatör (sihirbaz) bir röportaj yapmak için arayüzü kullanır. Sihirbaz, soru sorabilir ve sistem işlevlerini çağırarak da dahil olmak üzere yanıtlar oluşturabilir (örneğin, bir kullanıcının kişiliğini analiz etmek ve sonuçları görüntülemek için bir işlevi çağırarak gibi). Oz Büyücüsü Metodu (Wizard of Oz Method), yapılması planlanan bir teknolojinin veya ürün özelliklerinin bir insan yardımcısı tarafından taklit edilebilmesi kuralını esas alan bir prototipleme yöntemidir.

Genellikle dijital ya da karmaşık bir sistemin prototipi gibi davranır ve kullanıcıyı, gerçekte bir insan yardımcısı tarafından verilen tepkilerin aslında bir sistem tarafından veriliyor gibi düşünmesi için “kandırır”. Wizard of OZ yönteminin temelinde insan-diyalog sistemi etkileşimi sırasında insanın sistem gibi davranmasına, insanın sistemi taklit etmesine dayanmaktadır (Fraser ve Gilbert, 1991). Aynı isimli kitapta yer alan Oz Büyücüsü karakterinden esinlenen teknik, diyalog sistemin tasarım süreci için doğrulama ve veri toplama amacıyla kullanılmaktadır[39].

Oz Büyücüsü tekniği insan-insan etkileşimiyle toplanan gerekli verilere bakılarak veya tasarımcının kendi içgüdüleriyle şekil almış tasarımın işleyişine ilişkin veri toplama olanağı vermektedir.

Bununla birlikte Oz Büyücüsü Yöntemini asıl cazip kılan kullanımının kolay olmasıdır. Çözümleme süreci interaktif dökümleri üzerinden gerçekleştirilmektedirler. Veri toplama esnasında bilgisayarı ya da sistemi birebir taklit eden İnsan Yardımcısı (wizart) olarak adlandırılmaktadır. Oz Büyücüsü yönteminin başarısı teşkil edilen simülasyon ortamına büyük ölçüde bağlıdır. Kullanıcıdan gelen istek, insan yardımcısı aracılığıyla, sisteme eş değer bir biçimde verilmesi çok önemlidir. Örneğin telefon aracılığıyla sesle verilen komutlarla çalışan bir diyalog sisteminin simülasyonunda sesle verilecek olan cevaplar sanki bir makinadan çıkıyormuş görüntüsü yaratmalıdır.

Fraser ve Gilbert (1991) simülasyonun gerçekleşebilmesi için üç ön şartın olduğunu belirtmişlerdir:

- Sistem simülasyonunun sınırları insanın yetenekleri dahilinde kalmalı.
- Simülasyondan önce sistemin nasıl davranacağı detaylı bir şekilde öngörülebilmeli.
- Simülasyon ikna edici olmalı

Her ne kadar insan-makina etkileşiminin farklı şekillerle de gerçekleşebilme olasılığı olsa da diyalog sistemleri üzerinden gerçekleştirilen etkileşimlerin bağlama uygun olması gerekmektedir. Yani konuşarak veya yazışarak sağlanması öngörülmektedir.

Bununla beraber Fraser ve Gilbert (1991) insanın ve diyalog sisteminin ikisinin de doğal dil kullandığı durumlarda oluşabilecek 4 farklı iletişim şekli olduğunu belirtmişlerdir

İki taraf konuşarak iletişim kurar

- İki taraf yazışarak iletişim kurar
- Diyalog sistemi konuşarak, insan yazışarak iletişim kurar
- Diyalog sistemi yazışarak, insan konuşarak iletişim kurar

Google Duplex, insanları bir kişi olduğuna ikna etmeyi başardığında insanların bir sohbet botu ile sohbet ederken bilgilendirilmeleri gerekip gerekmediğine dair etik bir tartışma olmuştur. İnsan botu, insanları kandırmak için "ımm", "hımm" ve diğer tavırları yerleştirdi[40]. İnsanların sohbet botu olduklarını iddia etmek için işe alındığı keşfedildiğinde bir başka tartışma daha ortaya çıktı: "Sohbet botunun arkasına saklanan insanlar", ve "Sözde AI" nın yükselişi: teknoloji firmaları insanları bot yapmak için sessizce nasıl kullanıyor? X.ai ve GoButler.com gibi şirketler, istekleri gerçekten çalışanlar tarafından ele alındığında sohbet botları tarafından hizmet edildiğine inanan kişilere hizmet sundu. Facebook Messenger'ın amiral gemisi chatbot M, insan temsilcilerine işleyemediği istekleri sessizce iletti. Tüm geliştiriciler bunun kapsamını biliyor olsalardı, Messenger'da çalışmak için inşa edilen sohbet botlarından bazıları inşa edilmemiş olabilir. Benzer şekilde, calendar.help yanıtlayanın genellikle bir kişi olduğunu ortaya koymamıştır. İnsanlar çeşitli nedenlerle botlar gibi davranırlar. Yarım asırlık doğal dil anlama araştırması sayesinde araştırmacılar "Oz Büyücüsü" deneylerinde ("perdenin arkasındaki adama dikkat etmeyin") chatbotlar oynadılar. Planlanan sistem özelliklerinin, bunları oluşturma işini yapmadan önce

insan sorgularını işleyip işlemediğini test ederler. Önde gelen beş yapay zeka ve bilişsel bilim araştırmacısı tarafından 1977’de yayınlanan ‘çerçeve güdümlü bir diyalog sistemi’ GUS ’[6], bir Oz Büyücüsü araştırmasında test edilen bir havayolu rezervasyon botunu tanımladı ve bir yazar chatbot gibi davranıyordu[6]. Araştırmacılar, bu dar görev için bile doğal dili kullanmanın çok zor olduğunu keşfettiler

Chatbotun bir başka kullanımı da, yiyeceklerin teslim edilmesi, tiyatro bileti rezervasyonu yaptırması, son dakika hediyeleri satın alması, toplantıları planlaması vb. İçin günün her saati müsait olan bir yapay zeka hizmetçisini işe alacak müşterilerin ilgisini çekmek ve "sihir gibi görünmesini sağlamak." Bu müşteriler, taleplerinin "o kadar akıllara durgunluk veren bir işle uğraşan asgari ücretli gece vardiyası çalışanları tarafından karşılandığını bilselerdi kendilerini iyi hissetmeyebilirlerdi ki, insan çalışanlar botların yerini almayı dört gözle beklediklerini söyledi"[41].

Humbot’lar arayüzü kurtarabilir. “İşi yapmak için bir insanı kullanmak, bir dizi teknik ve iş geliştirme sorununu atlamanızı sağlar. Belli ki ölçeklenmiyor, ancak bir şeyler oluşturmanıza ve zor kısmı erkenden atlamanıza olanak tanıyor.”[41]. Birçok şirket, demoları karar vericileri etkilendikten sonra sohbet botlarının yakında çıkacağını duyurduktan sonra 2016 yılında demo tuzaklarına yakalandı. Humbots, ön uçta olma havasını koruyabilir ve laboratuvar dışında yapılan Oz Büyücüsü deneyleri olarak düşünülebilir ve çalışan bir sohbet botu geliştirmek için kullanılacak veri toplama planları vardır.

Görevlerinin tamamen otomatik olamayacağını bilengeliştiriciler tarafından bir çalışma dalgası var. Sohbet botları, doğru bir yanıtın yüksek bir güveni olan sorguları veya yorumları ele alır ve gerisini insanlara iletir. Botu daha da hassaslaştırmak veya eğitmek için zor vakalar kullanılabilir. Bu, insan ortaklığının uygun olduğu görevlere uygulanabilir[42]. 7.ai platformu kendisini "Temsilcilere yardım eden Sohbet Robotları, Sohbet Robotlarına Yardım Eden Temsilciler" olarak tanıtıyor. Rutin işler bir sohbet robotu tarafından gerçekleştirildiğinde, insan görevleri zihin uyuşturucu olmayabilir.

