

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Martin Garbuz

**Inimloetavate aruannete genereerimine
töötajate tagasiside
andmete põhjal Motichecki näitel**

Bakalaureusetöö

Juhendaja: Toomas Lepikult
PhD

Tallinn 2023

Autorideklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Martin Garbuz

24.04.2023

Annotatsioon

Käesolev lõputöö käsitleb inimloetavate aruannete genereerimist töötajate tagasiside küsitluste põhjal. Nimetatud funktsionaalsust arendatakse Eesti iduettevõttele Moticheck, mis pakub tarkvara töötajate rahulolu monitoorimiseks. Autor analüüsib erinevaid variante probleemi lahendamiseks ning otsustab ChatGPT keelemudeli kasuks. Töös kirjeldatakse, kuidas integreerida ChatGPT veebiteenus Motichecki rakendusse, ning milliste aspektidega peaks suurte keelemudelite kasutamise puhul arvestama. Muuhulgas käsitleb autor suurte keelemudelite kasutamisega seotud riske, ent samuti ka seda, kuidas luua täiendavat sünergiat ja lisandväärtust olemasoleva süsteemi ja ChatGPT integreerimise abil. Töö hõlmab kõnealuse funktsionaalsuse arendamise kirjeldust, sealhulgas analüüsi ja arendusprotsessi kirjeldust.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 30 leheküljel, 5 peatükki, 9 joonist, 2 tabelit.

Abstract

Generating Human-Readable Reports from Employee Feedback Data: the Case of Moticheck

The thesis discusses how to generate human-readable reports based on employee survey results. This feature was required for an Estonian startup Moticheck, which provides software for monitoring employee satisfaction. The thesis discusses different potential solutions to the problem, before opting to use the ChatGPT language model. It also describes, how to integrate the ChatGPT web service with Moticheck application, and what kind of questions should be considered when using large language models. The thesis also discusses some risks associated with using large language models, as well as how to create synergy and added value while integrating ChatGPT with existing applications. The development of the aforementioned functionality is described in detail, including analysis and development process.

The thesis is in Estonian language and contains 30 pages of text, 5 chapters, 9 figures, 2 tables.

Lühendite ja mõistete sõnastik

API	<i>Application programming interface</i> – liides, mille abil saab üks rakendus suhelda teisega
BERT	<i>Bidirectional Encoder Representations from Transformers</i>
GPT	<i>Generative pre-trained transformer</i>
HTTP	<i>Hypertext transfer protocol</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
LAMDA	<i>Language Model for Dialogue Applications</i>

Sisukord

1	Sissejuhatus.....	10
2	Probleemi kirjeldus.....	12
2.1	Tausta kirjeldus.....	12
2.1.1	Ettevõtte äriline taust.....	12
2.1.2	Tehniline taust.....	13
2.2	Eesmärgi üldkirjeldus.....	13
2.3	Lõputöö skoop.....	13
3	Analüüs.....	15
3.1	Ärilise vajaduse analüüs.....	15
3.2	Kasutajalugude kirjeldus.....	16
3.3	Lahenduse valik.....	18
3.3.1	Sisulise töö teevad ära inimesed.....	18
3.3.2	Tehisintellekti teenuse kasutamine.....	19
3.3.3	Keelemudeli treenimine.....	19
3.3.4	Aruannete genereerimine ette.....	19
3.3.5	Lahendusvariantide võrdlus.....	19
3.4	Välise teenuse valik.....	22
3.5	ChatGPT. Detailsem ülevaade.....	24
3.6	ChatGPT kasutamise riskid.....	25
3.7	Integratsiooni arhitektuuri valik.....	26
4	Töö käik.....	28
4.1	Töö teostamine.....	28
4.1.1	Eeltöö.....	28
4.1.2	Algse prototüübi loomine.....	29
4.1.3	Ettevalmistused testkeskkonda paigaldamiseks.....	29
4.1.4	Mentori sõnumite ajalugu.....	31
4.1.5	Testimine ja tagasiside.....	34
4.1.6	Interaktiivse vestluse funktsionaalsuse lisamine.....	34

4.2 Töö tulemused.....	36
5 Kokkuvõte.....	39
Kasutatud kirjandus.....	41
Lisa 1– Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks.....	42
Lisa 2 – Katsetused ChatGPT-ga.....	43
Lisa 3 – Töö käigu üksikasjalikum kirjeldus.....	50
Lisa 4 – Kasutajalood.....	60

Jooniste loetelu

Joonis 1. Mentori vestluste ajaloo andmebaasiskeem. Variant 1.....	32
Joonis 2. Mentori vestluste ajaloo andmebaasiskeem. Variant 2.....	33
Joonis 3. Klientide tagasiside mentori kasutamise kohta tulevikus.....	37
Joonis 4. Klientide valmisolek maksta mentori funktsionaalsuse eest.....	38
Joonis 5. ChatGPT vastus vene keeles.....	48
Joonis 6. ChatGPT vastus inglise keeles.....	48
Joonis 7. Küsimuse kaart Motichecki rakenduses.....	51
Joonis 8. Küsimuse kaart koos uue nupuga, mis kuvab tehisintellekti soovitusi.....	52
Joonis 9. Mentori küsimuste ajaloo kasutajaliides.....	58

Tabelite loetelu

Tabel 1. Lahendusvariantide võrdlus.....	22
Tabel 2. Katsetused ChatGPT-ga.....	43

1 Sissejuhatus

Käesolevas lõputöös uuritakse, kuidas genereerida töötajate tagasiside andmete põhjal inimloetavaid aruandeid, mis aitaksid tagasisidet tõlgendada ning anda soovitusi olukorra parandamiseks. Selline funktsionaalsus lisatakse Eesti idufirma Motichecki rakendusele, mis tegeleb töötajate rahulolu monitoorimisega.

Eesmärgi saavutamiseks loob autor Motichecki integratsiooni ChatGPT juturoboti veebiteenusega. Seetõttu on osa tööst pühendatud sellele, et analüüsida ChatGPT-d rakendusse integreerimise kontekstis.

ChatGPT on 2022. aastal üldsusele kättesaadavaks tehtud suurel keelemudelil põhinev juturobot, mis on pälvinud erakordselt suurt tähelepanu. See on siiani kõige kiiremini kasvanud kasutajate arvuga tarkvara. Tegemist on murrangulise tehnoloogiaga, mis pakub ühiskonnale suuri võimalusi, aga millega kaasnevad ka teatavad riskid.

Käesolev lõputöö võiks huvi pakkuda neile, kes kaaluvad ChatGPT integreerimist mõnda rakendusse. Lõputöö näitab ühte võimalust, kuidas sellisele küsimusele läheneda ning mis aspekte arvestada. Võib suure kindlusega oletada, et lähitulevikus soovitakse üha rohkem ChatGPT-d või muid suuri keelemudeleid kasutada väga erinevates rakendustes, kuna nende potentsiaalne kasutusala on väga lai. Loodetavasti annab käesolev lõputöö oma panuse selle kiiresti areneva valdkonna uurimisse.

Lõputöö koosneb 5 peatükist. Pärast sissejuhatust kirjeldab autor teises peatükis, millist probleemi täpsemalt lahendatakse. Samuti kirjeldab autor lühidalt Motichecki ettevõtet ning piiritleb töö skoopi.

Kolmandas peatükis analüüsib autor täpsemalt ärilist vajadust, mida tuleb lahendada. Selle pinnalt paneb autor paika loodava arenduse nõuded ja koostab kasutajalood. Järgnevalt kirjeldab autor erinevaid lahendusvariante paika pandud tarkvaranõuete saavutamiseks, ning jõuab järeldusele, et kõige otstarbekam on kasutada mõnda tehisintellekti teenust. Seejärel toimub konkreetse tehisintellekti teenuse valik, mis

langeb ChatGPT kasuks. Lisaks sisaldab kolmas peatükk detailsemat ülevaadet ChatGPT kohta ning eraldi alapeatükis analüüsitakse ka selle kasutamisega seotud riske. Kolmanda peatüki viimases alapeatükis vaadeldakse integratsiooni arhitektuuri.

Neljandas peatükis kirjeldab autor töö käiku: kuidas arendus teostati ning milline oli tulemus. See peatükk sisaldab lahenduse kohta tehnilisemaid üksikasju.

Lõputöö põhiosa lõpetab kokkuvõtte. Lisaks kokkuvõttele on lõputööl lisad, sealhulgas katsetuste näidised ChatGPT-ga, töö käigu üksikasjalikum kirjeldus, mis ei mahtunud põhiosa sisse, ning detailsemad kasutajalood.

2 Probleemi kirjeldus

Käesolevas töös otsitakse vastust küsimusele, kuidas genereerida töötajate tagasiside küsitluste põhjal inimloetav aruanne, mis annaks töötajate juhile tagasisidest ülevaate ning soovitusi olukorra parandamiseks.

Olukorra näitlikustamiseks võib ette kujutada mõne ettevõtte osakonnajuhti, kelle töötajate seas viidi läbi küsitlus „Kas te olete rahul oma töötingimustega”. Osakonnajuht näeb, et küsimusele on vastanud 20 töötajat, kellest 60% on vastanud jaatavalt. Eesmärk oleks leida viis genereerida nende andmete pinnalt osakonnajuhi jaoks inimloetav ülevaade, mis aitaks vastuseid tõlgendada ning annaks soovitusi, kuidas olukorda parandada.

2.1 Tausta kirjeldus

Järgnevalt kirjeldab autor lõputöö aluseks olevat tausta. Autor kirjeldab, millise ettevõtte jaoks seda lahendust vaja oli ning annab lühiülevaate ettevõtte tehnilisest taustast.

2.1.1 Ettevõtte äriline taust

Nimetatud probleem tekkis Eesti idufirmal Moticheck, mis pakub töötajate rahulolu monitooringu platvormi. Ettevõtte ärimudel on järgmine. Kui mõni ettevõtte soovib oma töötajate rahulolu paremini monitoorida ja hakkab Motichecki teenust kasutama, siis hakkavad selle ettevõtte töötajad saama iga päev e-mailile ühe küsimuse, millele saab vastata „jah” või „ei”. Küsimused puudutavad selliseid teemasid nagu rahulolu töökeskkonnaga, meeskonnasisese kommunikatsiooniga, palgaga jne. Soovi korral saavad töötajad lisada ka kommentaari. Tagasiside andmine on anonüümne. Ettevõttepoolne esindaja, milleks võib olla näiteks personalitöötaja või osakonnajuhataja, saab näha oma ettevõtte töötajate vastuste statistikat. See võimaldab tuvastada valdkondi, millega töötajad ei ole rahul ning võtta olukorra parandamiseks

midagi ette. Motichecki eelis traditsiooniliste töötajaküsitluste ees seisneb selles, et Motichack monitoorib töötajate rahulolu pideva protsessina, samas kui traditsioonilised küsitlused annavad pildi mingist konkreetsest ajahetkest. Tänu sellele saab Motichackis jälgida rahulolu muutumise dünaamikat.

2.1.2 Tehniline taust

Motichecki rakendus koosneb kahest osast. Tagarakendusest, mis on kirjutatud Kotlini keeles Ktor raamistiku abil ning esirakendusest, mis on kirjutatud Reacti teeki kasutades. Esirakendus võtab tagarakendusest andmeid HTTP protokolliga kaudu. Seega on tegemist tänapäeval väga levinud arhitektuuriga.

2.2 Eesmärgi üldkirjeldus

Järgnevalt vaatleb autor käesoleva töö raames tehtava arenduse eesmärgi. Eesmärk on leida viis genereerida töötajate tagasiside küsitluste põhjal inimloetav aruanne, mis annaks tagasisidest ülevaate ning soovitusi olukorra parandamiseks. See ülevaade tuleks teha Motichecki kasutajatele kasutajaliideses kättesaadavaks. Samuti võivad sellega kaasneda lisaarendused, et selline funktsionaalsus kasutajatele rohkem väärtust pakuks, näiteks võimalus küsida aruande kohta rakenduse käest täpsustavaid küsimusi, millele automaatselt vastatakse.

2.3 Lõputöö skoop

Lõputöö skoopi kuulub Motichecki rakenduse ja ChatGPT integratsioon. Motichecki rakendusse tuleb lisada võimalused, et kasutajad saaksid ChatGPT käest nõu küsida.

Lõputöö skoopi ei kuulu Motichecki rakenduse loomine nullist. See rakendus on juba olemas ja kasutusel. Käesoleva lõputöö skoopi kuulub üksnes olemasolevale rakendusele täiendava funktsionaalsuse lisamine.

Samuti ei kuulu lõputöö skoopi ChatGPT või muude keelemudelite arendus. Keelemudeleid käsitatakse lõputöö raames pigem kui musta kasti, mis saab tekstilise sisendi ning tagastab tekstilise väljundi. Siiski kuulub lõputöö skoopi see, millise

konkreetses sõnastusega ning milliste parameetritega keelemudeli teenuse poole pöörduda, et saada kõige kasulikum vastus Motichecki rakenduse kasutajate jaoks.

3 Analüüs

Käesolevas peatükis annab autor ülevaate ärilise vajaduse analüüsist, kasutajalugudest ning kirjeldab tehnoloogiate valikut. Iga tarkvara loomine peaks algama vajaduste kaardistamisest, mis tulenevad lahendamist vajavast probleemist. Probleemi saab täpsemalt sõnastada, kui vaadelda ärilist vajadust, mida on tarvis lahendada. Selle põhjal saab ära määratleda loodava tarkvara nõuded. Nõuete põhjal saab valida kontseptuaalse lahenduskäigu, ning seejärel saab valida tehnoloogilised vahendid, mille abil on lahendust kõige otstarbekam teostada. Seega on tegemist mitmesammulise analüüsiga, kus iga sammu tulemused on sisendiks järgmise sammu analüüsile.

3.1 Ärilise vajaduse analüüs

Motichecki rakendus pakub ettevõtetele võimalust monitoorida oma töötajate rahulolu ning selle muutumist. Ettevõtte juhid näevad statistikat, kuidas on töötajad erinevatele küsimustele vastanud ning kuidas on rahulolu dünaamika ajas muutunud. Küsimused on jagatud erinevatesse kategooriatesse, nagu vaimne tervis, kaasamine, missioon ja väärtused, sisekommunikatsioon jne. Seega on juhil võimalik jälgida ka ettevõtte käekäiku mõnes konkreetnes eelnimetatud valdkonnas. Vaatamata sellele, et Motichack annab ettevõtte juhtide käsutusse suure hulga statistikat töötajate rahulolu kohta, tekib mõnedel juhtidel sageli siiski küsimus, mida selle statistikaga peale hakata ja mis sellest täpsemalt järeldada võiks. Samuti võib tekkida küsimus, mida võtta ette olukorra parandamiseks.

Kui Motichack saaks juhte neis küsimusis aidata, pakuks see Motichecki klientide jaoks täiendavat väärtust. Nad saaksid töötajate tagasiside tulemusi paremini tõlgendada ning võtta ette samme olukorra parandamiseks. See võiks suurendada töötajate rahulolu ja motivatsiooni, mis kajastuks lõpuks ka suuremas produktiivsuses ja paremates töötulemustes.

