



Satakunnan ammattikorkeakoulu
Satakunta University of Applied Sciences

VIIVI RIIKONEN

Chatbotin käyttöönotto Satakunnan ammattikorkeakoulun verkkosivuilla

TIETOJENKÄSITTELYN TUTKINTO-OHJELMA
2024

TIIVISTELMÄ

Riikonen, Viivi: Chatbotin käyttöönotto Satakunnan ammattikorkeakoulun verkkosivuilla
Opinnäytetyö, AMK
Tietojenkäsittelyn tutkinto-ohjelma
Marraskuu 2024
Sivumäärä: 33

Tässä opinnäytetyössä käydään läpi chatbotin käyttöönottoa verkkosivuilla. Kyseinen chatbot otettiin käyttöön Satakunnan ammattikorkeakoulun verkkosivuilla kesäkuussa 2024. Opinnäytetyön toimeksiantajana toimii Satakunnan ammattikorkeakoulu.

Opinnäytetyössä käydään läpi chatbotin käyttöönottoprosessi projektin aloituksesta julkaisuun asti. Työn alussa perehdytään yleisesti chatbottien historiaan sekä kehitykseen. Käyttöönottoprosessin seuraamisen lisäksi tutustutaan tämän käyttöönotetun chatbotin ominaisuuksiin, toiminnallisuuksiin sekä projektissa ilmenneisiin haasteisiin. Lopuksi pohditaan vielä botin jatkokehitystä, tulevaisuutta ja hyödyllisyyttä organisaatiolle.

Avainsanat: Chatbot, Aie, Satakunnan ammattikorkeakoulu, Tekoäly

ABSTRACT

Riikonen, Viivi: Implementation of a chatbot on the Satakunta University of Applied Sciences website
Bachelor's thesis
Business Information Systems
November 2024
Number of pages: 33

This thesis discusses the implementation of a chatbot on a website. The chatbot was launched on the Satakunta University of Applied Sciences (SAMK) website in June 2024. The thesis commissioner is Satakunta University of Applied Sciences.

The thesis covers the chatbot implementation process from the project initiation to its release. The thesis starts with a general introduction to the history and development of chatbots. In addition to following the implementation process, the thesis explores the features, functionalities, and challenges encountered during the project. Lastly, the thesis reflects on the bot's further development, future prospects, and its usefulness to the organization.

Keywords: Chatbot, Intent, Satakunta University of Applied Sciences, Artificial Intelligence

SISÄLLYS

1 JOHDANTO	5
2 CHATBOT	6
2.1 Chatbottien historiaa.....	7
2.1.1 Turingin testi	7
2.1.2 Ensimmäinen chatbot	8
2.1.3 A.L.I.C.E. ja Jabberwacky.....	9
2.1.4 GPU	11
2.1.5 Myöhempi kehitys	11
2.2 Chatbotit Suomessa	13
2.3 Satakunnan ammattikorkeakoulun chatbot.....	13
3 PALVELUNTARJOAJA	15
3.1 Boost.ai	15
3.2 Front AI.....	16
4 PROJEKTIN ALOITUS	18
4.1 Kick-off	18
4.2 Botin kehitysalusta	18
4.3 Edistymisen seuranta	21
5 KÄYTTÖÖNOTTO	22
5.1 Aikeet	22
5.2 Synonyymit.....	23
5.3 Optimointi	26
6 BOTIN JULKAISU	28
7 HAASTEET PROJEKTISSA	29
7.1 Haasteet botin vastausten luomisessa	29
8 YHTEENVETO.....	31
LÄHTEET.....	32

1 JOHDANTO

Digitalisaation myötä palveluiden ja asiakaskokemuksen kehittäminen on noussut keskiöön monilla aloilla. Chatbot, eli keskusteleva tekoäly, on ollut yksi merkittävimmistä innovaatioista asiakaspalvelussa ja tiedonhakuprosesseissa. Tämä opinnäytetyö käsittelee chatbotin käyttöönottoa Satakunnan ammattikorkeakoulun (SAMK) verkkosivuilla, tavoitteenaan parantaa opiskelijoiden ja hakijoiden saavutettavuutta ja asiakaspalvelun tehokkuutta. Suoritin opintoihini kuuluvan harjoittelun SAMKissa ja tämän myötä pääsin mukaan chatbotin käyttöönottoprojektiin, sillä olin kiinnostunut tekemään aiheesta opinnäytetyön. Työskennellessäni SAMKissa syksyllä 2024, pääsin mukaan myös botin jatkokehitysvaiheeseen.