4.2. Teorik Arka Plan

Sohbet robotlarının yinelemeli, kullanıcı merkezli tasarımı, iyi performans ve teknolojinin amaçlanan dağıtım süreciyle alaka düzeyini sağlamak için gereklidir. Bu bölüm, sohbet robotu geliştirilmenin mevcut durumu ve prototip oluşturmanın buna karşılık gelen rolü ile ilgilidir. Ayrıca, WOZ deneyi prosedürü ve bunun sohbet robotu geliştirme ve araştırma içindeki uygulaması hakkında fikir vermektedir.

4.2.1. Chatbot Prototipleme ve Geliştirme

Sohbet robotları bir tür konuşma arayüzüdür ve İnsan- Bilgisayar Etkileşimi (HCI) araştırması alanına aittir . Kullanıcıları geliştirme sürecine dahil etme ihtiyacı, teknolojik sistemin genel uygunluğunu görmesi ve değer katması için onu uygun şekilde kullanabilmesi gereken insan kullanıcılar olduğu için ortaya çıkmaktadır. Yaygın uygulamaya göre, kullanıcı testleri, önemli bir geliştirme adımı olarak sistem tasarım sürecine entegre edilmektedir.

Genel olarak, bir sohbet robotu geliştirirken girdiye yeterli ve faydalı bir tepki, davranışsal uygunluk ve samimiyet gibi dikkate alınması gereken birçok gereksinim vardır. Ancak, bu tür prototipler doğrudan metin girişine tepki vermez ve bu nedenle sonraki kullanım senaryosundan güçlü bir şekilde soyutlanır. Bu nedenle, teknik olarak geliştirme, gerçek dünya senaryosunda test edilecek işlevsel bir sohbet robotu prototipi oluşturmak için halihazırda bir tür geliştirme platformu, yüksek düzeyde programlama becerileri ve geliştirme deneyimi [8] gerektirmektedir. İçerik açısından, niyet ve yanıt veritabanı esastır ve chatbot'un kalitesini yanıt uygunluğu biçiminde belirler[13]. Bu nedenle, yanıtların eşlik ettiği uygun, kapsamlı bir niyet listesinin oluşturulması, sohbet robotu geliştirmede çok önemlidir. Konuşma arayüzleri, görsel düzen ve etkileşim tasarımından bir ilerleme olarak görülebilir. Metin girişi ile girilen doğal dile dayalı olarak insan kullanıcılarla diyaloga izin veren bir arayüz görevi görürler. Bu nedenle, metin girişi oldukça statik olduğundan ve çok değişken olmadığından, ön uç kullanıcı arayüzü tasarımı için çok az yer bırakırlar. Bu nedenle, sohbet robotu konuşma tasarımında odak noktası içeriğin kendisidir ve aynı zamanda iletişim yoludur.

4.2.2. Teknolojik Yenilikler için Oz Büyücüsü Deneyleri

Ortaya atıldığı şekliyle bir Oz Büyücüsü deneyi, tam olarak işleyen bir teknoloji olarak görünürken, araştırmacıların kullanıcılarla gizli bir şekilde etkileşime girdiği bir simülasyondur. WOz çalışmaları, katılımcıların doğal dil diyaloglarını işleyen bir bilgisayar sistemiyle etkileşime girdiklerine inandırmak için yapılır, oysa gerçekte öyle değildirler. Bu tür bir deneyde sihirbaz olarak adlandırılan bir insan, günümüzün kısıtlamalarını aşmak için konuşmaya aracılık eder. Ancak, yetkin, becerikli bir sihirbazın entegrasyonu nedeniyle, bir WOz senaryosu, incelenen teknolojinin tamamen gerçekçi bir temsilini göstermez ve bütüncül bir test yaklaşımı olarak ele alınamayacak şekilde bir şekilde idealize edilir - daha ziyade ilk fikirleri vermektedir. Özellikle eksik işlevselliğe sahip prototiplenin ilk aşamalarında, WOz deneyleri, mevcut, potansiyel olarak hatalı sistemlere kıyasla gerçekçi, insan benzeri konuşma davranışına ve yetenekli diyalog yönetimine benzedikleri için avantajlıdır. WOz çalışmaları, çeşitli araştırma alanlarına entegre edilmiştir. Bu entegrasyon her türlü teknoloji geliştirme projesine değer katmaktadır. Yapay Zeka (AI) işlevlerini entegre eden sistemler gibi karmaşık teknolojiler bu yaklaşım için özellikle uygundur. Aşağıdaki bölüm, özel sohbet robotu geliştirme alanındaki WOz testini incelemektedir.

4.2.3. Chatbot Geliştirme için Oz Büyücüsü Deneyleri

Oz Sihirbazı kurulumları, gerçek uygulamadan önce test etmek için çeşitli sistemlere ve mimarilere uygulanabilir. WOz deneylerinin, sistemin ilgili tüm kullanıcı gereksinimlerine yakından uyması için erken kullanıcı geri bildirim ve (özellikle teknik) kaynaklardaki tasarruf gibi avantajları, sohbet robotu geliştirmede de kullanılabilir. İnsan kullanıcılarla doğal dilde konuşan sohbet sistemleri olarak, sohbet robotları çoğu zaman AI işlevlerini kapsar ve bu nedenle özellikle geliştirme sürecindeki WOz testleri için uygundur: AI bileşenleri, karmaşık AI çerçeve uygulaması gerekliliği olmadan taklit edilebilir. Chatbot geliştirmede, içerik ve tasarım gereksinimleri açısından dikkate alınması gereken çeşitli yönler vardır.

Bu ön koşulların yanı sıra, konuşma sistemlerinin oluşturulmasıyla ilgili bazı kısıtlamalar vardır: Sohbet robotları önceden tanımlanmış veri tabanlarına ve

dolayısıyla önceden belirlenmiş girdilere bağlıdır, bu da WOz tabanlı prototip testlerini beklenmedik ve dolayısıyla dikkate alınmayan içeriği kapsamakla alakalı hale getirir. Bu, ön konuşma tasarımının ilk kullanıcı algılarını değerlendirmek için ideal bir kurulumdur ve aynı zamanda WOz çalışmalarının veri toplamak için kullanılabilmesini belirten ve uyumlu olan yeni içerik derlemesine izin vermektedir. Bu çalışmada, WOz deneyi, bir sonraki adım olarak sohbet robotu prototipinde uygulanmak üzere eldeki sohbet robotu prototipinin test verilerinin yanı sıra ilgili amaçları ve beraberindekilerle bir sonraki aşama için eğitim verisi oluşturmaktadır. Arayüzün kendisi de sohbet robotları için en yaygın erişim noktası olarak belirli bir sosyal medya kanalı veya mesajlaşma uygulaması biçiminde önceden belirlenir. Bu nedenle, WOz çerçevesinin bu ortama entegre edilebilir olması ve beklenen tam gelişmiş sohbet robotundan ayırt edilemeyecek şekilde uyması gerekir. WOz yaklaşımı, sohbet robotu araştırmalarına yaygın olarak uygulanmıştır.

4.3. Yöntemsel Yaklaşım

Eldeki çalışma, gelişmiş bir sohbet robotunun simülasyonu için WOz testine odaklanıyor. Chatbot, elektronik (yani web tabanlı) bir tercih süreci içinde farklı konular ve süreç adımları hakkında SSS'leri yanıtlama kullanım durumuna uygulanır. Bu bölümde, amaçlarını, chatbot kavramını ve WOz çerçevesini içeren çalışmanın metodolojisi sunulmaktadır.

4.3.1. Oz Büyücüsü Çalışmasının Hedefleri

Bu çalışmanın üç kapsayıcı hedefi bulunmaktadır.