Üks võimalus juhtidele tulemuste tõlgendamisel abi pakkuda oleks kaasata professionaalseid ärinõustajaid, kes pakuksid juhtidele konsultatsioone. See on aga üsna kulukas lahendus, kuna vastavate professionaalide töötund on suhteliselt kallis.

Samuti selgus Motichecki esindajaga konsulteerides, et nõustajaid ei ole kohe võtta ja ühe isoleeritud mure lahendamiseks ei palgata mentoreid ja *coache*. Motichheck soovib aga toetada juhtimiskvaliteeti agiilselt ning ootamata, mil olukord suuremaks probleemiks eskaleerub. Praktikas on vaja otsida lahendusi, mis pakuvad juhile abi koheselt, on ligipääsetavad ja lihtsad kasutusele võtta. Klientide ootused viitavad vajadusele saada teenust nüüd ja kohe. Samuti on tehnoloogiasektori ettevõtetel raske integreerida oma toodetesse ja teenustesse lahendusi, mis sisaldavad klassikalisi konsultatsiooni teenuseid. Need ei ole skaleeritavad rahvusvahelisel turul.

Seega seisab Motichheck silmitsi väljakutsega automatiseerida juhtidele suunatud mentorlus ja coaching. Selleks soovitakse uurida tehisintellekti kasutamise võimalusi. Tehisintellekt on inimesest odavam kasutada, universaalsem, kiirem ja palju lihtsamini skaleeritav lahendus.

Motichheck kogub oma klientidele suures mahus statistikat mõistmaks kliendi töötajate tööelukogemust. Selle andmekogu integreerimine mõne tehisintellekti teenusega, mis aitaks ettevõtte juhtidel töötajate tagasiside tulemusi tõlgendada ning anda soovitusi olukorra parandamiseks on unikaalne võimalus nii Motichheckile kui ka tema klientidele. Seega ongi äriline vajadus luua integratsioon mõne tehisintellekti teenusega, mis aitaks töötajate küsitluste andmeid tõlgendada.

3.2 Kasutajalugude kirjeldus

Kui äriline vajadus on kaardistatud, on järgmiseks sammuks määratleda plaanitava tarkvaraarenduse nõuded. Nõuete sõnastamist võib pidada tarkvara loomise protsessi kõige keerulisemaks osaks; kõige keerulisem on otsustada, mida täpselt on tarvis arendada, kusjuures nõuete määratlemise ajal tehtud vead on kõige suurema mõjuga[6].

Nõuete spetsifitseerimiseks on erinevaid lähenemisi. Näiteks on olemas standard IEEE 830, mida peetakse äärmiselt põhjalikuks ning mis on kasutusel suure vastutusega

valdkondades, nagu pangandus, lennundus jne. Samas on olemas ka agiilsed lähenemised, mis põhinevad rohkem kasutajalugudel ning soodustavad pigem iteratiivset ja inkrementaalset arendust, kus vajadused ja eesmärgid võivad sageli muutuda. On teada, et tarkvaraarenduse projekti kulgu ei saa täpselt ette ennustada, ning sedamööda, kuidas tellijad näevad tarkvara esimesi versioone, tekivad neil sageli uued soovid ja nende arvamused muutuvad [7].

Antud töö eesmärkide saavutamiseks sobib just nimelt agiilne meetod, kuna Motichecki näol on tegemist idufirmaga ning arendatav funktsionaalsus on suuresti turu testimise ja avastamise faasis, kus täpsed vajadused võivad muutuda ja aja jooksul selgineda. Seetõttu on tarkvara nõudeid mõistlik määratleda kasutajalugude kaudu.

Kasutajalood on lühikesed kirjeldused funktsionaalsusest, mis loob tarkvara kasutaja jaoks väärtust. Kasutajalood koosnevad lühikirjeldusest, mis on aluseks arendaja ja tellija vahelisele vestlusele nõuete täpsustamiseks. Lisaks on kasutajaloo osaks vastuvõtukriteeriumid ehk testid. Kasutajaloo skoop peaks olema selline, et selle arendamiseks kuluks aega maksimaalselt kaks nädalat. Kui aega kulub rohkem, tuleks kasutajalugu jagada väiksemateks osadeks. [8]

Pärast esmaseid kohtumisi ja arutelusid Motichecki esindajaga, said määratletud järgmised kasutajalood.

Kasutajalugu 1.

Ettevõtte juhina ma tahaksin,

et töötajate tagasisidet kajastava küsimuse kaardil oleks nupp, millele vajutades näeksin soovitusi, mida selles olukorras teha,

et ma saaksin olukorda parandada.

Kasutajalugu 2.

Juhina ma tahaksin

näha oma AI chatiga peetud vestluste ajalugu,

et saaksin vajadusel huvipakkuva info hiljem üles leida.

Kasutajalugu 3.

Juhina ma tahaksin,

et küsimuse kaardil oleks nupp, millele vajutades saaksin alustada AI chatiga interaktiivset vestlust,

et ma saaksin täpsemat nõu mulle huvipakkuvate teemade kohta.

Antud kasutajalood on koostatud kasvava keerukuse põhimõttel. Kasutajalugu 1 on kõige lihtsam. Kasutajalugu 2 ja kasutajalugu 3 lisavad täiendavat keerukust. See võimaldab arendada agiilselt ja iteratiivselt, saades tihedamini tagasisidet. Sage tagasiside võimaldab vajadusel muuta kurssi varem kui muidu ning vältida ebaefektiivset ajakasutust. Arendusmetoodika seisukohast on eesmärgiks luua iga kasutajaloo kohta võimalikult kiiresti prototüüp, et Motichecki esindajad ning vajadusel ka kliendid saaksid hakata seda funktsionaalsust testimise ja tagasisidet andma. Saadud tagasiside põhjal saab arendusprotsessi suunata nii, et see looks lõppkasutajatele rohkem väärtust.

3.3 Lahenduse valik

Käesoleva alapeatüki eesmärk on analüüsida, kuidas on eelnevalt püstitatud ärilist eesmärki kõige mõistlikum saavutada. Kuigi esmane lahendus, mis esile kerkib, on integratsiooni loomine mõne tehisintellekti teenusega, siis tasub enne lõpliku otsuse tegemist vaadelda ka alternatiive. Siinkohal oleks sobilik vaadelda nelja võimalikku lahendust.

3.3.1 Sisulise töö teevad ära inimesed

Esimene võimalus oleks lahendada probleem hoopis inime tööjõu abil, näiteks kaasates ärinõustamise spetsialiste. Sellisel juhul ei oleks küll rakenduse kasutaja jaoks tagasiside väga kiire, kuid ette võib kujutada näiteks sellist süsteemi: ettevõtte juht vajutab teda huvitava küsimuse kaardi juures nupule, mille abil saab küsida nõu. Selle peale saadetakse mingis keskkonnas pöördumine ärinõustajale, kus on info töötajate tagasiside küsimuse ning töötajate vastuste kohta (mitu töötajat vastas positiivselt ja

mitu negatiivselt, samuti töötajate lisatud kommentaarid). Nende andmete põhjal kirjutab ärinõustaja töötajate tagasiside tõlgenduse, järeldused ning soovitused olukorra parandamiseks. Ärinõustaja poolt kirjeldatud tekst saadetakse ettevõtte juhile näiteks e-kirjaga või tekib see tal Motichecki rakendusse eraldi vaatesse.

3.3.2 Tehisintellekti teenuse kasutamine

Teine võimalus on kasutada inimtööjõu asemel tehisintellekti, eelkõige suurtel keelemudelitel põhinevaid teenuseid. Sellisel juhul saaks inimtööjõu asendada tehisintellektiga. Kui ettevõtte juht vajutab nupule, et küsida nõu, pöördub Motichecki rakendus suurel keelemudelil põhineva välise teenuse poole ning annab kaasa info küsimuse kohta. Keelemudelilt saadud vastus kuvatakse ettevõtte juhile.

3.3.3 Keelemudeli treenimine

Kolmas võimalus on sarnane teise võimalusega, kuid selle asemel, et pöörduda olemasoleva välise teenuse poole, treenida hoopis vastav keelemudel ise.

3.3.4 Aruannete genereerimine ette

Neljas võimalus on muuta lahendus lihtsamaks. Nimelt saaks iga küsimuse kohta koostada nõuandvad tekstid juba ette ära. Iga küsimuse kohta võiks olla näiteks kaks teksti. Üks oleks positiivse tulemuse kohta ja teine oleks negatiivse tulemuse kohta. Sellisel juhul tuleks inimtööjõudu kasutada tekstide koostamiseks ühekordselt. Kui tekstid on koostatud, saaks kasutajale näidata ette koostatud tekste näiteks sõltuvalt töötajate tagasiside tulemusest. Näiteks kui üle 70% vastustest olid positiivsed, siis näidata teksti positiivset versiooni. Kui positiivseid vastuseid oli alla selle, siis näidata negatiivset versiooni.

3.3.5 Lahendusvariantide võrdlus

Kolmanda variandi võib kohe kõrvale heita. Suure keelemudeli treenimine iseseisvalt on ilmselgelt liiga keeruline, ajamahukas ja kulukas ülesanne. Näiteks ChatGPT mudeli treenimine maksis hinnanguliselt 30 miljonit dollarit, olgugi et täpsed andmed puuduvad[4]. Tervikliku käsitluse huvides on sobilik see variant siiski ära mainida.

Esimese variandi puhul on selgeks miinuseks suur tööjõumahukus. Kuigi professionaalse inimese poolt tehtud analüüs, tõlgendus ja soovitused oleksid tõenäoliselt selgelt kõige parema kvaliteediga, oleks selline protsess ka oluliselt kallim kui tarkvaraline lahendus. Täpsed arvutused selle kohta puuduvad, kuid seda võib eeldada suure kindlusega, kuna kvalifitseeritud inime tööjõud on üldjuhul kallim kui tarkvaralised lahendused. Samuti oleks esimene variant kliendi jaoks aeglasem, kuna inimloetava aruande koostamine võtab inimesel kauem aega kui arvutil. Inimesel kuluks aruande kirjutamiseks vähemalt 10 minutit. Samuti peab arvestama, et tal ei pruugi olla võimalik kirjutamisega kohe alustada. Lisaks võib tekkida tööde kuhjumine. Võrdluseks, tehisintellekti abil kuluks sarnase teksti genereerimiseks umbes 1 minut ning neljanda variandi puhul, kus tekstid on eelnevalt valmis tehtud, oleks ajakulu minimaalne. Oluline miinus esimese variandi puhul seisneb ka selles, et selline lahendus ei oleks lihtsasti skaleeritav suurele kasutajate hulgale. Potentsiaalselt võiks sellest variandist olla kasu, et valideerida tehisintellekti poolt genereeritud sisu, s.o. paluda mõne üksiku küsimuse kohta kirjutada aruanne inimesel, ning võrrelda inimese tehtud aruannet sama küsimuse kohta tehtud tehisintellekti aruandega.

Teise variandi puhul oleks esimesega võrreldes eeliseks, et see on kiirem ja odavam. Töö kirjutamise ajal kehtinud hinnakirja järgi maksab nt ChatGPT API kasutamine 0,01 dollarit 5000 tokeni ees (üks token vastab natuke vähem kui ühele sõnale). Seega maksaks isegi väga pika aruande kirjutamine vaid 0,01 dollarit, tegelikkuses saaks aga selle eest tõenäoliselt mitu aruannet. Kui võrrelda, et inimesel kuluks ühe aruande kirjutamise peale näiteks 10 minutit ja arvestada inime tööjõu tunnihinnaks näiteks 60 eurot, siis kuluks ühe aruande peale ligikaudu 10 eurot. Vahe on mitu suurusjärku. Nagu eelpool mainitud, annaks suure keelemudeli kasutamine ka olulise kiiruse eelise. Täpne päringule vastuse saamise kiirus sõltub mitmest asjaolust, muuhulgas sellest, kui pikk peab vastus olema ja kui koormatud on parasjagu keelemudeli serverid, kuid peaks üldjuhul jääma ühe minuti piiresse. Eeliseks on ka see, et selline lahendus oleks lihtsasti skaleeritav suurele kasutajate hulgale. Potentsiaalne puudus võib olla seotud teenuse kättesaadavusega. Kuna suurte keelemudelite vastu on hetkel väga suur huvi ning kasutajaid palju, võib teenus olla ajutiselt aeglane või koguni kättesaamatu. Ent võib eeldada, et tulevikus peaksid sellised probleemid vähenema, kuna tegemist on uudse ja perspektiivika tehnoloogiaga ning kõigi eelduste kohaselt käib töö vastavate teenuste

jõudluse suurendamiseks. Suurte keelemudelite puhul on tõsine küsimus ka nende usaldusväärsus, mis vajab eraldi analüüsi.

Neljanda lahenduse puhul oleks vaja ühekordset mahukat inimtööjõu investeeringut, et kirjutada valmis aruande šabloonid erinevate küsimuste kohta. Edaspidi oleks aga protsess üsna odav, kiire ning kasutajate kasvuga skaleeritav. Siiski oleks täiendavat lisatööd vaja iga kord, kui Motichecki süsteemi lisanduvad mõned uued küsimused. Kõige suurem puudus on tõenäoliselt aga selles, et taoline süsteem oleks lihtsakoelisem. Esiteks tuleks käsitsi ära määratleda lävendid, alla mille peetakse tulemust negatiivseks ja üle mille positiivseks. Selline määratlus oleks teatud määral subjektiivne ja ei pruugi kehtida iga küsimuse kohta sarnaselt. Kui aga iga küsimuse kohta määratleda ka eraldi lävendid, läheks süsteem liiga keeruliseks. Samuti ei võimalda selline süsteem kajastada erinevaid nüansse ja varjundeid. Näiteks kui mõne tagasiside puhul ei tahaks seda klassifitseerida pelgalt positiivseks või negatiivseks, vaid tahaks välja tuua rohkem nüansse. Samuti oleksid aruanded konkreetse küsimuse kohta alati samasugused ja samasuguse sõnastusega, mis jätaks süsteemist masinlikuma mulje. Lisaks oleks esialgne investeering inimtööjõu näol väga mahukas, sest aruande šabloonid tuleks valmis kirjutada kõigi küsimuste kohta. Kui hiljem selgub, et kõnealune mentori funktsionaalsus ei leia kasutajate seas piisavalt nõudlust, läheb suur hulk tehtud tööd raisku. Teiste variantide puhul sellist riski taolisel kujul ei ole, kuna aruannete genereerimine toimub vastavalt kasutajate vajadusele jooksvalt.

Järgnevas tabelis on esitatud nelja variandi võrdlus aruannete kvaliteedi, rahalise maksumuse, kiiruse ja skaleeritavuse vaatepunktist. Igas kategoorias on kasutatud kolmepallilist hindamisskaalat.

Tabel 1. Lahendusvariantide võrdlus.