Opinnäytetyössä käsitellään chatbotin käyttöönottoprosessia vaihe vaiheelta, alkaen projektin alkuvaiheista ja päättyen chatbotin julkaisuun. Alussa tarkastellaan chatbotien historiaa ja kehitystä, eli kuinka chatbotit ovat kehittyneet yksinkertaisista ohjelmista nykyisiin edistyneisiin tekoälypohjaisiin ratkaisuihin. Tämä osio auttaa ymmärtämään, kuinka chatbotit ovat tulleet osaksi nykyteknologiaa ja miten ne ovat kehittyneet.

Käyttöönottoprojektin seuraamisen lisäksi työssä syvennytään juuri tähän SAMKissa käyttöönotettuun chatbotiin, sen ominaisuuksiin ja toiminnallisuuksiin, kuten siihen, kuinka se toimii käytännössä ja mitä se pystyy tekemään käyttäjän kanssa keskustellessa. Samalla tarkastellaan myös niitä haasteita, joita projektissa on kohdattu käyttöönoton aikana. Opinnäytetyön lopulla tarkastellaan chatbotin käyttäjämääriä julkaisun jälkeen sekä pohditaan sen jatkokehitystä ja tulevaisuutta. Yhtenä suurena kysymyksenä on tämän botin relevanttius jatkuvasti kehittyvän tekoälyn rinnalla sekä se, tuleeko tämä botti oikeasti hyödyttämään Satakunnan ammattikorkeakoulua niissä asioissa, joiden vuoksi se on hankittu.

2 CHATBOT

Lyhyesti selitettynä chatbot tarkoittaa tietokoneohjelmaa, jonka kanssa voidaan keskustella ja joka on suunniteltu interaktiiviseen tai automatisoituun keskusteluun ihmisten kanssa. Chatbot antaa vastauksia siltä kysytyihin kysymyksiin sekä esittää kysyjälle vasta- ja lisäkysymyksiä. Se voi tunnistaa tekstistä avainsanoja ja vastata tai muodostaa kysyjälle lisäkysymyksiä hänen syöttämänsä tekstin sekä aiempien vastausten perusteella. (Nieminen, 2022.)

Chatbot-ohjelmistot voidaan luokitella kolmeen eri ryhmään niiden toteutustavan mukaan:

- **Käsikirjoitetut botit:** Ihminen on miettinyt ja rakentanut keskustelupolut valmiiksi.
- **Älykkäät botit:** Botin taustalla toimii tekoäly, ja sitä opetetaan syötteiden sekä oikeiden vastausten avulla. Toisin kuin käsikirjoitetut botit, älykkäät botit voivat oppia ja kehittyä palautteen perusteella.
- **Hybridit:** Ohjelmistot, joissa on yhdistelty sekä tekoälyä että ihmisten rakentamia keskustelupolkuja. (Tyyskä, 2022.)

Chatbotin hyödyntämisessä on paljon etuja sekä yrityksen että käyttäjän näkökulmasta. Nykypäivän ihmisillä on tapana pitkästyä hyvin nopeasti. He vaativat helposti ja nopeasti saatavilla olevaa palvelua, eivätkä halua käyttää liikaa aikaa tai vaivaa etenkin yksinkertaisten asioiden selvittämiseen. (Nieminen, 2022.)

Yritykselle chatbot mahdollistaa keskustelun asiakkaiden kanssa ympäri vuorokauden, viikon jokaisena päivänä. Käyttäjän näkökulmasta suurin hyöty on erityisesti se, että chatbotin käyttö on hyvin helppoa, nopeaa ja ajasta riippumatonta, eikä se vaadi erikseen perehtymistä. Käyttäjälle jää myös positiivinen kuva palvelusta, jos yksinkertaisiin kysymyksiin vastauksen saaminen on ketterää, vastauksen voi saada viikonpäivästä ja kellonajasta huolimatta eikä asian selvittäminen välttämättä vaadi edes kontaktia oikean ihmisen kanssa.