-Amaç Eşleştirme ve Yanıt Uygunluğu Değerlendirmesi

Sohbet robotu kavramları bir WOz test ortamında simüle edilebilir. Gerçek bir kullanıcı geri bildirim almak için sihirbazın tepkileri, amaçlanan sohbet robotunun yalnızca işlevlerini değil aynı zamanda sınırlamalarını kapsayıcı da yansıtmalıdır. Sihirbaz bu nedenle özgürce yanıt vermemeli, ancak önceden tanımlanmış kuralları ve ayarları izlemesi gerekmektedir. Deney yoluyla, sihirbazın (insan yardımcısının) hangi kullanıcı sorguları ve kullanıcı isteğini yanıtlamak için mevcut bir kullanıcı amacı ile eşleşebileceğini, hangi durumlarda sihirbazın yanıtları değiştirmesi gerektiğini veya

önceden tanımlanmış hiçbir amaç bulunmadığını ve dolayısıyla bir yanıt bulunmadığını değerlendirmeye çalışılmıştır. İnsan yardımcısının uzmanlığına dayalı olarak formüle edilmesi gerekmektedir. Sonuç olarak, başlangıçtaki niyet setinin eksiksizliği ve uygunluğu değerlendirilmelidir.

-Konuşma Tasarımı Değerlendirmesi

Ek olarak, tercih SSS sohbet robotunun ilk versiyonunu kurduktan sonra, tercih ve diğer SSS'in özel uygulama alanındaki sohbet botu deneyimi kullanıcının memnuniyeti ile değerlendirilir ve algılanan kullanılabilirlik ve ayrıca gösterilen çözümün performansı, nicel (kullanıcı anketi) yaklaşımının geri bildirimleri toplanarak incelenmektedir.

-Niyet oluşturma

Tercih SSS sohbet robotu konseptinde hali hazırda uygulanan konuların test edilmesinin yanı sıra, daha fazla bilgi ihtiyacının ve ilgili kullanıcı niyetlerinin tanımlanması ve niyet setine entegre edilmesi gerekmektedir. Bu set, sohbet robotunun en yaygın sorular için ilgili yanıtları sağlaması amacıyla bir noktaya kadar genişletilmelidir. Amaç oluşturma dışında, potansiyel eğitim verileri, kullanıcı girdi ifadelerinin amaçlara entegre edilerek toplanması ve atanmasıyla WOz testinden yeni veriler türetilmektedir. Ayrıca, WOz deneylerinin bu şekilde kullanımı, chatbot çözümünün uygulandığı ve eğitilmesi gereken sonraki aşamalarda daha önemli ve üretken hale gelebilir. Gelişmiş sohbet robotlarının doğal dil anlayışı ve niyet eşleştirme bileşenleri (önceden eğitilmiş) makine öğrenimi algoritmalarına dayandığından ve dolayısıyla gelişmek ve iyileştirmek için alana özgü eğitim verilerine dayandığından eğitim mutlaka gereklidir.

WOz yaklaşımı, tercih ve genel kullanım SSS sohbet robotu prototipini yukarıda sunulduğu gibi ilgili perspektiflerden test etmek ve doğrulamak için kullanılır. Bulgulara dayanarak, sohbet robotu yinelemeli olarak uyarlanacak, iyileştirilecek ve daha da geliştirilecektir. WOz yaklaşımı, tercih ve genel kullanım SSS sohbet robotu prototipini yukarıda sunulduğu gibi ilgili perspektiflerden test etmek ve doğrulamak için kullanılır. Bulgulara dayanarak, sohbet robotu yinelemeli olarak uyarlanacak, iyileştirilecek ve daha da geliştirilecektir.

4.4. Chatbot Oluřturma ve Yapılandırma

Bu ön çalışma, katma deęerli sohbet robotu kullanım durumlarının belirlenmesine ve başvuran izleme sistemlerinde sohbet botu işlevlerinin uygulanmasına odaklanan araştırma projesinin bir parçasıdır. Tercih yapacak adayları desteklemek ve tercih sürecindeki soruları yanıtlamak için bir SSS sohbet robotunun özel kullanım durumunun genel uygunluęu, daha önceki arařtırmaların konusudur.

Hedef grubun ihtiyaçlarını karřılamak için farklı kaynaklardan niyetler toplanılmıřtır:

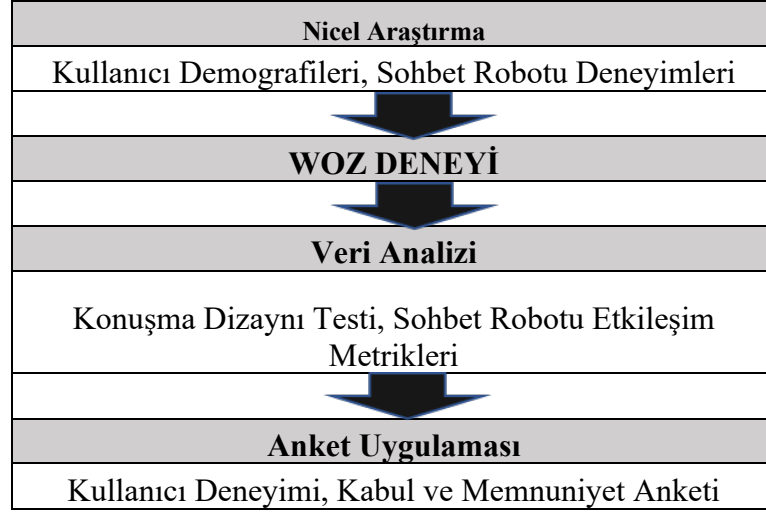
Tercih yapmak üzere olan potansiyel adaylardan bir başvuru esnasında ve sonrasında karřılařabilecekleri durumları ve problemleri hakkında sorular formüle edilmeye çalışılmıřtır.

Üniversitelerin Bilgisayar Mühendislięi bölüm siteleri ve Bilgisayar Mühendislięi sosyal medya tercih kanallarında yer alan mevcut SSS'lerdeki sorular ve cevaplar taranmıř ve konsolide edilmiřtir . Ayrıca dięer kanallardan gelen bilgi sorguları uygulanmıřtır.

Toplamda, SSS tercih sohbet robotu konseptine dahil edilecek 112 niyet-amaç seti belirlenmiřtir. Bu veriler deney boyunca sihirbazın veri tabanı olarak kullanılmıřtır.

4.4.1. Çalışma Tasarımı

Eldeki çalışma, tanımlanan hedeflerle uyumlu olacak şekilde tasarlanmıřtır: Yanıt uygunluęu deęerlendirmesi, konuşma tasarımı deęerlendirmesi ve niyet oluřturma şeklinde amaç eřleřtirme konularını içermektedir. Deney çalışma tasarımı, Őekil 13'te gösterildięi gibi dört alt ardışık bölümden oluşur ve ařaęıdaki paragraflarda açıklanmaktadır.



Şekil 4.1: WOZ Çalışma Yaklaşımına İlişkin Akış Şeması.

-Niceliksel Anket I

Chatbot deneyinden önce, katılımcıların bazı sosyo-demografik özellikleri (örneğin, mezuniyet vs.), sohbet robotlarıyla ilgili deneyimlerinin hakkında anket yapılmıştır.

-Veri Analizi

Bu aşamada WOz deneyi başlatılmış ve katılımcılardan (öğrencilerden) tercih edecekleri bölümle ilgili olarak karşılaşılabilecekleri soruları yöneltmeleri istenmiştir.

-Nicel Anket II

Katılımcılardan WOz deneyini tamamladıktan ve chatbot ile başarılı bir şekilde sohbetten sonra, chatbot desteğinden ve ilgili kullanıcı deneyiminden memnuniyetlerine odaklanan nicel bir anketi yanıtlamaları istenmiştir.