Lahendus	Aruannete kvaliteet	Rahaline maksumus	Kiirus	Skaleeritavus	Kokku
Töö teeb inimene	3	2	1	1	7
Tehisintellekti teenus	2	3	2	3	10
Enda treenitud keelemudel	2	1	2	3	8
Eelnevalt genereeritud aruanded	1	2	3	2	8

Arvestades eelnevalt esitatud analüüsi, on kõnealuse ärilise vajaduse lahendamiseks kõige mõistlikum valik kasutada välist tehisintellekti teenust. See tagab piisava aruannete kvaliteedi, on suhteliselt soodne, kiire ja lihtsalt skaleeritav. Kõigi ülejäänud lahenduste puhul kujunes vähemalt ühes kategoorias hindeks 1 pall, mis tähendab sisuliselt seda, et lahendus ei ole realselt vastuvõetav.

3.4 Välise teenuse valik

Eelmises alapeatükis jõudis autor järelduseni, et Motichecki ärilise vajaduse teostamiseks oleks kõige otstarbekam kasutada välist tehisintellekti teenust, mis aitaks genereerida töötajate tagasiside andmete põhjal inimloetavaid ülevaateid, tõlgendusi ning anda vajadusel soovitusi olukorra parandamiseks.

Kõige tuntum avalikkusele kättesaadav näide sellisest tehisintellekti teenusest on OpenAI poolt välja arendatud ChatGPT. See jõudis avalikkuseni 2022. aasta novembris [9]. 2023. aasta märtsis tuli välja ka API, mille abil saab ChatGPT-d liidestada muude rakendustega.

ChatGPT on juturobot, mis põhineb OpenAI poolt 2020. aastal avalikustatud GPT-3 keelemudelil. GPT-3 keelemudelit iseloomustab väga suur tekstide hulk, mille abil seda treeniti. Treenimise käigus kasutati umbes 175 miljardit parameetrit[4]. ChatGPT kasutab aga mitte GPT-3, vaid GPT-3.5-turbo mudelit.

Lõputöö kirjutamise hetkel on ChatGPT tõenäoliselt ainus avalikult kasutatav lahendus, mis suudaks käesolevas töös püstitatud probleemi lahendada. Ent kuna tegemist on revolutsioonilise tehnoloogiaga, mis on pälvinud ka väga palju avalikkuse tähelepanu ning kogunud kiiresti väga palju kasutajaid, üritavad OpenAI konkurendid tulla turule oma alternatiividega. Seetõttu on mõistlik vaadelda ka neid. Käesolevas töös vaatleb autor kahte alternatiivi: Google'i Bard ja Microsofti Bing.

Google on välja tulnud oma alternatiiviga Bard. Nii nagu ChatGPT aitab parendada Microsofti Bingi otsingumootori tulemusi, on tulevikus plaanis kasutada Bardi Google otsingu parendamiseks. Google Bard põhineb LaMDA (Language Model for Dialogue Applications) tehnoloogial. See kasutab transformeerivate närvivõrkude arhitektuuri, nagu ka GPT-3 keelemudel. Google Bardi tutvustati avalikkusele esmakordselt 6. veebruaril 2023, kuid esitluse käigus vastas ta valesti ühele kosmoseteemalisele küsimusele, mistõttu pälvis palju kriitikat. Selle tagajärjel langes Google'i aktsia väärtus mitu protsendipunkti. Tulevikus on plaanis teha võimalikuks Bardi liidestamine veebilehtede ja rakendustega, kuid hetkel sellist võimalust veel ei ole. [1]

Oma tootega on välja tulnud ka Microsoft: juturobot Bing. See põhineb samadel keelemudelitel, mis ChatGPT, kuid esimesed katsetused on andnud negatiivsemaid tulemusi. Näiteks on robot oma vestluskaaslasi ähvardanud[2] . See võib viidata asjaolule, et keelemudelit pole piisavalt treenitud inimeste tagasiside abil, et klassifitseerida sobimatud vastused.

Seega on hetkel probleeme nii Google'i Bardil kui ka Microsofti Bingil. Tõsi, probleeme esineb ka ChatGPT-ga. Näiteks ei osanud see vastata küsimusele: „Mis on Laura ema ainsa tütre nimi” [autori tõlge]. See näitlikustab selliste mudelite piiratust. [3]

Siiski võib väita, et ChatGPT on küpsem toode kui Bard või Bing, kuna seda on eeldatavasti kasutatud palju rohkem ja selle vigadest ei ole niivõrd markantseid näiteid nagu ülejäänud kahest juturobotist. Samas tuleb arvestada, et kuna tegemist on kiiresti arenevate tehnoloogiatega, võib olukord väga dünaamiliselt muutuda. Siiski on ka praktilisem põhjus, miks eelistada käesoleva lõputöö jaoks ChatGPT-d. Ülejäänud kaks ei ole veel isegi üldsusele vabalt kättesaadavad, vaid on avatud piiratud ringile testijatele. ChatGPT-d saab aga juba API abil liidestada olemasolevate rakendustega.

Seega on käesoleva töö kontekstis otstarbekas kasutada välise tehisintellekti teenusena just ChatGPT-d. Järgnevalt annab autor ChatGPT-st detailsema ülevaate ning käsitleb ka selle kasutamisega seotud riske.

3.5 ChatGPT. Detailsem ülevaade

Nagu eelnevalt mainitud, on ChatGPT OpenAI poolt välja arendatud juturobot, mis põhineb GPT-3.5 nimelisel keelemudelil. Teaduskirjandust, mis käsitleks ChatGPT tehnilisi aspekte, on siiani avaldatud suhteliselt vähe. Kui varasemalt on arenguid tehisintellekti valdkonnas kajastatud ka teadusajakirjades, näiteks Google'i poolt välja töötatud loomuliku keele töötlemiseks mõeldud närvivõrgu BERT puhul, siis ChatGPT puhul pole selle tehnilist poolt teadusajakirjanduses põhjalikult käsitletud ning üldsuse jaoks on avaldatud vaid lühike blogipostitus [10].

Selle tehnoloogia mõistmiseks, oleks esmalt kasulik vaadelda lähemalt suurte keelemudelite (ingl k *LLM – large language models*) olemust.

Tehnoloogiline murrang, mis tegi võimalikuks suurte keelemudelite kasutamise, toimus 2018. aastal. Siis võeti kasutusele enes juhendatud närvivõrkude masinõpe (*self-supervised neural network learning*), mis tähendas, et süsteem omandas teadmisi inimkeele kohta lihtsalt tänu sellele, et sai treeningandmeteks äärmiselt suure hulga tekste. Selle tekstihulga põhjal õppis süsteem ennustama, milline sõna järgneb mingisugusele sõnade jadale. Süsteem sai end iteratiivselt treenida ja seeläbi tulemusi pidevalt parandada. Piltlikult öeldes võis süsteem võtta mingi lause ja sealt ühe sõna kinni katta, et proovida seda ära arvata. Vigade põhjal sai süsteem oma mudelit muuta, et saavutada tulevikus paremaid tulemusi. [11]

Alates 2018. aastast on domineerivaks närvivõrgu tüübiks antud kontekstis transformeerivad närvivõrgud (*transformer neural networks*). Need närvivõrgud kujutavad kõike väga suuremõõtmeliste vektoritena ning suudavad õppida tagasilevi (*back-propagation*) abil. Kui kombineerida omavahel transformeerivatel närvivõrkudel põhinev mudel ning ülisuur hulk teksti, millega mudel saab harjutada ennustuste täpsust, on tulemuseks mudel, mis representeerib suurt osa inimkeele struktuurist. Sellised suured keelemudelid jätavad meelde palju fakte ning neid saab kasutada väga

erinevate ülesannete jaoks. Üheks sellistest mudelitest on GPT-3 (Generative Pre-trained Transformer-3). [11]

ChatGPT aluseks on niisil GPT-3.5 mudel, mille eelkäijaks on omakorda GPT-3. Lõputöö kirjutamise ajal ilmus aga juba ka GPT-4.

3.6 ChatGPT kasutamise riskid

ChatGPT ja sarnased suured keelemudelid üldisemalt on väga võimsad tööriistad, mida saab kasutada paljudes valdkondades. Ent nagu võimsate tööriistadega ikka, on selle kasutamisega seotud ka mitmeid tõsisid riske, millele on erinevates publikatsioonides juhitud üha rohkem tähelepanu.

2023. aasta märtsis tuli ettevõtte OpenAI välja uue mudeliga GPT-4. Seoses sellega avaldas ettevõtte ka publikatsiooni, milles käsitles GPT-4ga seonduvaid riske. Nende hulka kuuluvad hallutsinatsioonid, kahjulik sisu, valeinfo, konventsionaalsete ja ebakonventsionaalsete relvade kättesaadavuse soodustamine, isikuandmetega seotud riskid, küberturvalisus jm. [12]

Kõige tuntum suurte keelemudelitega seotud probleem on tõenäoliselt hallutsinatsioonid. Selle all mõistetakse sellise sisu loomist, mis on võrreldes teatavate allikatega vale, kusjuures see nähtus võib muutuda seda kahjulikumaks, mida enesekindlamaks mudelid muutuvad ja mida rohkem kasutajad neile toetuvad [12].

Seega on riskid tõsised ja enne sellise tehnoloogia kasutuselevõttu tuleks neid põhjalikult kaaluda. Samas võib välja tuua, et ChatGPT ja muude sarnaste teenuste kasutamise risk sõltub suurel määral ka sellest, millise konkreetse rakendusega neid liidestatakse. Näiteks on ilmselge, et kui kasutada sellist süsteemi meditsiinilises rakenduses, kus kasutajad saaksid küsida ChatGPT käest filtreerimata kujul meditsiinilist nõu, oleksid riskid suured. Samas on kahtlemata ka stsenaariumeid, kus riskid on suhteliselt väikesed, näiteks kui kasutada ChatGPT-d mõnes arvutimängus dialoogide või genereerimiseks, ennekõike seetõttu, et arvutimängude mängijad suhtuvad mängu sisusse eelkõige kui mängu, millest ei saa teha järeldusi reaalse elu kohta.

Motichecki rakenduse puhul tundub samuti, et riskid on väikesed. Kui vaadelda funktsionaalsust, kus ettevõtte juht saab vajutada nupule, et saada ChatGPT käest ülevaadet töötajate tagasiside kohta, ning soovitusi olukorra parandamiseks, siis on suhteliselt raske ette kujutada, mis sisu võiks ChatGPT genereerida, mis omaks kahjulikku mõju. Seni tehtud katsetuste käigus pole selliseid juhtumeid esinenud. Lisaks tuleb arvestada, et rakenduse kasutajad oleksid ettevõtte juhid ja osakonnajuhatajad, kellel on eeldatavasti piisavalt kriitilist mõtlemist, et mitte järgida pimesi iga ChatGPT soovitusi, vaid võtta neid pigem lisainfona, kust võib saada kasulikke mõtteid või ideid.

Kui aga vaadelda kasutajalugudes kirjeldatud funktsionaalsust, kus kasutaja saab lisaks ühekordsele vastusele pidada ChatGPT-ga ka interaktiivset vestlust ja küsida täpsustavaid küsimusi, siis seal tekib küll küsimus, kas peaks kasutaja küsimusi kuidagi piirama, sest teoorias võib kasutaja küsida ükskõik mida, mitte ainult töötajate juhtimisega seotud küsimusi. Samas ei suurendaks Moticheck tõenäoliselt riski oluliselt võrreldes kehtiva olukorraga, kuna ChatGPT on lõputöö kirjutamise hetkel tasuta teenus, mida kõik saavad kasutada. See, et seda saaks kasutada ka läbi Motichecki rakenduse, oleks ChatGPT praegust levikut arvestades vaid marginaalne lisandus.

3.7 Integratsiooni arhitektuuri valik

Eelnevalt jõudis autor järeldusele, et väliseks teenuseks, mis aitaks ülevaateid genereerida, oleks kõige mõistlikum valida ChatGPT. Kui väline teenus on valitud, on järgmine samm valida tehnoloogiad, mille abil ChatGPT olemasoleva Motichecki rakendusega integreerida.

Selleks on kaks peamist varianti:

1. Kirjutada ChatGPT integratsioon olemasoleva monoliitse rakenduse koodi sisse.
2. Luua ChatGPT integratsiooni jaoks eraldiseisev mikroteenus.

Esimese variandi puhul oleks programmeerimiskeel ette määratud olemasoleva rakenduse koodiga ja selleks oleks Kotlin. Kuigi autoril puudus eelnev kogemus Kotlini programmeerimiskeelega, ei olnud see suureks takistuseks, kuna Kotlin on sarnane Java keelega, millega autoril on olnud kõige pikaajalisem kokkupuude. Seetõttu ei lisandu

selle variandi jaoks suurt ajakulu uue keele õppimiseks, eriti arvestades asjaolu, et tegemist on juba olemasoleva projekti muutmisega, mitte uue kirjutamisega nullist. See tähendab, et saab vaadata, kuidas olemasolevas projektis on asju lahendatud, ning õppimine toimub kiiremini.

Teise variandi puhul saaks programmeerimiskeele valida ise. Ent asjaolu, et olemasolev koodibaas on juba Kotlinis, oleks sellegipoolest üks argument, et kasutada ka uues mikroteenuses Kotlinit. Kui kogu koodibaas on samas keeles, on seda lihtsam hallata, taaskasutada ja edasi arendada. Samas oleks see siiski vaid üks argument mitmete seas. Eraldi mikroteenuse kasutuselevõtt oleks ajamahukam, kuna uue mikroteenuse tegemine on aeglasem kui olemasoleva koodi täiendamine. Mitmed funktsionaalsused, nt turvalisus ja autentimine, tuleks duplitseerida. Positiivse poole pealt aitaks eraldi mikroteenuse tegemine ChatGPT-ga seotud loogikat rohkem isoleerida.

Kaaludes poolt- ja vastuargumente, otsustas autor siiski esimese variandi kasuks: täiendada juba olemasoleva rakenduse koodi ja mitte luua eraldi mikroteenust.

4 Töö käik

Järgnevalt kirjeldab autor töö käiku. Kuidas toimus vajaliku funktsionaalsuse arendus, millised takistused esinesid, ja kuidas need lahendati. Selles peatükis on esitatud töö kokkuvõtlikum kirjeldus. Üksikasjalikuma kirjeldusega on võimalik tutvuda lisades (vt lisa 3). Samuti on lisades võimalik tutvuda üksikasjalikumate kasutajalugudega (vt lisa 4), kus on ära toodud näiteks mõned tehnilised detailid.

4.1 Töö teostamine

Töö algas esimese koosolekuga Motichecki esindajaga, kes selgitas ettevõtte vajadust lähemalt. Tema soov oli luua Motichecki kasutavatele juhtidele rohkem väärtust, pakkudes töötajate tagasiside põhjal soovitusi olukorra parandamiseks. Tundus, et selleks võiks kõige paremini sobida ChatGPT, mis on tegelikult sisuliselt ka ainus piisavalt võimas avalikult kättesaadav tööriist sellise ülesande lahendamiseks.

4.1.1 Eeltöö

Enne järgmist koosolekut tegi autor ChatGPT-ga mõned katsetused töötajate tagasiside teemadel (vt lisa 2). Selleks uuris autor Motichecki testkeskkonda, kus oli üks testandmetega ettevõtte, millele oli testandmetena lisatud ka mitmeid töötajate tagasiside vastuseid. Testandmete põhjal proovis autor välja mõelda mõned küsimused, mis võiksid tekkida töötajate tagasiside statistikat uurival ettevõtte juhil, ning katsetas nende esitamist ChatGPT-le. Saadud tulemused koguti Google Sheets keskkonda, et neid saaks hiljem analüüsida ja töötada välja küsimused, millele ChatGPT annaks kõige kasulikumad vastused. Kuna ChatGPT puhul sõltub palju just sellest, kuidas küsimus sõnastada.