2.1 Chatbottien historiaa

Alkeelliset chatbotit kehitettiin ensimmäisen kerran 1900-luvun puolivälissä ja lopulla, ja niistä tuli teknisesti järkevämpiä sekä laajemmin saatavilla olevia 2010-luvun lopulla sekä 2020-luvun alussa, varsinkin, kun tekoäly alkoi saada näkyvyyttä teknologia-alalla. (Gisonna, 2024.)

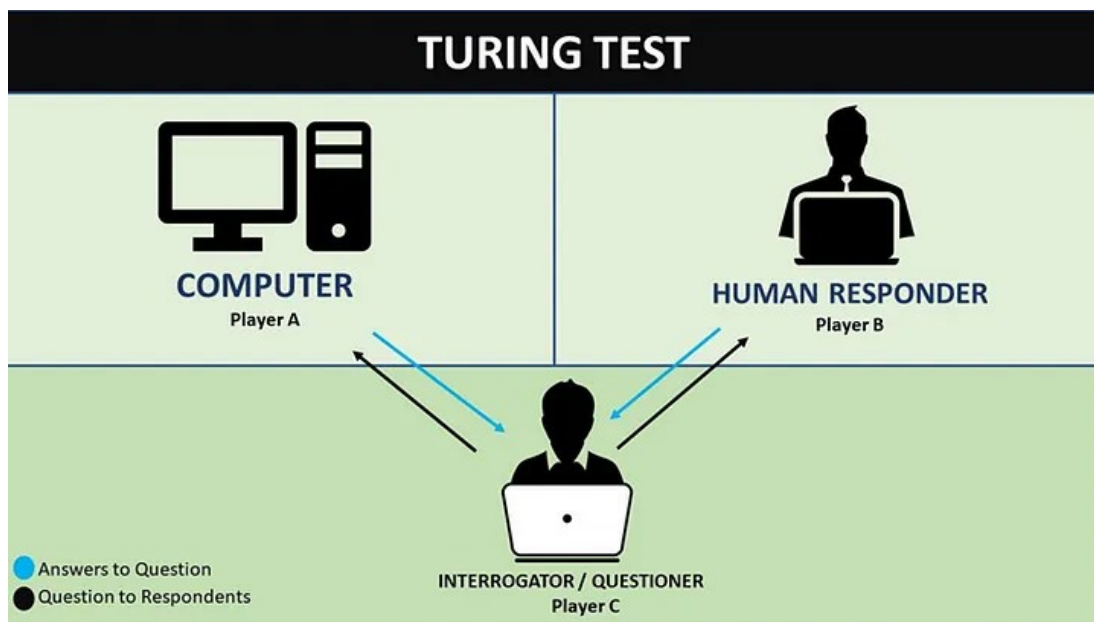
Ensimmäiset chatbotit olivat hyvinkin alkeellisia. Ne olivat joko sääntöpohjaisia, jolloin käyttäjän piti valita sopivin vastausvaihtoehto tai ne olivat tekoälypohjaisia, jotka yrittivät liikaa olla ihmisen kaltaisia. Alkeellisen sääntöpohjaisen chatbotin kanssa keskustelu oli, kuin olisi puhunut robotille. Botin vastaukset olivat hyvin rajallisia ja ennalta arvattavia. Ensimmäisten tekoälypohjaisten chatbottien suunniteltiin olevan ihmisen kaltaisia, mutta todellisuudessa botit eivät kovin usein ymmärtäneet kysymystä oikein ja tämä aiheutti käyttäjissä ymmärrettävästi turhautumista. Nykypäivän chatbotit ovat onneksi paljon älykkäämpiä ja tarjoavat käyttäjälle oikeasti apua sekä oikeita vastauksia, eivätkä yritä olla ihmisen kaltaisia. Bottien älykkyys perustuu siihen, että nykypäivän chatboteissa voidaan käyttää hybridimallia. Tämä tarkoittaa sitä, että chatbot-ohjelma on yhdistelmä sääntö- ja tekoälypohjaista tekniikkaa. (Sachdeva, 2024.)