Bu süreçte (yerinde veya uzaktan) katılımcılara bir moderatör eşlik ederken, araştırmacılardan biri WOz çerçevesinde sihirbaz rolündeydi. Ayrıntıları, WOz deney kurulumunun aşağıdaki açıklamasında detaylandırılacaktır.

4.4.2. Oz Büyücüsü Deneyinin Kurulumu ve Teknik Altyapı

Bir otomasyon aracısıyla WOz testi üzerine bir çalışma yürüten sihirbaz için, katı bir davranış planının çok önemli olmakla birlikte, bu tutarlı davranış, dolayısıyla deneysel güvenilirliği korumak için çok önemlidir. Sunulan deneyin dört bölümünde, birkaç bileşen kavramsallaştırılmıştır.

Genel olarak, aşağıdaki roller atanmıştır:

-Katılımcı

Aday öğrenci süreçlerinde chatbot ile sohbet eden, potansiyel adaylardan oluşan hedef gruba ait tercih yapacak chatbot kullanıcıları.

-Sihirbaz

Deneysel çalışma çerçevesine dayalı olarak ayrı bir odadan/uzaktan uzaktan, önceden formüle edilmiş mesajlar göndererek veya görünüşte yapay zeka tabanlı otomatik yanıtlar olarak geçici yanıtlar oluşturarak WOz çerçevesini çalıştıran araştırma projesine ait bir araştırmacıdır.

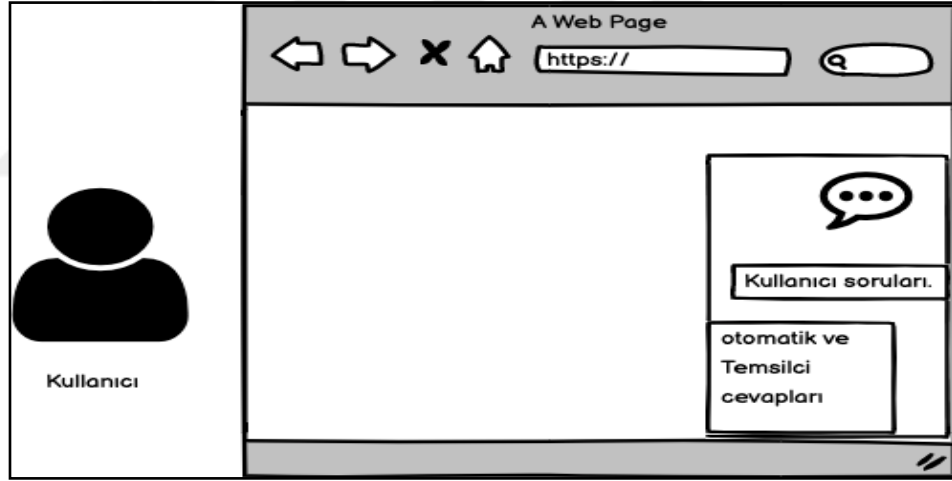
Katılımcılar, sohbet robotuna ve onunla etkileşimlerine yönelik yaklaşan tüm düşüncelerini, algılarını ve duygularını ifade etmeye teşvik edildi. WOz deneyinden sonraki nicel anket, UX ögesi (tasarıma odaklanmayan arayüz aracılığıyla etkileşime ilişkin), katılımcıların yanıt kalitesine (cevapların yeterliliği, yetkinlik ve hız) ilişkin memnuniyetlerine ilişkin sorular, algılanan katma değere ilişkin sorular içermektedir. Genel olarak, aday öğrenci sürecinin her bir uygulama adımı için chatbot desteği verilmiştir.

WOz testi için gerekli olan tek bileşen, arayüz yazılımı ve veri tabanlarıdır. Buna uygun olarak, WOz çerçevesi teknik olarak, sihirbaz ve deney katılımcısı için arka uç ve ön uç sohbet arayüzlerine izin veren ücretsiz bir açık kaynak sohbet platformu olan Rocket.Chat aracılığıyla kurulmuştur.

Ayrıca, bu sohbet sunucusu sistemi, verilerin saklanması, yani mesajların daha sonra analiz edilmek üzere ek bilgilerle (örneğin, zaman damgaları) günlüğe kaydedilmesi için işlevler sağlamıştır. Sohbet sunucusu olarak Rocket.Chat, gelişmiş programlama uzmanlığına ihtiyaç duymadan basit kullanımı ve konfigürasyonu nedeniyle seçilmiştir. Platformun Canlı Sohbet özelliği, katılımcılar için bir iletişim kanalı olarak kullanılmıştır. Sohbet penceresi bir HTML sayfasına entegre edilmiştir ve daha sonra katılımcılar tarafından bir web tarayıcısı aracılığıyla erişilebilir hale getirilmiştir.

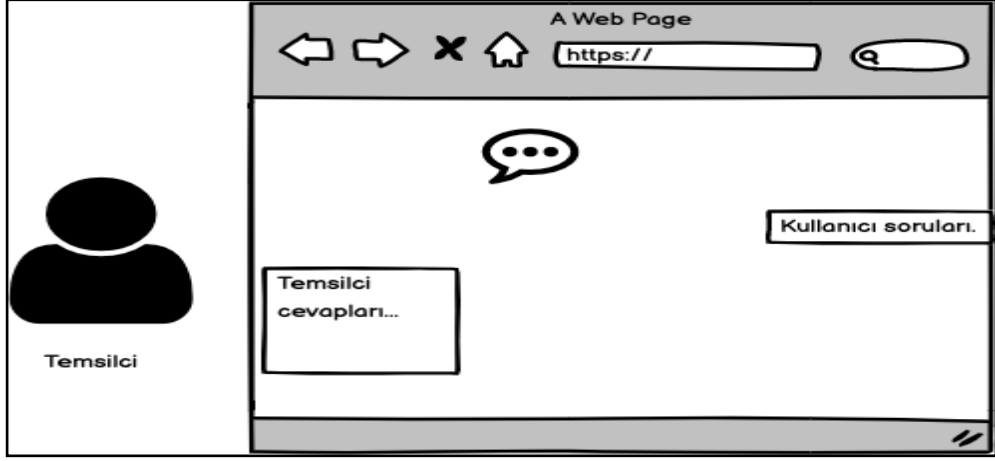
Rocket.Chat, yüksek standartlarda veri korumasına sahip kuruluşlar için özelleştirilebilir bir açık kaynaklı iletişim platformudur. İş arkadaşları arasında, diğer şirketlerle veya müşterilerinizle web, masaüstü veya mobil cihazlarda gerçek zamanlı konuşmalar sağlar.

Katılımcının ön uç konfigürasyonu Şekil 13'te gösterilmektedir. Sol tarafta gösterildiği gibi kullanıcı çalışırken, insan temsilcisi chatbot'a tüm süreç boyunca sağ alt köşede erişilebilir durumdadır. Bu şekilde, soru veya rahatsızlık(cevap yetersizliği, hız vs.)şeklinde ortaya çıkan tüm sorunlar, katılımcılardan chatbot'a yönlendirilebilir. Katılımcılar tarafından sohbet penceresinde gönderilen mesajlar Rocket.Chat'e gönderilmiş ve burada saklanmıştır. Sohbetlere belirli bir arayüz üzerinden erişilebilir ve kaydedilebilir. JavaScript aracılığıyla, JSON verileri daha sonra Microsoft Excel aracılığıyla değerlendirilmiştir. Sihirbaz olarak görev yapan Temsilci, bir chatbot olarak görünürken arka planda gelen soruları almak ve işlemek için Rocket.Chat yönetim arayüzünü kullanılmıştır.