Samuti uuris autor, millised on tehnilised võimalused sellega liidestuda. Tuli välja, et ChatGPT-l ei ole veel ametlikku API't või muud masinliidest, mille abil seda teistesse

rakendustesse lihtsasti integreerida saaks. ChatGPT välja töötanud ettevõtte OpenAI veebilehel oli kirjas, et varsti on kavas selline API turule tuua.

Siiski oli API kaudu saadaval GPT-3 teenus, mis kasutab DaVinci keelemudelit ning toimib üsna sarnaselt ChatGPT-ga. Erinevus seisneb aga selles, et ChatGPT on optimeeritud just interaktiivse vestluse jaoks. GPT3 sobib aga pigem ühekordsete küsimuste esitamiseks. Sellegipoolest oli selle API olemasolu positiivne avastus, kuna tekkis lootus, et kuni ChatGPT avaliku API väljatulekuni saab kasutada GPT3 APIt.

Ühel järgneval koosolekul arutas autor Motichecki esindajaga juba lähemalt, kuidas probleemi lahendada. Vaadati välja koht kasutajaliideses, kuhu võiks tekitada nupu, mis käivitaks vestlusakna GPT3-ga. Samuti valmisid koosoleku tulemusena esimesed kasutajalood.

4.1.2 Algse prototüübi loomine

Järgmise sammuna toimus algse prototüübi loomine. Selle käigus lisas autor nimetatud küsimuse kaardile nupu, mille abil saab tehisintellekti teenuselt nõu küsida.

GPT-3 poolt genereeritud vastuse kuvamiseks hüpinkaknas kasutas autor sweetalert2 teeki. See võimaldab väga lihtsa vaevaga kuvada suhteliselt hea visuaaliga hüpinkaknaid.

Seejärel toimus järjekordne koosolek Motichecki esindajaga, kus autor näitas GPT-3 nupu prototüüpi. Katsetamise käigus ilmnis, et GPT-3 poolt genereeritud vastused olid palju lakoonilisemad, kui ChatGPT omad. Lisaks tundus, et eestikeelsete vastuste kvaliteet on madalam kui ingliskeelsete puhul. Üldiselt jäi Motichecki esindaja prototüübiga rahule.

4.1.3 Ettevalmistused testkeskkonda paigaldamiseks

Kui algne prototüüp oli valmis, siis järgmiseks oli vaja teha 2 sammu: kirjutada prototüübi kood ümber, järgides parimaid praktikaid; seejärel paigaldada koodimuudatused testkeskkonda, et ChatGPT nupu funktsionaalsust saaks testida rohkem inimesi.

Testkeskkonda paigaldamisega seoses tuli läbi mõelda järgmine küsimus. Eesmärk oli paigaldada uus funktsionaalsus testkeskkonda, aga samas tagada, et see kogemata toodangusse ei jõuaks. Selleks oli kaks võimalust:

1. Luua versioonihaldustarkvaras eraldi haru, kus hoida uut funktsionaalsust, ning paigaldada testkeskkonda just see haru, luues sisuliselt uue testkeskkonna.
2. Piirata ChatGPT nupu nähtavust rakenduse siseselt, et sellele pääseksid ligi näiteks ainult konkreetse ettevõtte esindajad või konkreetsed kasutajad. Tegemist oleks nn *feature flagiga*, mis võimaldab töötavas rakenduses teatava konfiguratsiooni abil lülitada sisse või välja teatavat funktsionaalsust.[5]

Valitud sai teine variant, kuna see on antud olukorras mõistlikuma ja lihtsam, ning võimaldab kiiremini testimise asuda.

Sel ajal tuli välja ka ChatGPT API, mida varem ei olnud. Seetõttu sai autor asendada senise GPT-3 text-davinci mudeli põhise päringu ChatGPT päringuga. Päringut tuli selleks muuta minimaalselt, kuna liides oli suhteliselt sarnane. ChatGPT API-le tuli sisendiks anda mitte lihtsalt üks sõnum, vaid sõnumite järjend, mis sisaldab käesoleva vestluse ajalugu. Ent kuna selles faasis oli iga küsimus ja vastus sisuliselt eraldi vestlus, siis suuri muudatusi teha ei tulnud. Juba esimesed katsetused näitasid, et ChatGPT vastus on pikem, põhjalikum, sisukam ja sujuvama keelekasutusega.

Tarkvaraarenduses on väga levinud põhimõte, kus loogika ja kasutajaliides hoitakse lahus. Laiemal tasandil väljendub see põhimõte juba selles, kui esirakendus ja tagarakendus on eraldi. Sellisel juhul peaks loogika asuma valdavalt tagarakenduses. Kuid sama põhimõte võib olla kasutusel ka esirakenduse sees, kuna enamasti ei õnnestu esirakendusest loogikat täielikult ära kaotada. Üldjuhul on ikkagi vaja teha kas või päringuid tagarakendusse, mingeid andmeid uuendada jne. Sellega kaasneb teatav kogus loogikat, mis asub esirakenduses.

Sellise lahenduse eesmärk on ühtlasi üritada ChatGPT funktsionaalsust ülejäänud rakendusest võimalikult isoleerida. Isoleeritus on kasulik kahel põhjusel. Esiteks on sellisel juhul lihtsam ChatGPT funktsionaalsust hiljem vajadusel eemaldada, näiteks kui tuleb välja, et selle vastu pole klientidel siiski loodetud huvi. Teiseks aitab koodi

isoleerimine vähendada arenduse käigus tekkida võivaid koodikonflikte, kui mitu arendajat muudavad korraga sama faili, ning seeläbi arendusprotsessi kiirendada ja tõhustada.

Selleks, et funktsionaalsust saaks kasutajapõhiselt sisse ja välja lülitada, tuli täiendada andmebaasi mudelit ning lisada kasutaja seadete tabelisse uus veerg, mis näitab, kas selle kasutaja jaoks on ChatGPT sisse lülitatud või mitte.

Mentor on kõnealuse tehisintellekti funktsionaalsuse äriline nimetus, mida kasutatakse eka koodis. Kui Motichecki kasutaja tahab saada küsimuse kohta nõu, siis ta pöördub (robot)mentori poole, kelle rolli täidab tegelikult ChatGPT.

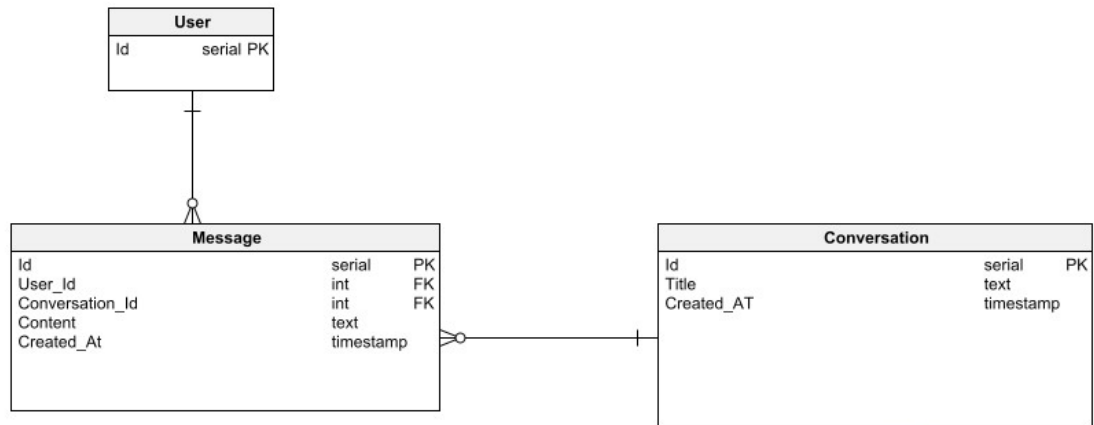
Kui esmane mentori funktsionaalsus oli valmis, paigaldati see testkeskkonda, et koguda tagasisidet äripoole testijatelt. Testimise käigus laekus kasulikke tagasisidet. See puudutas nii kasutajaliidese poolt kui ka ChatGPT poolt genereeritud vastuste sisulist poolt.

Järgmise sammuna lisas autor funktsionaalsuse, et ChatGPT suudaks vastata erinevates keeltes sõltuvalt sellest, mis keel on kasutajal parasjagu valitud. See oli suhteliselt lihtne muudatus. Tuli lihtsalt võtta kasutusele olemasolev tõlgete süsteem ning tõlkida ChatGPT-le esitatav küsimus enne teele saatmist selle süsteemi abil. See tõlkimine toimub esirakenduses. Tagarakenduses aga midagi muutma ei pidanud, kuna see võtab parameetrina sisse lihtsalt sõne.

4.1.4 Mentori sõnumite ajalugu

Järgmiseks oli tarvis lisada funktsionaalsus, et kasutaja saaks kuskil näha ajalugu küsimustest, mida ta on mentori käest küsinud, ning vastuseid, mida mentor on andnud. Selle jaoks tuli otsustada, millisel kujul neid andmeid andmebaasis hoida. Kaalumisel oli kaks alternatiivi.

Esimene kaalumisel olnud alternatiiv on näha joonisel 1.



Joonis 1. Mentori vestluste ajaloo andmebaasiskeem. Variant 1.

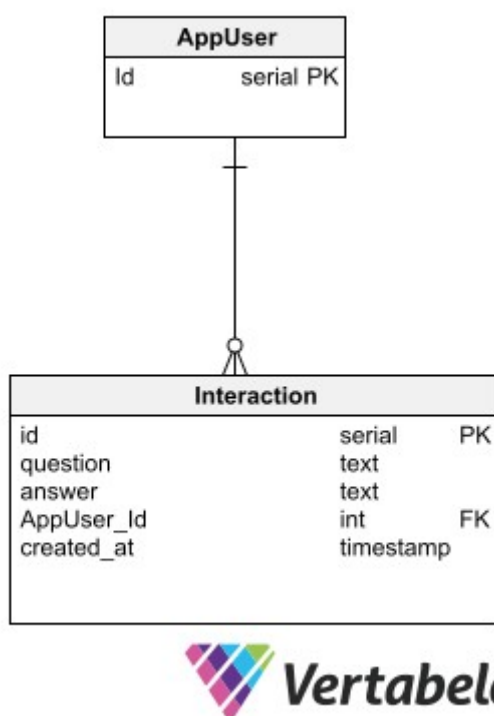
Selle variandi puhul kasutatakse küllaltki klassikalist sõnumite andmemudelit, kus sõnumid on jaotatud vestlustesse. Iga vestlus võib sisaldada mitu sõnumit, mille on saatnud erinevad kasutajad. Sellise andmemudeli puhul on tõenäoliselt vaja täiendusi jõudluse optimeerimiseks. Kui vaadelda näiteks stsenaariumit, kus kasutaja soovib näha kõiki vestlusi, mille osaline ta on, siis tuleks sellise andmemudeli puhul vaadata läbi kõikide andmebaasis olevate vestluste kõik sõnumid, et leida vestlused, kuhu on saatnud sõnumeid ka konkreetne kasutaja. Kuna vestluseid ja sõnumeid sisaldavas rakenduses võib sõnumeid olla äärmiselt palju, oleks selline lahendus suure kasutajate hulga korral praktikas teostamatu. Seega oleks vaja päringuid optimeerida. Selleks on mitmeid võimalusi.

Üks lihtsamaid oleks hoida iga vestluse juures eraldi andmeid selle vestluse osaliste kohta. Selleks oleks vaja vahetabelit *Conversation_participant*, mis seoks omavahel vestlused ja kasutajad. Rakenduse loogika kihis saaks sinna tabelisse lisada kande iga kord, kui vestlusesse postitab mõni kasutaja, kes sellesse vestlusesse varem veel ei postitanud. Sellisel juhul piisaks kasutaja vestluste ülesleidmiseks vaadata läbi kõik vestlused ja nende vestluste osalejad. Kui lisada ka vastavad indeksid, ei tohiks jõudlusega probleeme olla.

Teine võimalus oleks sisuliselt kasutada vahemälu tehnikat, ehk arvutada kasutaja vestluste info ette ära ning kasutada väljaarvutatud väärtust, mitte iga kord eraldi arvutada. Selleks on omakorda mitmeid võimalusi, näiteks andmebaasides vaadete funktsioon.

Kokkuvõtteks võib esimese andmemudeli kohta öelda, et see on küllaltki universaalne ning võimaldab tulevikus lisada rakendusse mitmeid uusi funktsionaalsusi, näiteks kas või kasutajate omavaheliste sõnumite funktsionaalsust. Samas vajab see jõudlusele mõtlemist ja optimeerimist, tõenäoliselt juba praeguse kasutajate hulga juures.

Teine alternatiiv on näha joonisel 2.



Joonis 2. Mentori vestluste ajaloo andmebaasiskeem. Variant 2.

See andmemudel on esimesega võrreldes lihtsam ning kohandatud konkreetse kasutajaloo tarbeks. Selle mudeli kohaselt ei salvestata kasutaja ja ChatGPT vahelisi interaktsioone mitte vestluste formaadis, vaid ühekordsete interaktsioonidena. Üks interaktsioon koosneb kasutaja küsimusest ning ChatGPT poolt genereeritud vastusest.

Üks erinevus esimese andmemudeliga on näiteks see, et kui tulevikus peaks tekkima kasutajal võimalus esitada ChatGPT-le täpsustavaid küsimusi, siis ei oleks need küsimused grupeeritud algse küsimuse ja vastusega ühtsesse olemisse ehk vestlusesse. Sellega läheks kaduma osa andmete struktuurist. Samas oleks muidugi võimalik vestluste olem hiljem juurde tekitada, kui selline vajadus peaks tekkima. Ent sellegipoolest piirab vaatlusalune andmemudel vestluste struktuuri, kuna igale küsimusele peab järgnema vastus. Seega ei toeta selline andmemudel näiteks vestlusi, kus üks kasutaja kirjutab mitu sõnumit järjest.

Pärast mõlema alternatiivi analüüsimist otsustas autor valida teise, lihtsama alternatiivi. Esimene mudel on küll universaalsem ja võimaldab tulevikus rakendusele rohkem funktsionaalsuseid lisada, kuid selle teostamiseks kuluks rohkem aega. Antud juhul on aga oluline esimese prototüübiga võimalikult kiirelt valmis saada, et saaks seda testima hakata, kuna tegemist on iduettevõttega ja uue funktsionaalsusega, millele on oluline saada klientidelt ehk reaalselt kasutajatelt võimalikult kiiresti tagasisidet.