2.1.1 Turingin testi

Iso-Britannialainen matemaatikko ja logiikan tutkija Alan Turing oli ensimmäinen, joka konseptoi idean chatbot-ohjelmia edeltäneeseen teknologiaan. Vuonna 1950 Turing esitteli ideansa tekoälyn testaamiseen. Tätä ideaa kutsutaan Turingin testiksi. (Gisonna, 2024.)

Turingin testin tavoitteena oli tutkia, voiko tietokone käydä vuorovaikutusta tuntevan olennon kanssa. Testissä ihmiskuulustelijan on tietyn ajan kuluessa erotettava tietokone ihmisestä sen perusteella, miten he vastaavat kuulustelijan esittämiin kysymyksiin. Testi katsottiin tietokoneen kannalta onnistuneeksi, mikäli ihmiskuulustelija ei testin lopussa erottanut, kumpi

vastaajista on ihminen ja kumpi tietokone. (Lindström, 2020.) Kuvassa 1. on havainnollistettu Turingin testi visuaalisena.



Kuva 1. What is a Turing Test (Visualwebz, 2024.)

2.1.2 Ensimmäinen chatbot

Ensimmäisen chatbotiksi tunnistetun ohjelmiston kehitti saksalaisamerikkalainen tietojenkäsittelytieteilijä Joseph Weizenbaum vuonna 1966. Weizenbaumin kehittämä Eliza-niminen ohjelma pystyi jäljittelemään keskustelua tietokoneen käyttäjän kanssa. Käyttäjä pystyi keskustelemaan ohjelman kanssa käyttämällä keskusyksikköön kytkettyä sähköistä kirjoituskonetta, ja Eliza sitten tarkasteli lauseen hahmontunnistusalgoritmin avulla. Algoritmi vertasi käyttäjän antamaa syötettä sääntöjoukkoon, joka sitten tuotti sopivia lauseita vastaukseksi. Vaikka tätä pidettiin historiallisena kehityksenä tietojenkäsittelytieteen alalla, Eliza ei kyennyt läpäisemään aiemmin mainittua Turingin testiä, koska se tuotti ajoittain liian epä johdonmukaisia vastauksia. (Gisonna, 2024.)

Toisin kun voisi luulla, Weizenbaum ei suunnitellut Elizaa avuksi asiakaspalveluun tai virtuaaliseksi ystäväksi, vaan sen tarkoitus oli tutkia

ihmisen ja tietokoneen välisen viestinnän luonnetta. Eliza käytti tekniikkaa, jota kutsutaan kuvioiden yhdistämiseksi ja avainsanojen tunnistamiseksi. Käyttäjä pystyi kirjoittamaan tekstiä, josta Eliza etsi avainsanoja tai kuvioita. Ohjelma vastasi käyttäjälle ennalta kirjoitetuilla lauseilla tai kysymyksillä, jotka muotoilivat uudelleen käyttäjän itse syöttämää tekstiä. Tällainen yksinkertainen tekniikka loi käyttäjälle illuusion keskustelusta ja ymmärryksestä, vaikka todellisuudessa Eliza ei ymmärtänyt käyttäjän tunteita tai tekstin syvyyttä. Rajoitteistansa huolimatta Eliza pystyi yllättämään monet, mukaan lukien sen kehittäjän. Ihmiset osallistuivat ohjelmaan yllättävän henkilökohtaisella tavalla, luottaen ajatuksensa ja tunteensa Elizalle. Tämä odottamaton lopputulos pakotti ihmiset kohtaamaan inhimillisen halun löytää merkitys viestinnässä, jopa koneiden kanssa. Eliza näytti tietä tulevaisuuden chatbot-kehitykselle, osoittaen, että ihmisen ja tietokoneen välillä voi olla luonnollista vuorovaikutusta. Se sai aikaan myös tärkeitä keskusteluja tekoälystä ja ihmissuhteiden luonteesta teknologian aikakaudella. (Gandhi, 2024.)