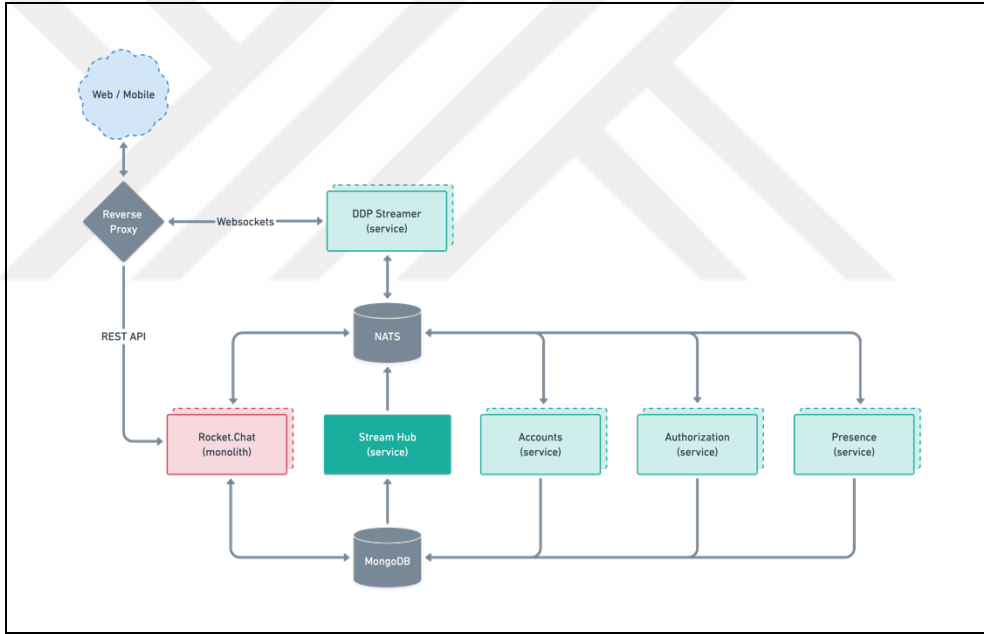


Şekil 4.2 : Oz Sihirbazı Kullanıcı Ön Uç Yapılandırması.

Şekil 14'te WOZ Deneyinde Sihirbaz (Temsilci) Ön Uç Yapılandırması gösterilmiştir. Şekil 14'te temsilci kullanıcının sorularını daha öncesinde oluşturulan Niyet veri setine göre cevaplamaktadır. Ayrıca kullanılan microservisin sistem mimarisi aşağıda Şekil 15'te verilmiştir

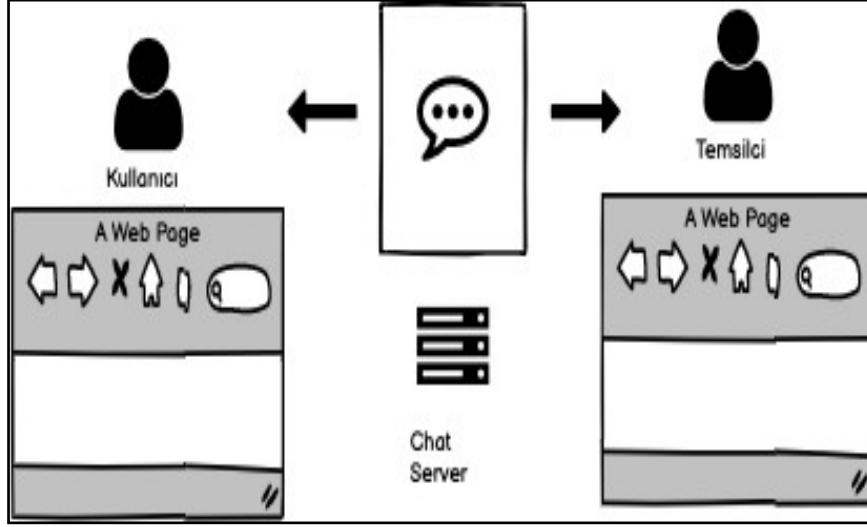


Şekil 4.3: Oz Sihirbazı Sihirbazı Ön Uç Yapılandırması.



Şekil 4.4: Şekil Rocket.Chat Microservis Sistem Mimarisi.

Şekil 17, çerçeveyi, farklı rolleri, belgeler ve süreçler dahil olmak üzere genel çalışma tasarımına yerleştirilmiş prosedür olarak göstermektedir. Bütünsel olarak Rocket.Chat sohbet ortamında sunucuyla, kullanıcılar çerçeveye ön uç perspektiften (sol taraf) erişirken, sihirbaz beklenen SSS'yi taklit ederek arka uç perspektiften gizlice çalışmıştır.



Şekil 4.5: Çalışmanın Oz Büyücüsü Deneyi Çerçevesi.

Ayrıca Js ile arayüz GTÜ Bilgisayar Mühendisliği web sayfasında pop-up ile eklenmiş olup sohbetlerden küçük bir kesitin ekran görüntüsü aşağıdadır.



Şekil 4.6: Chatbot Arayüzüne ilişkin bir kesit.

5. BULGULAR VE SONUÇLAR

5.1. Elde Edilen Bulgular

Otomatik Yanıt Sistemi ve bununla birlikte İnsan Temsilcinin eklendiği toplam 62 sohbet botu oturumunda sihirbazla etkileşime girdi. Oturum başına sohbet robotu etkileşimlerinin oranı (c sütunu) 0,75 ile 1,00 arasında değişmekte olup, ortalama değer 0.9 'dur. Bu değerlendirme, her şeyden önce, kullanım için aktivasyonun ve kullanım yoğunluğunun çalışmadaki katılımcılar arasında büyük ölçüde değiştiğini göstermektedir.

Ayrıca, bazı katılımcıların (bir web sitesindeki bir arama isteğine kıyasla) hızlı bir yanıt almayı beklediği, diğerlerinin ise amaçlanan bilgileri almak için sohbet robotu ile etkileşimli bir diyaloga daha fazla dahil olduğu açıktır.(d) ile (e) sütunlarında da sadece Sohbet Robotu ve İnsan Temsilcinin yanıt süreleri gösterilmiştir. (f) sütununda ise sadece Sohbet robotunun bitirdiği oturumlar listelenmiştir.

Maalesef sadece 4 katılımcı oturumu tek başına (sadece sohbet robotuyla) bitirebilmiştir. Toplam 914 dakikalık deneyde toplamda 795 dakika etkileşim olup bunun da 436 dakikasında Temsilci cevap vermiştir. Bu da ortalama %55'ini İnsan Temsilci'nin yaptığını göstermektedir.

Tablo 5.1: WOZ Deneyi Metrikleri.

(#)	(a) Oturum Süresi (dk)	(b) Etkileşim Süresi (dk)	(c) Oturum Başına Etkileşim süresi (dk)	(d) SR cevap Süre (dk)	(e) Temsilci cevap Süresi (dk)	(f) Sadece Sohbet Robotuyla bitirilip bitirilemediği
1	16	14	0,88	6	8	
2	14	13	0,93	7	6	
3	17	15	0,88	9	6	
4	12	11	0,92	8	3	
5	14	14	1,00	8	6	
6	13	12	0,92	5	7	
7	9	9	1,00	8	1	

Tablo 5.1: Devam.