4.1.5 Testimine ja tagasiside

Pärast ajaloo funktsionaalsuse lisamist oli mentori featuur piisav, et alustada aktiivsemat ettevõttesisest testimist. Selle käigus tuli väärtuslikku tagasisidet ChatGPT poolt genereeritud tekstide, näiteks nende pikkuse ja tooni kohta. Tagasiside põhjal toimus ChatGPT-le saadetava prompti seadistamine. Näiteks sai lisatud funktsionaalsus, kus üle 80% positiivsete vastustega küsimuse puhul antakse ChatGPT-le kaasa ka juhised olla oma vastuses positiivsem ja kiitvam, kuna vastasel juhul tundus testijatele, et ChatGPT ei erista piisavalt erineva protsentuaalse tulemusega küsitlusi, vaid annab sageli üsna standardse vastuse, kus käsitleb peamiselt töötajatelt küsitud küsimuse teemat, mitte niivõrd konkreetset protsentuaalset tulemust.

Sarnaseid peenhäälestusi oli teisigi, et tagada ChatGPT vastuste parem vastavus ootustega.

4.1.6 Interaktiivse vestluse funktsionaalsuse lisamine

Järgmise sammuna tuli lisada funktsionaalsus, kus kasutaja saab pidada ChatGPT-ga interaktiivset vestlust, et esitada vajadusel täpsustavaid küsimusi. Kogu seni tehtud

funktsionaalsus pidi jääma samaks: kui kasutaja vajutab küsimuse kaardil vastavale nupule, suunatakse ta mentori vestluse vaatesse, kus ChatGPT genereerib ülevaate ja soovitused küsimuse kaardil olnud küsimuse kohta. Ent kui varem sellega tegevus piirduski, siis nüüd pidi kasutajal olema võimalus esitada vaba tekstina täiendavaid küsimusi ning pidada ChatGPT-ga interaktiivset vestlust.

Selleks tuli teha täiendusi ni eesirakenduses kui ka tagarakenduses. Tagarakenduses tuli lisada API *endpoint*, mille abil saab esitada ChatGPT-le vabatekstilise küsimuse. Seejuures oli tarvis kuidagi defineerida vestluse skoop, ehk see, millised sõnumid OpenAI API-le päringuga kaasa pannakse. Nimelt saab OpenAI ChatGPT API-le päringut saates lisaks küsimusele endale panna kaasa ka eelneva küsimuste-vastuste ajaloo. See võimaldab ChatGPT-l võtta arvesse senise vestluse ajalugu ning vajadusel varasematele sõnumitele tagasi viidata. Küsimus seisnes selles, kuidas otsustada, millised sõnumid OpenAI API-le täpselt kaasa panna. Naiivne lahendus oleks näiteks alati panna kaasa kogu sõnumite ajalugu, mida konkreetne kasutaja on ChatGPT-ga pidanud, ent see lahendus põrkuks kokku praktilise takistusega, kuna ChatGPT-le saadetava päringu maht on piiratud. Seega sai otsustatud, et rakenduse jaoks algab eraldiseisev vestlus iga kord, kui kasutaja vajutab küsimuse kaardil nupule ja küsib seeläbi nõu konkreetse teema kohta. Kõik järgnevad täpsustavad küsimused lähevad sama vestluse alla, ning vestlus lõpeb siis, kui kasutaja peaks uuesti mõnel küsimuse kaardil mentori nuppu vajutama. Kui aga kasutaja esitab nii palju täpsustavaid küsimusi, et kõik sõnumid ei mahu enam ChatGPT API päringusse ära, siis pannakse kaasa nii palju sõnumeid, kui mahub, alustades ajaliselts kõige värskematest sõnumitest.

Selle loogika realiseerimiseks sai lisatud *interaction* tabelile uus veerg *conversation_id*, mis grupeerib interaktsioonid loogilisteks vestlusteks. Tagarakendus saadab esirakendusele iga interaktsiooni kohta kaasa ka selle *conversation_id*, ning kui kasutaja esitab täpsustava küsimuse, paneb esirakendus tagarakendusele samuti kaasa käimasoleva vestluse *conversation_id*, et tagarakendus saaks enne OpenAI API-le päringu tegemist võtta andmebaasist kõik eelnevad selle vestluse sõnumid ning need päringuga kaasa panna.

Samuti vajas kõnealuse funktsionaalsuse realiseerimise käigus lahendamist küsimus, kuidas takistada vestluseid täiesti kõrvalistel teemadel. Kuna Motichack on töötajate

rahulolu monitoorimiseks mõeldud rakendus, siis peaksid kasutajate poolt ChatGPT-ga peetud vestlused olema seotud selle teemaga. Täiesti kõrvalistel teemadel peetavad vestlused oleksid ebasoovitavad, kuna need juhiksid kasutajat eemale põhitegevusest, mida ta rakendusse tegema tuli. Samuti võib tõstatuda küsimus, kes vastutab ChatGPT vastustest tulenevate võimalike kahjulike tagajärgede eest. Näiteks kui kasutaja peaks küsima ChatGPT käest näiteks meditsiinilist nõu, ning kui saadud nõu järgimine peaks näiteks tema tervist hoopis kahjustama. Seetõttu oli oluline lisada funktsionaalsus, et kui kasutaja kaldub vestluse teemast kõrvale, siis ChatGPT paluks tal viisakalt teemas püsida.

Küsimuste teema piiramiseks kaalus autor kahte peamist võimalust. Üks võimalus oleks teha enne seda, kui rakendus päringu OpenAI API-le saadab, veel üks täiendav päring, ning paluda selles päringus ChatGPT-l tuvastada, kas kasutaja küsimus on seotud töötajate ja töökeskkonna teemadega või mitte. Kui nimetatud kaks päringut teha üksteise järel, siis pikendaks see oluliselt kasutaja jaoks vastuse saamise aega, mis oli juba niigi suhteliselt pikk (tavaliselt 10-60 sekundit). Võimalik olnuks teha optimeering, kus mõlemad päringud saadetakse teele samaaegselt erinevates lõimedes, kuid kui temakontrolli päringu vastus peaks olema negatiivne, siis teises päringus saadud vastust kasutajale ei tagastata, vaid selle asemel tagastatakse teade teemavälise küsimuse kohta.

Teine võimalus oli see, et tehakse küll üks päring nagu ennegi, kuid Motichecki rakendus lisab kasutaja küsimusele omalt poolt prompti sisse ChatGPT jaoks täiendava juhtnööri, kus on öeldud, et kui järgnev küsimus ei puuduta tööelu ja töökeskkonda, siis palutakse ChatGPT-l vastamisest viisakalt keelduda ning paluda kasutajal püsida teemas.

Autor testis teist lahendust ning kuna see toimis, otsustas valida just selle, kuna see on lihtsam ning vajab vähem arendamist.

4.2 Töö tulemused

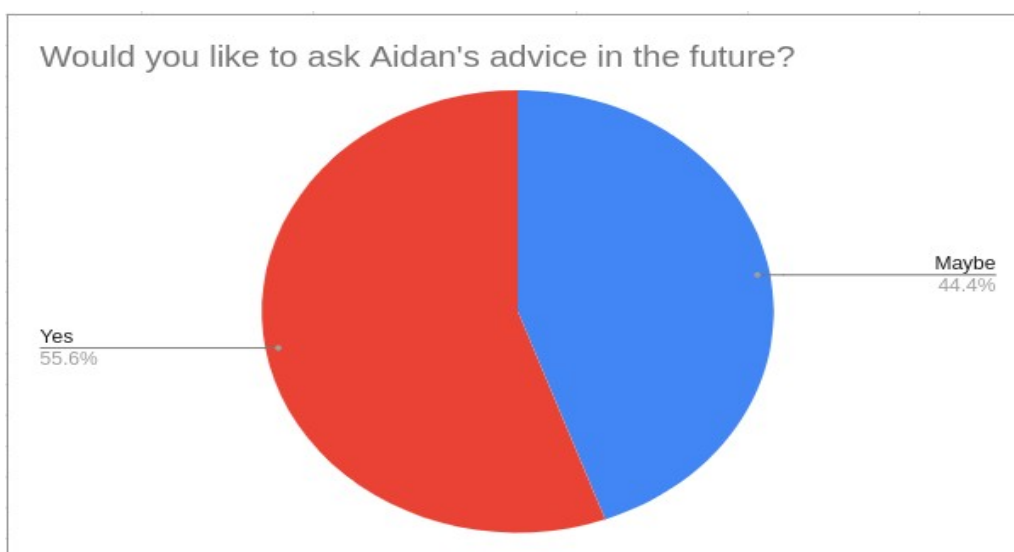
Töö tulemusena sai Motichecki rakendusele lisatud uus mentori funktsionaalsus. See tähendab, et Motichecki kasutajad saavad pöörduda tehisintellektil põhineva ChatGPT

nn mentori poole, kes aitab neil töötajate tagasiside andmeid tõlgendada ning anda soovitusi, kuidas olukorda parandada.

Motichecki kasutajad saavad küsida nõu konkreetse töötajate tagasiside kohta. Lisaks saavad kasutajad alustada mentoriga interaktiivset vestlust, et küsida täpsustavaid küsimusi. Mentoriga toimunud interaktsioone saavad kasutajad hiljem näha eraldi vaates, kus näeb kõiki salvestatud vestluseid.

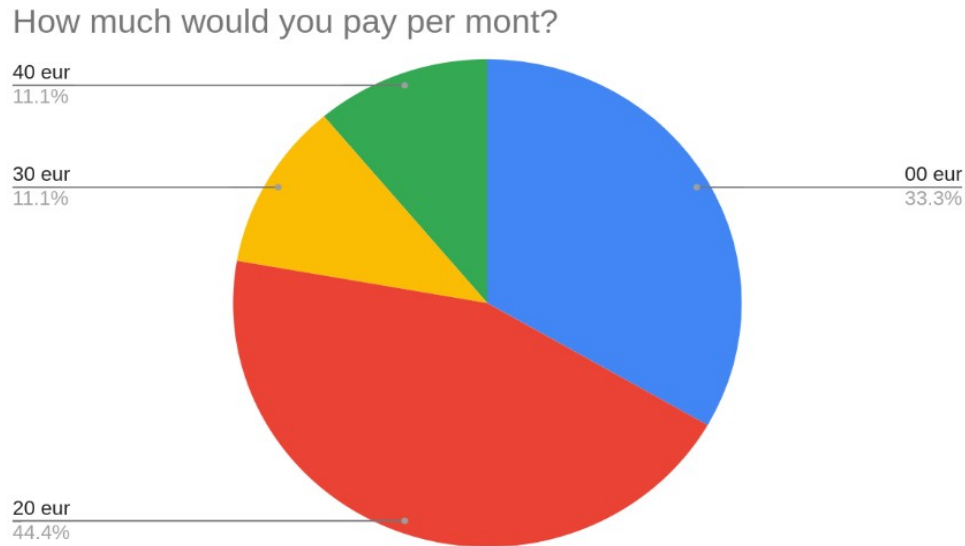
Teostatud töö tulemusena arenes Motichecki rakendus järgmisele tasemele ning pakub selle kasutajatele rohkem väärtust. Kasutajad saavad töötajate tagasiside andmeid enda jaoks paremini lahti mõtestada ning samuti soovitusi, kuidas olukorda parandada. See võimaldab efektiivsemalt tõsta töötajate rahulolu ja luua paremat töökeskkonda. Motichecki kasutavate ettevõtete jaoks väljendub see loodetavasti lõpuks paremates majandustulemustes, kuna motiveeritumad töötajad on produktiivsemad ja aitavad ettevõtte eesmärke paremini saavutada. Motichecki enda jaoks aga võimaldab see küsida klientidelt mentori teenuse kasutamise eest täiendavat tasu ning muuta Motichecki potentsiaalsete uute klientide jaoks atraktiivsemaks lahenduseks.

Pärast arenduse valmimist anti uut funktsionaalsust testimiseks kasutada kümnele Motichecki kliendile, kes andsid selle kohta ka tagasisidet. Nagu näha jooniselt nr 3, siis sooviksid 55,6% vastanutest mentori funktsionaalsust ka tulevikus kasutada. 44,4% sooviksid seda võibolla tulevikus uuesti kasutada.



Joonis 3. Klientide tagasiside mentori kasutamise kohta tulevikus.

78% pilootklientidest vastas, et mentori funktsionaalsusest oli neile kasu. Nagu näha jooniselt 4, oleks kaks kolmandikku klientidest valmis maksma selle funktsionaalsuse eest 20 eurot kuus või rohkem.



Joonis 4. Klientide valmisolek maksta mentori funktsionaalsuse eest.

Lisaks võiks klientide tagasiside poole pealt välja tuua, et 33% klientidest hindas mentori antud soovitusid liialt üldiseks. Seega tasuks edaspidi uurida, kas oleks võimalik muuta genereeritavaid vastuseid veelgi konkreetsemaks.

Samuti oleks täiendava arendusena võimalik lisada rakendusele mentori funktsionaalsuse hinnastamine ja sellega seotud arvepidamine.

5 Kokkuvõte

Käesolevas lõputöös uuris autor, kuidas genereerida töötajate tagasiside andmete põhjal inimloetavaid aruandeid, mis aitaksid tagasisidet tõlgendada ning anda soovitusi olukorra parandamiseks. Eesti iduettevõtte Motichecki jaoks teostas autor vastava analüüsi ning arenduse, luues integratsiooni ChatGPT juturoboti veebiteenusega. Pärast arendust viidi läbi testimine Motichecki klientidega. Suurem osa klientidest pidas uut funktsionaalsust kasulikuks ning väljendas valmisolekut selle eest igakuiselt maksta.

Muuhulgas analüüsis autor konkreetset ärilist vajadust, mida tuleb lahendada, pani paika loodava arenduse nõuded ja koostas kasutajalood, kirjeldas erinevaid lahendusvariante paika pandud tarkvaranõuete saavutamiseks ning analüüsis erinevaid tehisintellekti teenuseid, mis võiksid aidata probleemi lahendada. Samuti esitas autor detailsema ülevaate ChatGPT kohta ning analüüsis eraldi alapeatükis ka selle kasutamisega seotud riske. Lisaks analüüsile esitas autor ülevaate tehtud arenduse käigust, arenduse käigus tekkinud küsimustest ning valitud lahendustest.

Kuna suured keelemudelid, mille hulka kuulub ka ChatGPT, on kiirelt arenev ja võimas tehnoloogia, siis võib oletada, et tulevikus hakatakse seda integreerima üha rohkematesse rakendustesse. Antud lõputöö pinnalt võiks tuua esile, et taoliste integratsioonide puhul oleks oluline eelnevalt läbi kaaluda võimalik kasu ja kahju, ning teha otsus sellest lähtuvalt, arvestades rakenduse spetsiifikat. Kuna selliste keelemudelite puhul jääb alati risk, et nad genereerivad tõe mittevastavat sisu, siis tuleb analüüsida võimalikke tagajärgi, mis sellisel juhul võiksid tekkida. Seetõttu ei sobi need tõenäoliselt kriitilise tähtsusega infosüsteemidesse, küll aga võib neist kasu olla väiksema riskiga infosüsteemides.

ChatGPT puhul on oluline, milline täpselt on küsimuse sõnastus, kuna sellest võib sõltuda vastuse kvaliteet. Lõputööst nähtub, et ettevõtted saavad ChatGPT-d oma rakendusse integreerides luua lõppkasutajale täiendavat väärtust, kui mõtleavad hoolikalt läbi, kuidas koostada prompt, mis tagaks kasutaja jaoks võimalikult kasuliku vastuse.