(#)	(a)	(b)	(c)	(d)	(e)	(f)
	Oturum Süresi (dk)	Etkileşim Süresi (dk)	Oturum Başına Etkileşim süresi (dk)	SR cevap Süre (dk)	Temsilci cevap Süresi (dk)	Sadece Sohbet Robotuyla bitirilip bitirilemediği
8	14	11	0,79	8	3	
9	14	12	0,86	8	4	
10	12	11	0,92	9	2	
11	15	12	0,80	9	3	
12	17	15	0,88	10	5	
13	16	14	0,88	9	5	
14	15	11	0,73	9	2	
15	15	14	0,93	11	3	
16	10	10	1,00	9	1	
17	13	12	0,92	10	2	
18	14	11	0,79	9	2	
19	15	14	0,93	11	3	
20	16	15	0,94	11	4	
21	17	14	0,82	8	6	
22	12	11	0,92	7	4	
23	11	11	1,00	9	2	
24	14	12	0,86	9	3	
25	16	14	0,88	8	6	
26	12	11	0,92	9	2	
27	15	14	0,93	8	6	
28	13	12	0,92	9	3	
29	11	11	1,00	8	3	
30	16	12	0,75	10	2	
31	14	13	0,93	9	4	
32	18	16	0,89	9	7	
33	14	13	0,93	11	2	
34	17	13	0,76	12	1	

Tablo 5.1: Devam.

(#)	(a)	(b)	(c)	(d)	(e)	(f)
	Oturum Süresi (dk)	Etkileşim Süresi (dk)	Oturum Başına Etkileşim süresi (dk)	SR cevap Süre (dk)	Temsilci cevap Süresi (dk)	Sadece Sohbet Robotuyla bitirilip bitirilemediği
35	15	14	0,93	10	4	
36	12	10	0,83	9	1	
37	13	12	0,92	10	2	
38	18	13	0,72	11	2	
39	14	12	0,86	10	2	
40	13	12	0,92	11	1	
41	15	14	0,93	12	2	
42	16	12	0,75	10	2	
43	12	12	1,00	9	3	
44	10	10	1,00	9	1	
45	17	13	0,76	12	1	
46	14	12	0,86	12	0	*
47	17	16	0,94	13	3	
48	16	11	0,69	10	1	
49	15	15	1,00	14	1	
50	16	14	0,88	13	1	
51	14	13	0,93	12	1	
52	15	13	0,87	11	2	
53	17	12	0,71	10	2	
54	16	14	0,88	13	1	
55	15	14	0,93	12	2	
56	12	10	0,83	10	0	*
57	18	13	0,72	12	1	
58	14	13	0,93	12	1	
59	17	12	0,71	12	0	*
60	19	15	0,79	13	2	
61	18	15	0,83	14	1	

Tablo 5.1: Devam.

	Oturum Süresi (dk)	Etkileşim Süresi (dk)	Oturum Başına Etkileşim süresi (dk)	SR cevap Süre (dk)	Temsilci cevap Süresi (dk)	Sadece Sohbet Robotuyla bitirilip bitirilemediği
62	14	12	0,86	12	0	*
Toplam	904	785		359	426	
Ortalama	28,25	24,53	0,89	8,59	15,94	

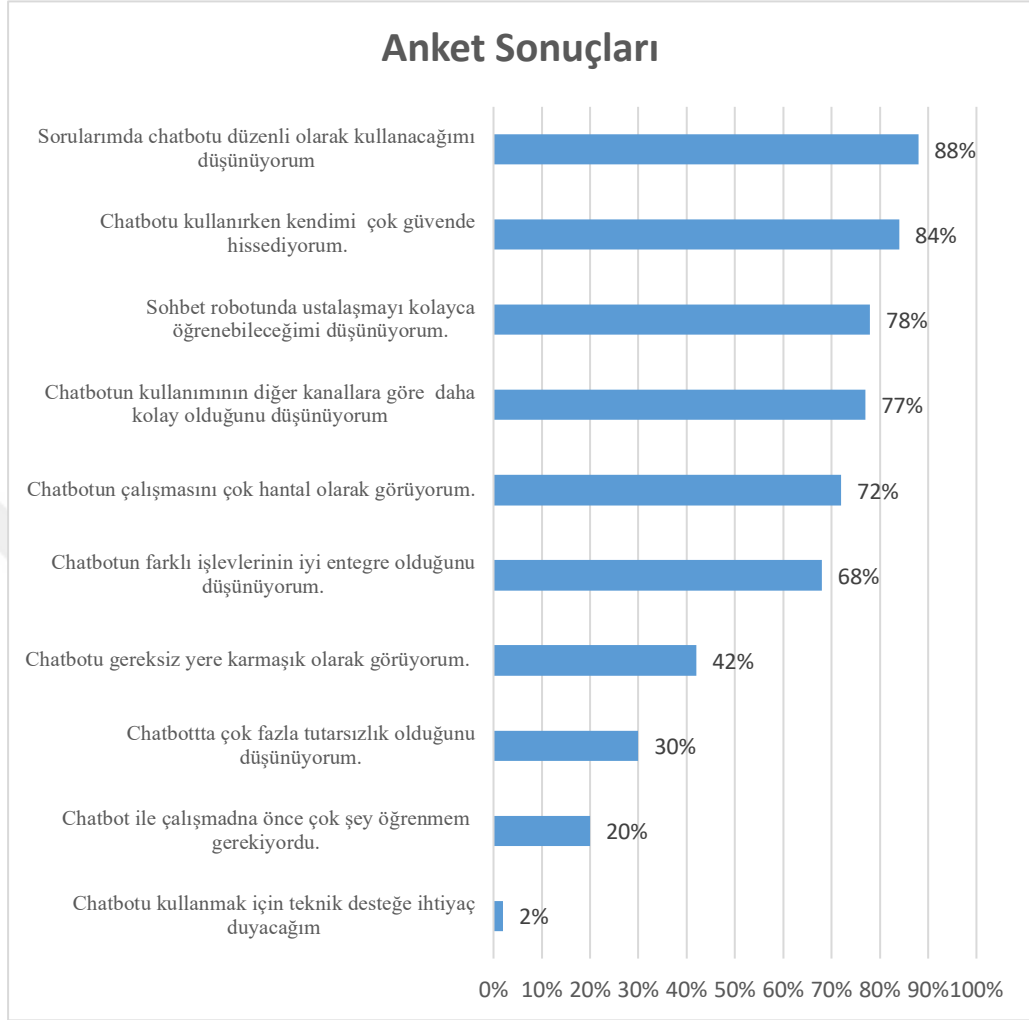
Tablo 2, bu anketin sonucunun bir özetini göstermektedir. Tablo, simüle edilmiş sohbet robotuyla etkileşime giren 62 katılımcı için anket sonuçlarını göstermektedir. Cevap eksiksizliğine ilişkin memnuniyetin oldukça memnun ve orta düzeyde memnun olduğu ve sonraki kullanım ölçüsü olarak, sihirbazın genel olarak, yanıt sürelerinin otomatik bir yanıtlama sisteminden beklenenden daha önemli olmuştur.

Tablo 5.2: Chatbot Prototipinin Kullanıcı Deneyimi Değerlendirmesi (%)

(#)	(a)	(b)	(c)	(d)	(e)
	Tamamen Memnun	Oldukça Memnun	Orta derecede memnun	Memnun değil	Hiç Memnun değil
Cevap Yeterliliği	11,3	22,6	24,2	22,6	19,4
Hız	24,2	25,8	19,4	17,7	12,9
Performans	19,4	17,7	30,6	19,4	12,9
	Çok Yüksek	Yüksek	Ortalama	Oldukça Düşük	Çok Düşük
Genel Katma Değer	22,6	27,4	25,8	12,9	11,3

Tablo 2’de ise en belirgin sorun, sohbet robotunu simüle eden insan sihirbazının neden olduğu uzun gecikme süreleriydi. Bununla birlikte, katılımcılardan Bu sohbet robotunun sordukları sorulara verilen cevaplar çerçevesinde genel bir katma değeri sorulduğunda %22,6’sı Çok Yüksek olduğunu ve %27,4’ü ise Yüksek olduğunu belirtmişlerdir. Bu da yaklaşık olarak %50’sinin bu konsepti Yüksek ve Çok Yüksek genel katma değeri olduğunu beyan etmişlerdir.