Samuti tasub kaaluda võimalust panna prompti kaasa täiendavaid asjakohaseid andmeid, mis on vastavas infosüsteemis olemas, et muuta ChatGPT vastus veelgi informatiivsemaks ja kasulikumaks.

Lõputöö võiks huvi pakkuda kõigile, kes soovivad integreerida ChatGPT mõnda rakendusse. Loodetavasti saab sellest lõputööst selliseks juhuks kasulikke mõtteid.

Kasutatud kirjandus

- [1] Jon Martindale. (2023). How to use Google Bard, the latest AI chatbot service [Online, kasutatud 21. märtsil 2023]. <https://www.digitaltrends.com/computing/how-to-use-google-bard/>
- [2] Billy Perrigo. (2023). The New AI-Powered Bing Is Threatening Users. That's No Laughing Matter [Online, kasutatud 21. märtsil 2023]. <https://time.com/6256529/bing-openai-chatgpt-danger-alignment/>
- [3] Floridi, L. (2023). AI as Agency without Intelligence: On ChatGPT, large language models, and other generative models. *Philosophy and Technology*.
- [4] Eleri Aedmaa, Helen Kaljumäe. (2023). Skandaalne ChatGPT – kuidas panna see eesti keele heaks tööle? [Online, kasutatud 21. märtsil 2023]. <https://novaator.err.ee/1608889298/skandaalne-chatgpt-kuidas-panna-see-eesti-keele-heaks-toole>
- [5] Meinicke, J., Wong, C. P., Vasilescu, B., & Kästner, C. (2020, June). Exploring differences and commonalities between feature flags and configuration options. In Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering in Practice (pp. 233-242).
- [6] Brooks, F., & Kugler, H. (1987). No silver bullet (pp. 1-14). April.
- [7] Cohn, M. (2004). User stories applied: For agile software development. Addison-Wesley Professional. Lk 3.
- [8] Cohn, M. (2004). User stories applied: For agile software development. Addison-Wesley Professional. Lk 4.
- [9] Liebrez, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *The Lancet Digital Health*, 5(3), e105-e106.
- [10] Toby Walsh. (2022). Everyone's having a field day with ChatGPT – but nobody knows how it actually works [Online, kasutatud 21. märtsil 2023]. <https://theconversation.com/everyones-having-a-field-day-with-chatgpt-but-nobody-knows-how-it-actually-works-196378>
- [11] Manning, Christopher D. (2022). "Human Language Understanding & Reasoning". Daedalus.
- [12] OpenAI. (2023). "GPT-4 System Card".

Lisa 1– Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina, Martin Garbuz

- 1 Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Inimloetavate aruannete genereerimine töötajate tagasiside andmete põhjal Motichecki näitel” mille juhendaja on Toomas Lepikult
 - 1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2 üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
- 2 Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
- 3 Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

24.04.2023

1 Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

Lisa 2 – Katsetused ChatGPT-ga

Järgnevalt on välja toodud mõned katsetused ChatGPT-ga. Sama taustavärviga küsimused tähistavad loogiliselt sama tervikvestlust. Ilma taustavärvita on eraldiseisvad küsimused.

Järgnevad katsetused on tehtud 16. veebruaril 2023.

Tabel 2. Katsetused ChatGPT-ga.

Küsimus (prompt)	Vastus	Hinnang (kasulik või mitte?)
Vestlus: töötaja on väsinud	<p>Töötajate tagasiside kuulamine on oluline samm nende rahulolu ja heaolu tagamisel. Kui töötajad tunnevad end vaimselt ebamugavalt, võib see mõjutada nende tervist ja jõudlust, mis omakorda võib mõjutada teie ettevõtte tulemuslikkust. Seetõttu on oluline, et tegeleksite selle probleemiga õigeaegselt ja tõhusalt.</p> <p>Siin on mõned sammud, mida saate kaaluda, et lahendada töötaja esitatud probleemi:</p> <p>Vestelge töötajaga: Vestelge töötajaga avatult ja kuulake tema muresid ja kahtlusi. Küsige temalt üksikasjalikult, miks ta tunneb end nii väsinuna ja mis teda tööl kõige rohkem häirib. Kuulates tema vastuseid, saate paremini mõista tema murede olemust ja leida sobivad lahendused.</p> <p>Vaadake üle töötingimused: Vaadake üle töötingimused, et veenduda, et need vastavad töötervishoiu- ja ohutusnõuetele ning et töökeskkond on optimaalne. Võib-olla saab tööprotsesse optimeerida, et vähendada töökoormust ja tagada töö ja puhkuse õige tasakaal.</p> <p>Pöörduge spetsialisti poole: Võib-olla võiksite töötajale soovitada kohtumist psühholoogi või terapeutiga, kui ta tunneb, et tema vaimsed probleemid mõjutavad tema elukvaliteeti. Psühholoog või terapeut saab pakkuda vajalikku tuge ja nõuandeid, et töötaja saaks oma muredega toime tulla.</p> <p>Looge positiivne töökultuur: Loo positiivne töökultuur, kus töötajad tunnevad end toetatuna ja väärtustatuna. See võib hõlmata motivatsioonipakette, tunnustusprogramme, võimalusi töökoha arendamiseks ja karjääriredelil</p>	

tõusmiseks.

Korraldage uuesti küsitlus: Töötajate uuesti küsitlus, et hinnata, kas hiljutised muudatused on aidanud parandada nende tööolusid ja vaimset heaolu.

On oluline, et tegeleksite töötaja esitatud probleemiga võimalikult kiiresti, et vältida võimalikku negatiivset mõju töötaja tervisele ja ettevõtte tulemustele. Kui te pole kindel, kuidas lahendada probleemi, võiksite kaaluda professionaalse konsultandi või spetsialisti poole pöördumist, et saada nõu ja soovitusi.

Kui töötaja ei oska öelda, miks ta tunneb end nii väsinuna, võib selle põhjus olla midagi muud kui tööga seotud probleemid. Võib-olla on tal mingisugune terviseprobleem, mis vajab tähelepanu, või ta võib kogeda muid isiklikke probleeme, mis mõjutavad tema vaimset ja füüsilist heaolu.

Siiski on oluline, et te ei jäta töötaja esitatud probleemi tähelepanuta. Võite kaaluda muid võimalusi, kuidas saate teda toetada, näiteks:

Pakkuge töötajale lühikest puhkust: Võib-olla võiks töötaja võtta paar päeva puhkust, et ta saaks taastuda ja värskenduda.

Pöörduge arsti poole: Võite soovitada töötajal pöörduda arsti poole, et ta saaks oma väsimuse põhjuseid selgitada ja vajadusel ravi alustada.

Ma rääkisin selle töötajaga ja küsisin, miks ta tunneb end nii väsinuna. Ta vastas, et ta ei tea, miks ta nii väsinud on.

Pakkuge vaimse tervise tuge: Võib-olla võiks töötaja saada tuge psühholoogi või terapeudi käest, kui ta kogeb vaimse tervise probleeme.

Töötaja paindlikkuse pakkumine: Kaaluge võimalust, et töötaja saaks teha paindlikumat tööaega või töötada kaugtöona, kui see aitaks tema töö- ja puhkeaja tasakaalu saavutamisele.

Töökoormuse vähendamine: Võib-olla saaksite töötaja töökoormust ajutiselt vähendada, kuni tema vaimne ja füüsiline seisund paraneb.

Kui töötaja endiselt ei suuda välja selgitada, miks ta tunneb end nii väsinuna, võite kaaluda ametliku aruande esitamist tema tervise seisundi kohta ja võimalusel soovitada, et ta saaks tervisekontrolli. Samuti võite kaaluda teiste töötajate üldise rahulolu ja heaolu uurimist, et näha, kas mõni teine töötaja kogeb sarnaseid probleeme.

Kui teised töötajad sarnaseid probleeme ei koge, võib see viidata, et probleem ei ole seotud töökeskkonnaga, vaid võib olla seotud individuaalsete faktoritega, mis mõjutavad töötaja heaolu. Sel juhul võite kaaluda järgmisi samme:

Teised töötajad sarnaseid probleeme ei koge.

Vestelda töötajaga üks-ühele: Kuna töötaja probleem ei pruugi olla seotud tööga, võite kaaluda individuaalset vestlust töötajaga, et uurida, kas tal on muid isiklikke probleeme, mis mõjutavad tema vaimset ja füüsilist heaolu.

Pakkuda lisatuge: Võite töötajale pakkuda lisatuge, näiteks psühholoogi või terapeudi konsultatsioone, kui ta kogeb vaimse tervise probleeme. Samuti võite pakkuda töötajale täiendavat puhkust või paindlikku tööaega, kui see aitaks tema töö- ja puhkeaja tasakaalu saavutamisele.

Motiveerida ja julgustada: Võite julgustada töötajat leidma motivatsiooni ja rõõmu oma tööst, näiteks pakkudes talle uusi väljakutseid või ülesandeid, mis vastavad tema oskustele ja huvidele.

Jälgida edasist olukorda: Oluline on jälgida töötaja olukorda ja teha regulaarselt kindlaks, kas tema seisund on paranenud. Kui probleem püsib, võite kaaluda individuaalsete meetmete kohandamist või muutmist vastavalt töötaja vajadustele.

Lõpuks on oluline mõista, et töötajate heaolu on oluline ja võib mõjutada nende tootlikkust ja pikaajalist panust ettevõttesse. Seetõttu on oluline võtta töötaja mured tõsiselt ning pakkuda talle vajalikku tuge ja abi. Muidugi, siin on näidiskiri, mida võite kasutada, et leppida kokku üks-ühele vestlus Peetriga:

Tere Peeter,

Oleme saanud tagasisidet meie töötajate küsitluse tulemuste kohta, kus mainisid, et tunned end tööl väga väsinuna. Tahaksin sinuga kohtuda, et arutada, kuidas saaksime aidata sind end paremini tunda ning leida lahendusi, mis võiksid su tööolukorda parandada.

Palun anna mulle teada, millal sulle sobib üks-ühele vestluseks aega leida. Kohtumine võib toimuda kas kontoris või virtuaalselt, vastavalt sinu eelistusele.

Ootan sinu vastust ja loodan, et saame peagi kokku saada, et arutada sinu heaolu töökohal.

Parimate soovidega,

[Teie nimi]

Hea küsimus! Läbipaistev kommunikatsioon on äärmiselt oluline igas organisatsioonis, eriti kui soovite tagada, et teie töötajad oleksid informeeritud ja motiveeritud.

Siin on mõned ideed, mida saate proovida:

Ma olen ettevõtte juht ja minu töötajad kurdavad, et töö juures ei ole piisavalt läbipaistev kommunikatsioon. Mida ma peaksin tegema?

Selgitage oma ootusi töötajatele: Selgitage selgelt välja, millist teavet peate jagama ja millal ning millist teavet ootate oma töötajatelt. Veenduge, et teie ootused oleksid kõigile töötajatele selged.

Korraldage regulaarseid koosolekuid: Korraldage regulaarseid koosolekuid, et jagada uudiseid, uuendusi ja ettevõtte arengut. See annab töötajatele võimaluse küsida küsimusi ja saada vajalikku tagasisidet.

Loo avatud ukse poliitika: Julgustage oma töötajaid pöörduma teie poole küsimuste ja muredega. Veenduge, et

teie töötajad tunneksid end vabalt rääkima ja jagama oma mõtteid.

Kasutage erinevaid kommunikatsioonivahendeid: Kasutage erinevaid vahendeid, näiteks e-kirju, siseportaale või sotsiaalmeediat, et jagada teavet ja hoida oma töötajaid kursis ettevõtte uudistega.

Anna tagasisidet: Andke oma töötajatele tagasisidet nende töö kohta ja veenduge, et nad teaksid, kuidas nad saavad oma tööd paremini teha.

Koolitage oma juhte: Koolitage oma juhte tõhusa kommunikatsiooni osas, et nad saaksid suhelda oma alluvatega selgelt ja tõhusalt.

Need on vaid mõned ideed, kuidas saate suurendada oma organisatsiooni läbipaistvat kommunikatsiooni. Loodan, et see aitab!

Kui kolm töötajat kümnest ütlevad, et nad pole oma töötasuga rahul, on see murettekitav signaal, et peate võib-olla vaatama üle oma palgapoliitika. Siin on mõned sammud, mida võiksite kaaluda:

Analüüsige oma palgapoliitikat: Vaadake üle oma ettevõtte palgapoliitika, et näha, kuidas see võrdub teiste ettevõtetega teie tööstusharus. Võite kaaluda ka töötajate palgaküsitluste korraldamist, et saada aimu sellest, milline on tööturul tavapärase palk teie ettevõtte asukoha ja töövaldkonna kohta.

Arutage töötajatega nende muret: Arutage töötajatega nende muret ja küsige neilt, mis on nende arvates õiglane palk, võrreldes nende töö panusega. Kuulake hoolikalt nende seisukohti ja püüdke leida võimalusi, kuidas saate nende muret leevendada.

Kolm töötajat kümnest ütlesid, et nad pole rahul oma töötasuga. Mida peaksin tegema?

Käivitage palgakorrektsioon: Kui teie ettevõtte suudab endale seda lubada, võite kaaluda töötasude korrigeerimist, et vastata töötajate murele. Selleks võite kaaluda palgakorrektsiooni kõigile töötajatele või suurendada palka ainult nendele töötajatele, kes seda kõige rohkem vajavad.

Pakkuge töötajatele muid hüvesid: Kui palgatõus pole võimalik, võite kaaluda töötajatele muude hüvede pakkumist, nagu boonused, paindlik tööaeg, lisapuhkusepäevad või muud soodustused. Need hüved võivad aidata kaasa töötajate rahulolule ja motivatsioonile.

Töötage välja pikemaajaline strateegia: Kui palgatõus või hüved ei ole kohe võimalikud, võite kaaluda pikemaajalise strateegia väljatöötamist, et tõsta ettevõtte kasumlikkust ja võimaldada tulevikus suuremaid investeeringuid töötajate palkadesse ja hüvedesse.

Oluline on meeles pidada, et töötajate rahulolu on võtmeks ettevõtte edule. Kui teie töötajad pole oma töötasuga rahul, võib see mõjutada nende motivatsiooni ja jõudlust, mis võib omakorda mõjutada ettevõtte kasumlikkust ja tulemusi.

Seetõttu on oluline töötada välja lahendused, mis aitavad parandada töötajate rahulolu ja luua positiivset tö

Me küsisime töötajatelt sellise küsimuse: "Kas sinu meeskonnas aktsepteeritakse juhtkonna otsuseid meelsasti?"

Kas sa saaksid teha mulle töötajate vastustest kokkuvõtte?

Vastused olid sellised:
16.02.2023
See nädal oli väga ok

19.01.2023
Thank you for this good program!

Küsimusele "Kas sinu meeskonnas aktsepteeritakse juhtkonna otsuseid meelsasti?" vastasid töötajad järgmiselt:

23.12.2022
100%! Hea on teha tööd, mille väärtus ja roll on hästi tuntav

3 töötajat vastasid positiivselt ja väljendasid, et meeskonnas aktsepteeritakse juhtkonna otsuseid meelsasti.
1 töötaja väljendas oma rahulolu tööga, kuid ei maininud otseselt küsimusele vastates midagi.