Tablo 5.3: SSS Tercih, Chatbot Prototipinin Kullanıcı Deneyimi Memnuniyeti Derecelendirmesi



Ayrıca, sihirbaz kısmını optimize ederek, cevapları seçmek için bir tavsiye sistemi entegre ederek (yanıtları webde aramak yerine) veya cevaplar için metin yazmaktan kaçınmak için dilden metne ara yüzün entegrasyonu ile yapılabilir.

Daha kapsamlı bir gözlem de soruların karmaşıklığıyla ilgilidir. Chatbot'un uygulanmasına yönelik çözüm, tek bir kullanıcı isteminde birden çok amacın tanımlanmasını desteklemez. Gerçekçi bir senaryo için, sohbet robotu soruları birer birer yanıtladı. Daha sonraki bir uygulamada, bu tür sorunları tanımlamak ve kullanıcılara soruları basitleştirmek için uygun geri bildirim sağlamak için çözümler bulunmalıdır.

Sıklıkla dile getirilen bir başka yorum da genel olarak kullanıcıları tatmin etmeyen, aksine onları hüsrana uğratan ve bazı durumlarda beceriksiz olarak algılanan

birkaç chatbot cevabının algılanan yüzeyselliği. Bu, deneyde kullanılan amaç kümesinin yanı sıra karşılık gelen yanıtların iyileştirilmesi gerektiğini gösterir.

5.2. Sonuç ve Tartışma

Bu çalışma, WOz deneylerinin, sohbet robotları için kavramları doğrulamak için sistem geliştirme'nin ilk aşamalarında kullanılabileceğini göstermiştir. Uygun altyapılar, mevcut açık kaynaklı web ve sohbet sunucusu çözümlerine dayalı olarak yönetilebilir miktarda kaynakla uygulanacaktır. Bu tür WOz kurulumlarında, katılımcılar gerçek bir sohbet robotuyla etkileşime girmeye ikna edilebilir. Bununla birlikte, seçilen sohbet robotu tasarımının kısıtlamaları deneyde korunacaksa uygun yanıtları seçmek için gereken süre sorunludur ve sihirbaz yalnızca elle yanıt vermemelidir. Deney ayrıca, kullanıcıların bir sohbet robotu tarafından sunulan desteği otomatik olarak kabul etmediklerini ve böyle bir sistemle daha kapsamlı bir diyaloga girmeleri gerekmediğini de göstermiştir.

Çalışmanın bulguları, başvuru süreçlerinde SSS sohbet robotlarının uygulanmasının katılımcılar tarafından ustalaşması kolay ve değerli görüldüğünü göstermektedir. Ancak, bir kullanıcıyla etkileşime girerken, sohbet robotu yalnızca sorulara uygun cevaplar sağlamakla kalmamalı, aynı zamanda karmaşık sorgulamalar durumunda gerekli basitleştirmeleri belirtmeli ve hatta insan uzmanlar tarafından ikinci seviye bir destek sunmak için inisiyatif almalıdır.

WOz deneyleri, sohbet robotu konseptinin gerekli içeriği ve kapsamı hakkında da önemli bilgiler sağlayabilir. Kullanıcı sorularının çoğu, sohbet robotu kavramının mevcut durumunu yansıtan belirli bir amaç kümesinden önceden tanımlanmış yanıtlarla ele alınabilirken, deneydeki kullanıcı sorgularının % 30'undan fazlasının, süreyi genişletme gereğini gösteren sihirbaz tarafından serbest olarak yanıtlanması gerekiyordu. Bu, tam olarak tatmin edici olmayan ve bu nedenle iyileştirilmesi gereken yanıt bütünlüğüne ilişkin nicel anketin bulgularıyla desteklenmektedir. Sohbet robotu yanıtlarının algılanan yüzeyselliği, deneyde tanımlanan sohbet robotu kavramının, niyet/yanıt kümelerinde daha fazla iyileştirme olduğunu gösteren kaliteyle ilgili başka bir sorundur.

Ayrıca bazı sınırlamaların da dikkate alınması gerekir: Bulgularımız, yalnızca 62 katılımcıdan oluşan çok küçük bir örnekleme yapılan bir WOz çalışmasına

dayanmaktadır. Bununla birlikte, geliřtirmenin erken ařamalarında ve genelleřtirilebilir sonuçlardan daha çok genel fizibilite ve kullanılabilirlięe odaklanan alıřmalarda, kk test kiři grupları olduka yaygındır[43]. Chatbot konseptinin uygulanmasına iliřkin bir dięer yanlıřlık ise sihirbazın niyet eřleřtirme performansının piyasada mevcut olan gnmz chatbot platformları tarafından saęlanıp saęlanamayacaęıdır. Kullanıcı sorgularının kapsamı ve sohbet robotunun yanıtları, niyet seti tarafından gereki bir řekilde sınırlandırılmıř olsa da, daha sonraki bir uygulamada niyet eřleřtirmesi hala nemli lde deęiřebilir ve bu da kullanıcı memnuniyetini de etkileyebilir. Genel olarak, WOz deneyleri iin, tutarlı sihirbaz davranıřını srdrmek ve hataları veya en dřk sistem performansını simle etme yetersizlięi bu tr alıřmaların sınırlayıcı ynleridir[14].

Yazılımla telepatik olarak etkileřim kurabilene kadar, konuřma arayzleri, yazılımın insanlara zg olması iin sahip olduęumuz en iyi yol olacaktır .

alıřmamızda Wizard Of Oz metodunun hem tasarlanırken hem de geliřtirilirken kullanılabilereęini, bu řekilde bir yntemle daha hızlı, etkin ve kullanıcı dostu chatbotlar tasarlanabileceęi sonucuna elde edilen kullanıcı memnuniyet anketleri sonularıyla varılmıřtır.

KAYNAKLAR

- [1] Turing A. M., (1950), “Can a Machine Think. Mind”, 59(236), 433-460.
- [2] Abdul-Kader S.A., Woods D. J., (2015), “Survey on Chatbot design techniques in speech conversation systems”, International Journal of Advanced Computer Science and Applications, 6 (7), 72–80.
- [3] Weizenbaum J., (1966), “ELIZA—a computer program for the study of natural language communication between man and machine”, Communications of the ACM, 9(1), 36-45.
- [4] Weizenbaum J., (1967). “Contextual understanding by computers” Communications of the ACM, 10(8), 474-480.
- [5] Schildknecht L., Eißer J., Böhm, S., (2018), “Motivators and barriers of chatbot usage in recruiting: An empirical study on the job candidates’ perspective in Germany” , Journal of E-Technology, 9(4),109–123.
- [6] Gnewuch U., Feine J., Morana S., Maedche A, (2020), “Soziotechnische gestaltung von chatbots (socio-technical design of chatbots)”, In Cognitive Computing, Springer Fachmedien Wiesbaden, 12(2) 169–189.
- [7] Jurafsky D., Martin J., H. (2018), “Dialog systems and chatbots”, 2nd ed. Upper Saddle River, New Jersey: Prentice Hall.
- [8] McTear M. F., (2017), “The rise of the conversational interface: A new kid on the block?” In Future and Emerging Trends in Language Technology. Machine Learning and Big Data: Second International Workshop, 38–49, Seville, Spain, November 30–December 2.
- [9] Nielsen J., (1992), “The usability engineering life cycle”, Computer, 25(3), 12–22.
- [10] Böhm S., Eißer J., Meurer S., Linnyk O., Kohl J., Locke, H., Teetz I., (2020), “Intent Identification and Analysis for User-centered Chatbot Design: A Case Study on the Example of Recruiting Chatbots in Germany”, In Proceedings of the Thirteenth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC), 34-43, Porto, Portugal, 18-22 October.
- [11] Schlögl S., Doherty G., Luz S., (2014), “Wizard of oz experimentation for language technology applications: Challenges and tools”, Interacting with Computers, 27(6), 592–619.
- [12] Dahlback N., Jonsson A., Ahrenberg L., (1993), “Wizard of Oz studies: why and how”, In Proceedings of the 1st international conference on Intelligent user interfaces - IUI '93, ACM Press, 9(4), 193–200.

- [13] Hmoud B., Laszlo V., (2019), “Will artificial intelligence take over human resources recruitment and selection”, *Network Intelligence Studies*, 7(13), 21–30.
- [14] Dudler L., (2020), “Wenn Bots übernehmen–Chatbots im Recruiting. Digitalisierung im Recruiting: Wie sich Recruiting durch künstliche Intelligenz”, *Algorithmen und Bots verändert*, 101-111.
- [15] Maulsby D., Greenberg S., Mander R., (1993), “Prototyping an intelligent agent through wizard of oz”, In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 277–284, Amsterdam, Netherlands, 24-29 April.
- [16] Zamora J., (2017), “I’m sorry, dave, i’m afraid i can’t do that: Chatbot perception and expectations” In *Proceedings of the 5th international conference on human agent interaction*, 253-260, Newyork, USA, October.
- [17] Packialatham A., Thilagavathy R., Sarulatha, (2019), “An Interactive Medical Assistant using Natural Language Processing”, *5th International Conference On Science Technology Engineering & Mathematics (ICONSTEM 2019)*, 9 (1), 1-5, India, 14-15 March.
- [18] Luger E., Sellen A., (2016), "Like Having a Really Bad PA" *The Gulf between User Expectation and Experience of Conversational Agents*”, In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 5286-5297, San Jose, CA, USA, 7-12 May.
- [19] Grudin J., Jacques R., (2019), “Chatbots, Humbots, and the Quest for Artificial General Intelligence”, In *CHI 2019 Conference Proceedings 1-12*, Glasgow, UK, 4-9 May.
- [20] Toxtli C., Monroy-Hernández A., Cranshaw J., (2018), “Understanding chatbot-mediated task management”, In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 542-554, Montreal, CA, 21-26 April.
- [21] Eißer J., Meurer S., (2020), “Wizard-of-Oz Testing as an Instrument for Chatbot Development: An experimental Pre-study for Setting up a Recruiting Chatbot Prototype” In *CENTRIC 2020 : The Thirteenth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services*, 48-56, Porto, Portugal, 18-22 October.
- [22] Bradeško L., Mladenčić D., (2012), “A survey of chatbot systems through a Loebner Prize competition”, In *Proceedings of the Slovenian Language Technologies Society Eighth Conference of Language Technologies*, 2(5), 34-37.
- [23] Weizenbaum J., (1966), “Eliza- a Computer Program for the Study of Natural Language Communication between Man and Machine”, *Communications of the ACM*, 9(1), 36-45.

- [24] Pinar Saygin A., Cicekli I., Akman V, (2000), "Turing test: 50 years later", *Minds and Machines*, 10(4), 463-518.
- [25] Shawar B. A., Atwell E., (2007), "Chatbots: Are They Really Useful?", *Journal for Language Technology and Computational Linguistics*, 22(1), 29-49.
- [26] Ferrucci D., Brown E., Chu-Carroll J., Fan J., Gondek D., Kalyanpur A. A., Welty C., (2010), "Building Watson: An overview of the DeepQA Project" *AI magazine*, 31(3), 59-79.
- [27] Lewis S. C., Guzman A. L., Schmidt T. R., (2019), "Automation, journalism, and human-machine communication: Rethinking roles and relationships of humans and machines in news", *Digital Journalism*, 7(4), 409-427.
- [28] Web1, (2023), <https://www.businessofapps.com/data/chatgpt-statistics/>, (Erişim Tarihi: 22/02/2023).
- [29] Ulrich K. T., Eppinger S. D., (2012), "Concept selection", *Prod Des Dev 5th Ed Phila McGraw-HillIrwin*, 1, 145-61.
- [30] de Los Santos C. B., Krause-Jensen D., Alcoverro, T. Marbà N., Duarte C. M., Van Katwijk M. M., Santos R., (2019), "Recent trend reversal for declining European seagrass meadows", *Nature communications*, 10(1), 3356.
- [31] Meyer von Wolff R., Hobert S., Schumann M., (2019), "How may i help you?—state of the art and open research questions for chatbots at the digital workplace", *Hawaii International Conference on System Sciences*, 95-104, Grand Wailea, Hawaii, 8 – 11 January.
- [32] Schumaker R. P., Liu Y., Ginsburg M., Chen H., (2006), "Evaluating mass knowledge acquisition using the alice chatterbot: The az-alice dialog system", *International journal of human-computer studies*, 64(11), 1132-1140.
- [33] Web 2, (2023), <https://medium.com/nerd-for-tech/nlp-nlu-and-nlg-7aa3d8d64e2f>, (Erişim Tarihi: 06/02/2023).
- [34] Mueller J., Thyagarajan A., (2016), "Siamese recurrent architectures for learning sentence similarity", In *Proceedings of the AAAI conference on artificial intelligence*, 30 (1), 2786-2792.
- [35] Maroengsit W., Piyakulpinyo T., Phonyiam K., Pongnumkul S., Chaovalit P., Theeramunkong T., (2019), "A survey on evaluation methods for chatbots" In *Proceedings of the 2019 7th International conference on information and education technology*, 111-119, Aizu-Wakamatsu, Japan, 29-31 March.
- [36] Sari A. C., Virnilia N., Susanto J. T., Phiedono K. A., Hartono T. K., (2020), "Chatbot developments in the business World", *Advances in Science, Technology and Engineering Systems Journal*, 5 (6), 627-635.

- [37] Gupta A., Hathwar D., Vijayakumar A., (2020), “Introduction to AI chatbots”, *International Journal of Engineering Research and Technology*, 9 (7), 255-258.
- [38] Feine J., Morana S., MaedcheA., (2020), “Designing interactive Chatbot development systems”, In *Proceedings of the 41st International Conference on Information Systems (ICIS)*, 1–18, AISel, India, 13-16 December.
- [39] Bernsen N. O., Dybkjær H., Dybkjær L., (1996), “Cooperativity in human-machine and human-human spoken dialogue”, *Discourse processes*, 21(2), 213-236.
- [40] Landauer T. K., (1996), “The Trouble with Computers”, Cambridge, Massachusetts: The MIT Press.
- [41] Web 3, (2020), <https://rocket.chat/> (Erişim Tarihi: 16/07/2022)
- [42] Riek L., (2012), “Wizard of oz studies in HRI: A systematic review and new reporting guidelines,” *Journal of Human-Robot Interaction*, 1 (1), 119–136.
- [43] Web 4,(2012), <https://www.nngroup.com/articles/howmany-test-users/>, (Erişim Tarihi 07/06/2022).

ÖZGEÇMİŞ

Osman Utar, 2012 yılında başladığı Sakarya Üniversitesi Bilgisayar ve Bilişim Bilimleri Fakültesi Bilgisayar Mühendisliği Bölümünü 2016 yılında başarıyla tamamlayarak 2019 yüksek lisans eğitimine Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim dalında başladı. 2012 -2019 yılları arasında GTÜ Bilgi İşlem Dairesinde çalışmıştır. Halen GTÜ Bilgisayar Mühendisliği Bölümünde çalışmaktadır.



EKLER

Ek A: Tez Çalışması Kapsamında Yapılan Yayınlar

Utar O., Gokturk M., (2022), “FAQ Chatbot Prototype Creation with Wizard of Oz Method” in International Eastern Conference on Human-Computer Interaction, IECHI 2022 (Nahçıvan -Azerbaijan)

Utar O., Gokturk M., (2022), “Comparison Of Chatbots With Wizard Of Oz Technique” in International Conference on Engineering Natural and Applied Science, ICENAS 2021, (Osmaniye- TURKEY)