23.12.2022
100%! Hea on teha tööd, mille väärtus ja roll on hästi tuntav

1 töötaja vastas positiivselt, kuid väljendas samal ajal muret mustuse ja koristamise üle ühes konkreetse koha köögis.

21.12.2022
Thank you for this good program!

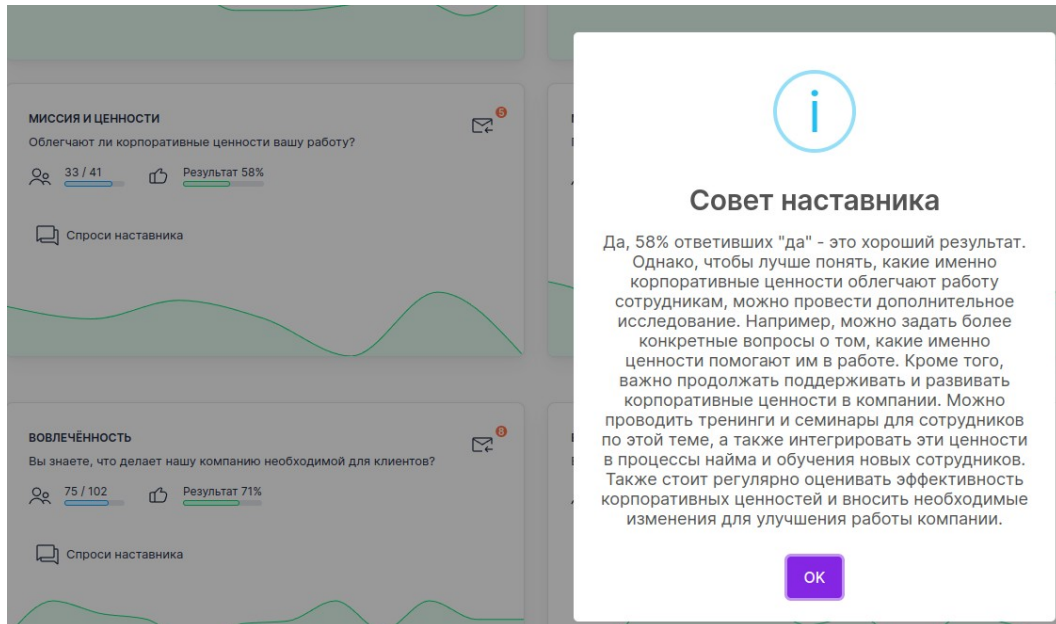
1 töötaja vastas vene keeles, väljendades rahulolu oma tööga.
2 töötajat väljendasid positiivsust ja tänulikkust programmi kohta, kuid ei maininud otseselt küsimusele vastates midagi.
Kokkuvõttes võib öelda, et enamuse töötajatest vastas küsimusele positiivselt, väljendades, et meeskonnas aktsepteeritakse juhtkonna otsuseid meelsasti. Siiski oli üks töötaja mures koristamise ja mustuse pärast ühes konkreetse koha köögis, mis võib vajada tähelepanu ja lahendamist.

09.12.2022
Я очень доволен своей работой

09.12.2022
Sellega on hästi, aga eile oli 3. korruse köök väga räpane. Palun võtke sellega seoses midagi ette. See ei ole normaalne, et meil pidevalt on köögis asjad igal pool laial..

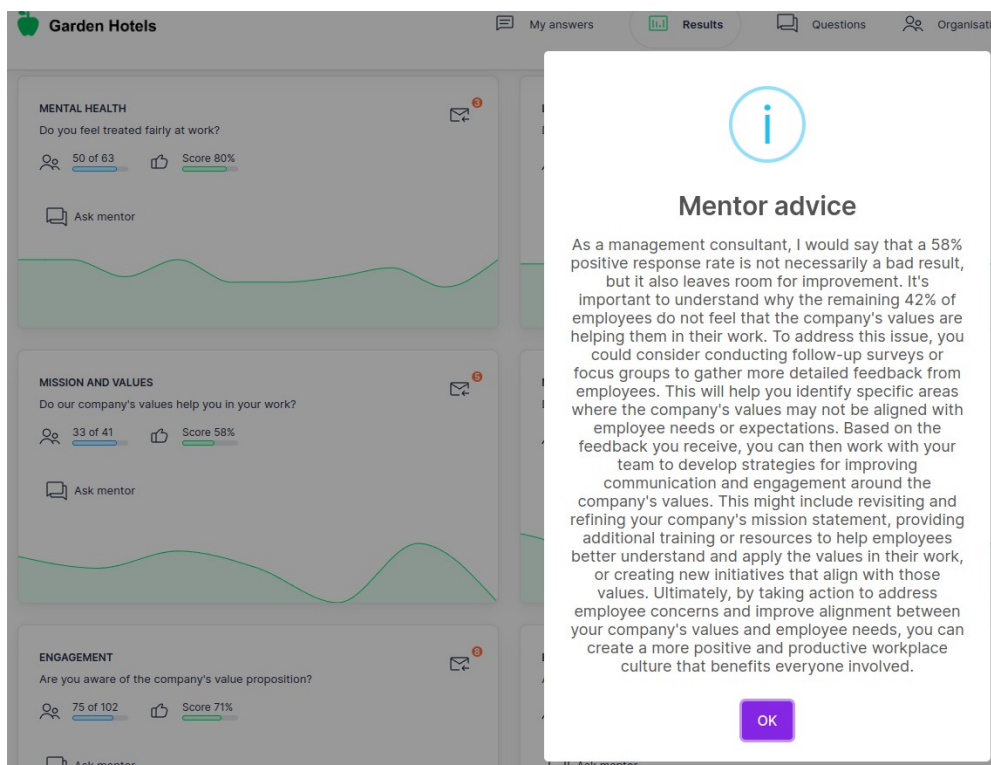
29.11.2022
Good enough

Testimise käigus ilmnes tähelepanuväärne erinevus, kuidas ChatGPT vastas samale küsimusele, mis oli esitatud esmalt vene ja siis inglise keeles, erineval moel. Seejuures hindas ta vene keeles vastates, et 58% positiivselt vastanud on hea tulemus, samas kui inglise keeles pidas ta seda pigem negatiivseks tulemuseks.



Joonis 5. ChatGPT vastus vene keeles.

Vastus samale küsimusele inglise keeles oli üsna erinev.



Joonis 6. ChatGPT vastus inglise keeles

Selline tulemuste erinevus on küllaltki märkimisväärne.

Lisa 3 – Töö käigu üksikasjalikum kirjeldus

Antud lisa esitatakse päevikusarnases vormis üksikasjalikum kirjeldus töö käigust.

Töö algas esimese koosolekuga Motichecki esindajaga, kes selgitas ettevõtte vajadust lähemalt. Tema soov oli luua Motichecki kasutavatele juhtidele rohkem väärtust, pakkudes töötajate tagasiside põhjal soovitusi olukorra parandamiseks. Tundus, et selleks võiks kõige paremini sobida ChatGPT, mis on tegelikult sisuliselt ka ainus piisavalt võimas avalikult kättesaadav tööriist sellise ülesande lahendamiseks.

Eeltöö

Enne järgmist koosolekut tegi autor ChatGPT-ga mõned katsetused töötajate tagasiside teemadel. Selleks uuris autor Motichecki testkeskkonda, kus oli üks testandmetega ettevõtte, millele oli testandmetena lisatud ka mitmeid töötajate tagasiside vastuseid. Testandmete põhjal proovis autor välja mõelda mõned küsimused, mis võiksid tekkida töötajate tagasiside statistikat uurival ettevõtte juhil, ning katsetas nende esitamist ChatGPT-le. Saadud tulemused koguti Google Sheets keskkonda, et neid saaks hiljem analüüsida ja töötada välja küsimused, millele ChatGPT annaks kõige kasulikumad vastused. Kuna ChatGPT puhul sõltub palju just sellest, kuidas küsimus sõnastada.

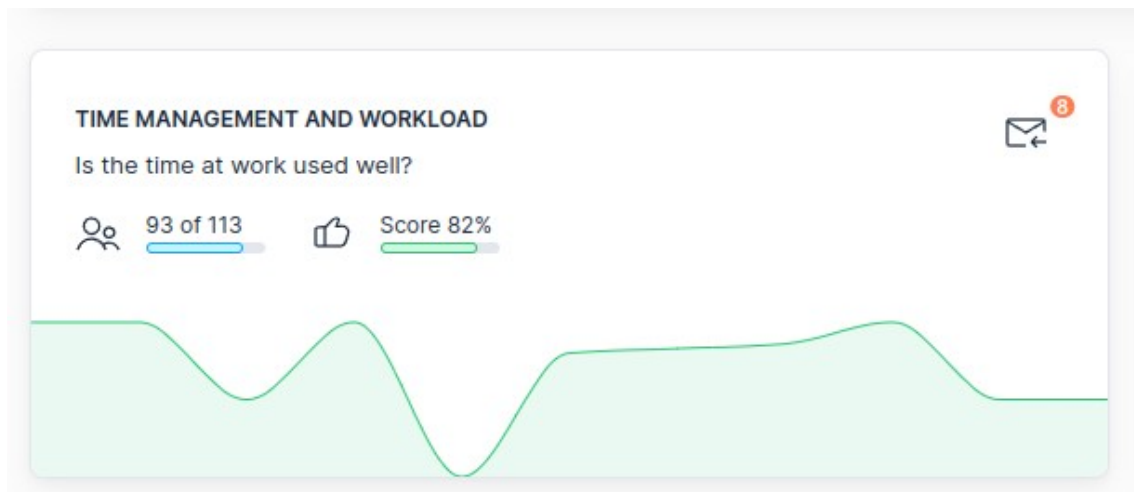
Samuti uuris autor, millised on tehnilised võimalused sellega liidestuda. Tuli välja, et ChatGPT-l ei ole veel ametlikku REST APIt või muud masinliidest, mille abil seda teistesse rakendustesse lihtsasti integreerida saaks. ChatGPT välja töötanud ettevõtte OpenAI veebilehel oli kirjas, et varsti on kavas selline REST API turule tuua.

Siiski oli REST API kaudu saadaval GPT-3 teenus, mis kasutab DaVinci keelemudelit ning toimib üsna sarnaselt ChatGPT-ga. Erinevus seisneb aga selles, et ChatGPT on optimeeritud just interaktiivse vestluse jaoks. GPT3 sobib aga pigem ühekordsete küsimuste esitamiseks. Sellegipoolest oli selle API olemasolu positiivne avastus, kuna tekkis lootus, et kuni ChatGPT avaliku API väljatulekuni saab kasutada GPT3 APIt.

Ühel järgneval koosolekul arutas autor Motichecki esindajaga juba lähemalt, kuidas probleemi lahendada. Vaadati välja koht kasutajaliideses, kuhu võiks tekitada nupu, mis käivitaks vestlusakna GPT3-ga. Samuti valmisid koosoleku tulemusena esimesed kasutajalood.

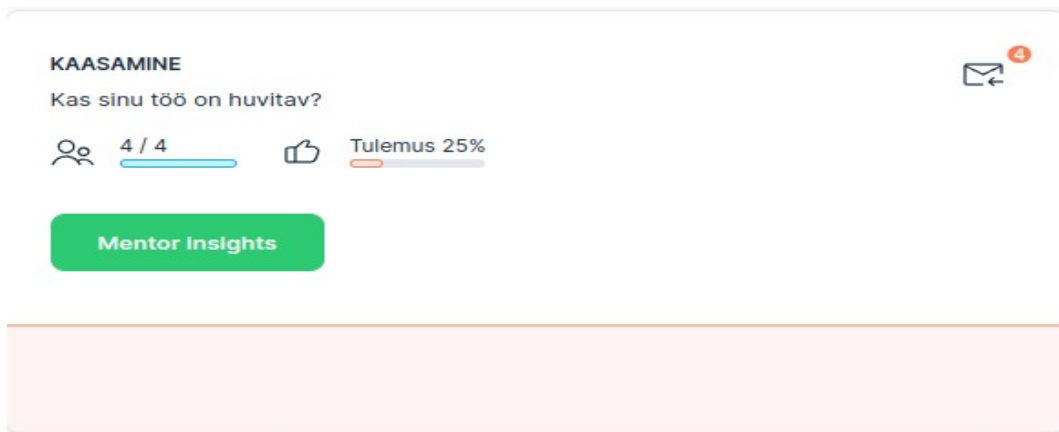
Algse prototüübi loomine

Järgmise sammuna laadis autor enda arvutisse Motichecki lähtekoodi ning käivitas rakenduse lokaalselt. Seejärel oli võimalik tuvastada, millist kohta oleks vaja muuta, et lisada GPT-3 vestluse nupp. Selleks kohaks osutus küsimuse kaart, mida kuvatakse ettevõtte juhi vaates.



Joonis 7. Küsimuse kaart Motichecki rakenduses.

Järgmise sammuna lisas autor nimetatud küsimuse kaardile nupu, mille abil saab tehisintellekti teenuselt nõu küsida.



Joonis 8. Küsimuse kaart koos uue nupuga, mis kuvab tehisintellekti soovitusi.

Seejärel tuli valmis kirjutada funktsionaalsus, mis pärast nupule vajutamist käivitaks päringu GPT-3 api suunas. Selleks tuli päringule kaasa panna ka konteksti loov küsimus, Motichecki enda küsimus, mille kohta sooviti roboti käest infot, ning vastuste tulemus. Prototüüpimise faasis toimus kõik esirakenduse koodis, kuna seda oli plaanis vaid autori enda arvutist Motichecki esindajale näidata. Esialgne küsimus, mis GPT-3 mudelile saadeti, oli järgmine:

```
`I will now speak in Estonian. Ma olen ettevõtte juht. Me küsisime oma töötajatelt: '{question}'. ${answerResult} vastasid 'jah'. Mida peaksin tegema?`
```

Antud näites on question muutuja, mis saab väärtuse sõltuvalt konkreetsest küsimusest, mille kaardil nupp parasjagu asub. Muutuja answerResult näitab positiivse vastuse andnute arvu ja vastanute koguarvu, nt 3/5.

GPT-3 poolt genereeritud vastuse kuvamiseks hüplikaknas kasutas autor sweetalert2 teeki (<https://github.com/sweetalert2/sweetalert2>). See võimaldab väga lihtsa vaevaga kuvada küllaltki hea visuaaliga hüplikaknaid.

Seejärel toimus järjekordne koosolek Motichecki esindajaga, kus autor näitas GPT-3 nupu prototüüpi. Katsetamise käigus ilmnnes, et GPT-3 poolt genereeritud vastused olid palju lakoonilisemad, kui ChatGPT omad. Lisaks tundus, et eestikeelsete vastuste kvaliteet on madalam kui ingliskeelsete puhul. Üldiselt jäi Motichecki esindaja prototüübiga rahule. Lepiti kokku, et tehakse mõned muudatused disainis, et nupp ei oleks nii silmatorkav. Järgmise sammuna aga tuleks koodimuudatused paigaldada

testkeskkonda, et GPT-3 nupu funktsionaalsust saaks testida rohkem inimesi, sealhulgas klientide fookusgrupp.

Selleks, et muuta vastuseid põhjalikumaks, lisas autor vastava korralduse ka GPT-3 küsimusepüstitusse. See toimis, vastus oli juba pikem.

`I will now speak in Estonian. Ma olen ettevõtte juht. Me küsisime oma töötajatelt: '\$ {question}'. \$ {answerResult} vastasid 'jah'. Palun vasta pikalt ja põhjalikult. Kas see on hea tulemus? Mida peaksin tegema?`

Ettevalmistused testkeskkonda paigaldamiseks

Kui algne prototüüp oli valmis, siis järgmiseks oli vaja teha 2 sammu: kirjutada prototüübi kood ümber, järgides parimaid praktikaid; seejärel paigaldada koodimuudatused testkeskkonda, et ChatGPT nupu funktsionaalsust saaks testida rohkem inimesi.

Testkeskkonda paigaldamisega seoses tuli läbi mõelda järgmine küsimus. Eesmärk oli paigaldada uus funktsionaalsus testkeskkonda, aga samas tagada, et see kogemata toodangusse ei jõuaks. Selleks oli kaks võimalust:

1. Luua versioonihaldustarkvaras eraldi haru, kus hoida uut funktsionaalsust, ning paigaldada testkeskkonda just see haru, luues sisuliselt uue testkeskkonna.
2. Piirata ChatGPT nupu nähtavust rakenduse siseselt, et sellele pääseksid ligi näiteks ainult konkreetse ettevõtte esindajad või konkreetsed kasutajad. Tegemist oleks nn *feature flagiga*, mis võimaldab töötavas rakenduses teatava konfiguratsiooni abil lülitada sisse või välja teatavat funktsionaalsust.[5]

Valitud sai teine variant, kuna see on antud olukorras mõistlikuma ja lihtsam, ning võimaldab kiiremini testimise asuda.

Sel ajal tuli välja ka ChatGPT API, mida varem ei olnud. Seetõttu sai autor asendada senise GPT-3 text-davinci mudeli põhise päringu ChatGPT päringuga. Päringut tuli selleks muuta minimaalselt, kuna liides oli suhteliselt sarnane. Juba esimesed katsetused näitasid, et ChatGPT vastus on pikem, põhjalikum, sisukam ja sujuvama keelekasutusega.

Arenduse käigus oli eesmärk ChatGPTfunktsionaalsust ülejäänud rakendusest võimalikult isoleerida. Isoleeritus on kasulik kahel põhjusel. Esiteks on sellisel juhul lihtsam ChatGPT funktsionaalsust hiljem vajadusel eemaldada, näiteks kui tuleb välja, et selle vastu pole klientidel siiski loodetud huvi. Teiseks aitab koodi isoleerimine vähendada arenduse käigus tekkida võivaid koodikonflikte, kui mitu arendajat muudavad korraga sama faili, ning seeläbi arendusprotsessi kiirendada ja tõhustada.

Seejärel tuli tagarakendusse teha uus API *endpoint* mentor/advice, millele saaks kaasa panna ChatGPT-le esitatava küsimuse (prompt), ning mis tagastaks ChatGPT vastuse. Uut *endpoint*-i implementeeris MentorAdviceController, mis tegeleb kasutaja sessiooni kontrollimisega (kas kasutaja on sisse logitud või mitte). Lisaks kontrollib see, kas konkreetse kasutaja jaoks on ChatGPT funktsionaalsus aktiveeritud või mitte. Kui need kontrollid on tehtud, delegeerib see töö edasi MentorService-le, mis paneb kokku päringu ja saadab selle ChatGPT API-le, seejärel aga tagastab sealt saadud vastuse.

Selleks, et funktsionaalsust saaks kasutajapõhiselt sisse ja välja lülitada, tuli täiendada andmebaasi mudelit ning lisada kasutaja seadete tabelisse uus veerg, mis näitab, kas selle kasutaja jaoks on ChatGPT sisse lülitatud või mitte.

Mentor on kõnealuse tehisintellekti funktsionaalsuse äripline nimetus, mida kasutatakse eka koodis. Kui Motichecki kasutaja tahab saada küsimuse kohta nõu, siis ta pöördub (robot)mentori poole, kelle rolli täidab tegelikult ChatGPT.

Arenduse käigus said tagarakendusele lisatud mentori funktsionaalsuse testimiseks ka integratsioonitesti. Integratsioonitesti tegemiseks kasutas autor ChatGPT API päringu simuleerimiseks laialt levinud *mock*-imise tehnikat, kus testi jaoks realselt ChatGPT poole ei pöördutud, kuid veenduti, et rakenduse töö käigus käivitatakse päring õigete parameetritega. *Mock*-imise jaoks oli kasutusel Mockito raamistik.

Kui esmane mentori funktsionaalsus oli valmis, paigaldati see testkeskkonda, et koguda tagasisidet äripoole testijatelt. Testimise käigus laekus kasulikku tagasisidet. See puudutas nii kasutajaliidese poolt kui ka ChatGPT poolt genereeritud vastuste sisulist poolt.

Järgmise sammuna lisas autor funktsionaalsuse, et ChatGPT suudaks vastata erinevates keeltes sõltuvalt sellest, mis keel on kasutajal parasjagu valitud. See oli suhteliselt lihtne muudatus. Tuli lihtsalt võtta kasutusele olemasolev tõlgete süsteem ning tõlkida ChatGPT-le esitatav küsimus enne teele saatmist selle süsteemi abil. See tõlkimine toimub esirakenduses. Tagarakenduses aga midagi muutma ei pidanud, kuna see võtab parameetrina sisse lihtsalt sõne.

Sellega sai esimene kasutajalugu valmis ning võis liikuda järgmise poole.

ChatGPT küsimuste peenhäälestamine

Nagu teada, on ChatGPT puhul, samamoodi nagu ka mitmete muude keelemudelite puhul, üsna oluline, kuidas oma küsimus täpsemalt sõnastada. Pärast esmast testimist testkeskkonnas oli tagasiside üldiselt positiivne, kuid Motichecki esindajad andsid teada, et küsimust võiks muuta ja panna kaasa rohkem infot ja konteksti.

Väljapakutud küsimus oli selline.

I will ask you a question in English. Please provide your response in the `x` language.

I am a manager of a department at a company.

We are monitoring our employee experience daily in 13 categories. The average score of these categories for a period of '`up_to_last_3_months`' is `{av score}`%.

The score for '`{category name}`' in my team is `{category score}`%.

More specifically, we asked our people: '`{question_text}`'. The rate of positive answers was `{score}`%.

Only if filters are applied!

These specific people are working in (`country` / `x` roles)
They have been working for (`less than 1`, `1 to 3 years`, `3 to 10 years`, `more than 10 years`) in the organization.

Only if comments are available!

Employees' comments included:

'(comment 1)'

'(comment 2)'

'(comment 3)'

'(comment n)'

How would you comment on that?

What should be the 3 steps to take next?

What should I focus on first?

Nagu näha, sisaldab ChatGPT prompti teine versioon oluliselt rohkem täiendavaid andmeid. Kui esimeses versioonis oli ära toodud vaid küsimuse tekst ning positiivsete vastuste protsent, siis teises versioonis on lisatud viimase kolme kuu keskmine skoor üle kõigi valdkondade, samuti keskmine skoor selle valdkonna kohta, millesse kuulub vaatlusalune küsimus. Samuti lisatakse küsimusele kaasa filtrid, juhul kui Motichecki kasutaja on neid rakendanud, näiteks kui on valitud vaid teatud riigi või ametikoha töötajad. Lisaks on küsimusele kaasa pandud töötajate poolt jäetud kommentaarid, kui neid on.

Selleks, et eelmainitud vajalikud andmed kokku koguda, oli vajalik lisaarendus. Sellega aitas Motichecki arendaja, et saaks tööd teha paralleelselt, samas kui autor tegeles muude funktsionaalsustega. Eesmärk oli saada kogu mentori funktsionaalsus kiiremini testkeskkonda.

Sellest küsimuse peenhäälestamise arendusest nähtub ka lisandväärtus, mida ChatGPT-d integreerivad rakendused võivad pakkuda. Võiks ju küsida, milles seisneb Motichecki lisandväärtus ChatGPT vahendajana? Selle asemel, et vajutada nupule Motichecki keskkonnas, võiks kasutaja sama hästi minna ise ChatGPT lehele, ning seal hoopis ise küsimuse esitada. Ainsaks lisavaevaks oleks vajadusel ChatGPT lehel konto tegemine ja sisselogimine. Antud funktsionaalsuses on aga näha, kuidas Motichecki rakendus paneb kasutaja eest ise kokku põhjaliku ChatGPT prompti ning lisab sinna olulisel määral taustainfot (3 kuu keskmised skoorid jne), mis on Motichecki rakenduses olemas. Tänu sellele on lootus, et ChatGPT poolt antud vastus saab sisukam ja informatiivsem. Sellise prompti käsitsi kokkupanemine oleks kasutaja jaoks küllaltki märkimisväärne töö.

Samuti ei pruugi kasutajal olla teadmisi, kuidas tõhusat prompti koostada ja millistele kriteeriumitele peaks üks tõhus prompt vastama. Mida enam aga suured keelemudelid laia kasutusse jõuavad, seda olulisemaks oskuseks muutub just keelemudelite jaoks promptide koostamine. Seda kinnitavad ka mitmed artiklid.

Mentori ajaloo kasutajaliides

Kui mentori ajaloo andmemudel oli valitud, oli järgmiseks sammuks realiseerida see tagarakenduse koodis ning esirakenduses välja kuvada.

Tagarakenduses olid täiendused üsna lihtsad. Rakendusele tuli lihtsalt luua võimekus uude mentori interaktsioonide tabelisse andmeid sisestada ja neid sealt lugeda. Mentori teenusekihis, kus küsitakse mentori nõuandeid, tuli aga lisada täiendava sammuna küsimuste ja vastuste salvestamine. Samuti tuli lisada uus *API endpoint*, mille kaudu saab esirakendus küsida välja kõik konkreetse kasutajaga seotud mentori interaktsioonid, et neid saaks hiljem kasutajaliideses välja näidata.

Esirakenduses oli ärinõue, et mentori sõnumite ajaloo väljakuvamiseks kasutataks juba olemasolevat juturoboti kasutajaliidest, mida seni kasutati peamiselt küsimuste küsimiseks. See vajas mõningat aega, et mõista, kuidas juturoboti kasutajaliides täpsemalt toimib ning mis kujul tuleks talle anda sisse mentori interaktsioonide ajalugu, et ta seda sobivalt välja kuvas. Tulemus on näha järgneval joonisel.

communication between management and employees. It's important to remember that improving employee engagement is an ongoing process that requires consistent effort and attention from management. By taking steps to address the issue of disengagement, you can create a more productive and fulfilling workplace for your employees.

I am a manager at a company. We asked our employees: 'Does your manager support your development goals?'. 100% answered 'yes'. Is this a good result? What should I do?



Congratulations on receiving a 100% positive response from your employees! This is a great result and indicates that your management team is doing an excellent job of supporting the development goals of your employees. To maintain this positive result, you should continue to prioritize employee development and ensure that managers are providing regular feedback, coaching, and training opportunities to their team members. You may also want to consider conducting regular surveys or check-ins with employees to ensure that they feel supported in their career growth. Additionally, you could explore ways to enhance your employee development programs, such as offering mentorship opportunities or providing access to online learning resources. By investing in the growth and development of your employees, you can help them reach their full potential while also benefiting your company through increased productivity and employee retention.

Joonis 9. Mentori küsimuste ajaloo kasutajaliides.

Nagu näha, kuvatakse mentori küsimuste ajalugu sõnumirakendustele omase kasutajaliidese abil. See kasutajaliides oli Motichecki rakenduses juba olemas, ning autoril tuli seda vaid natuke modifitseerida ja kasutusele võtta.

Lühendatud küsimuse kuvamine kasutajale

Kuna ajapikku olime muutnud ChatGPT-le saadetava küsimuse üsna pikaks ja põhjalikuks ning varustanud selle taustainfoga (vt ptk 5.1.4), siis oli mõistlik muuta mentori ajaloo vaadet nii, et kuvada kasutajale küsimuse lühiversioon, mitte kogu prompt, mille rakendus ChatGPT-le tegelikult saadab. Peamine eesmärk on selles, et pilt ei läheks kasutaja jaoks liiga kirjuks ning fookus püsiks olulisel.

Selle arenduse realiseerimine oli üsna lihtne. Mentor_interactions tabelile tuli lisada täiendav veerg, mis sisaldaks kasutajale kuvatavat küsimust. Seega salvestas süsteem nii täispika küsimuse, mis ChatGPT-le tegelikult saadeti, kui ka lühendatud küsimuse, mida hiljem ajaloo vaates kasutajale kuvatakse. See arendus oli küllaltki lihtne ja triviaalne.

Lisa 4 – Kasutajalood

Lisada nupp küsimuse kaardile, mis käivitab chati päringu

Kasutajaloo kirjeldus

Juhina ma tahaksin,

et küsimuse kaardil oleks nupp, millele vajutades näeksin soovitusi, mida selles olukorras teha,

et ma saaksin olukorda parandada.

Lisanõuded

Nuppu peaksid nägema ainult kasutajad, kelle jaoks on see featuur sisse lülitatud.

Esiolgu testimiseks sobib, kui vastused on alati eesti keeles.

Tehnilised detailid

Võimalik on kasutada OpenAI endpointi <https://api.openai.com/v1/chat/completions>

Päringusse saab kaasa panna küsimuse teksti ning vastuste statistika (skoor). Need andmed on küsimuse kaardil olemas.

Teha backendis eraldi endpoint, mis võtab sisse prompti, ja tagastab chatGPT vastuse.

Employee tabelisse lisada uus veerg isMentorEnabled, mis võimaldaks piirata kasutuskonda.

Lahtised küsimused

- Mida me Chati käest täpsemalt küsime. Nt “Mida peaksin selles olukorras tegema?”, “Tee mulle
- kokkuvõtte vastustest ja kommentaaridest.” vms.
- Kas peaks kaasa panema ka kommentaarid?

- Kuhu ja kuidas kuvada vastus?
- Tõlked?
- Kas see nupp peaks olema kõigil klientidel või peaks kuidagi eristama (Premium feature)?

Lisame eraldi chati vaate, kuhu jääb kogu ajalugu ja on menüüs eraldi nupu all

Kasutajaloo kirjeldus

Juhina ma tahaksin,

näha mentori vestluste ajalugu,

et saaksin hiljem üle vaadata, mida ta on mulle soovitanud.

Lisanõuded

Ajaloo näitamiseks tuleks kasutada olemasolevat chatboti vaadet.

Ajaloo vaates peaks kasutajale kuvama lihtsustatud versiooni promptist, mida realselt ChatGPT-le saatsime.

Tehnilised detailid

Luu eraldi andmebaasi tabel `mentor_interaction`.

MentorService peaks ChatGPT-le saadetud küsimused ning saadud vastused salvestama andmebaasi.

Luu eraldi API endpoint, mille kaudu saab küsida kõik ühe töötaja mentori interaktsioonid.

Lahtised küsimused

- kuidas arvestada mitmekeelsusega?