2.1.3 A.L.I.C.E. ja Jabberwacky

Vuonna 1995, amerikkalainen tutkija Richard Wallace julkaisi chatbotin nimeltä A.L.I.C.E. Tämä nimi muodostuu sanoista Artificial Linguistic Internet Computer Entity. Luodessaan A.L.I.C.E.:n Wallace paransi Weizenbaumin aiemmin kehittämän Elizan käyttämää tekoälyä, seurattuaan sen keskusteluja. Kun A.L.I.C.E.:lle esitti lauseen, jota se ei osannut tunnistaa, Wallace lisäsi sille käytettäväksi uuden vastauksen, sen sijaan, että antaisi sen tuottaa kerta toisensa jälkeen väärän vastauksen. A.L.I.C.E.:n suosion kasvaessa muut ohjelmistosuunnittelijat pystyivät lisäämään omia vastauksiaan chatbotin niin sanottuun spin-off-ohjelmistoon. Siten botti pystyi vastaamaan aiempaa monipuolisemmin sille esitettyihin kysymyksiin. (Gisonna, 2024.)

Brittiläinen tietojenkäsittelytieteilijä Rollo Carpenter käytti yli vuosikymmenen chatbotin kehittämiseen ja lopulta julkaisi Jabberwacky nimisen chatbotin

vuonna 1997. Kun se julkaistiin ensimmäisen kerran, Jabberwacky käytti samanlaista sääntöpohjaista lähestymistapaa vastausten luomiseen, kuin sitä edeltäneet chatbotit, esimerkiksi jo aiemmin mainittu Eliza. (Gisonna, 2024.)

Carpenterin keskeinen tavoite oli kehittää tekoälypohjainen ohjelmisto, joka pystyy oppimaan vuorovaikutuksesta ihmisten kanssa. Tämä tarkoittaa sitä, että ohjelmisto ei vain reagoi ennalta määriteltyihin kysymyksiin, vaan se pystyy myös ymmärtämään ja muistamaan, mitä ihminen on sille sanonut. Tämän chatbotin erityinen piirre on sen tekoälypohjainen malli. Monet chatboteista seuraavat ennalta määrättyjä kaavoja, mikä tekee niiden vastauksista helposti tunnistettavia ja jopa hieman mekaanisia. Jabberwacky sen sijaan käyttää lainattua älykkyyttä, mikä tarkoittaa, että se voi oppia ja sopeutua ihmisten keskustelutyyliin. Tämä tekee siitä erottuvan ja ainutlaatuisen verrattuna useimpiin muihin botteihin, sillä se pystyy tarjoamaan monipuolisempia ja inhimillisempiä vastauksia, mikä tekee keskusteluista mielenkiintoisempia ja luontevampia ihmisen näkökulmasta. (Dey, 2019.)

Jabberwackya on kehitetty jatkuvasti sen ensimmäisestä julkaisusta lähtien ja vuonna 2008 ohjelma nimettiin uudelleen, jolloin se sain nimekseen Cleverbot. Cleverbot on tunnettu kyvystään tallentaa ihmisten vastauksia aiemmista keskusteluista, mikä, toisin kuin A.L.I.C.E., antaa sille mahdollisuuden oppia ja kehittyä ilman kehittäjän väliintuloa. (Gisonna, 2024.)

Cleverbot on niin onnistunut ohjelmisto, että se pystyi vakuuttamaan monet ihmiset siitä, että he keskustelevat toisen ihmisen kanssa, eivätkä ohjelmiston. Tämä on merkittävä saavutus, sillä Cleverbot läpäisi Turingin testin 59 prosentilla, kun läpäisyraja on 30 prosenttia. Nykypäivänä Cleverbot pystyy käsittelemään tehokkaasti ja samanaikaisesti jopa noin 7 miljoonaa keskustelua. Vuodesta 2014 lähtien Cleverbot on hyödyntänyt GPU-tekniikoita, mikä parantaa sen suorituskykyä. (Dey, 2019.)

2.1.4 GPU

GPU, eli *Graphics Processing Unit*, on erityisesti graafisten laskentatehtävien suorittamiseen suunniteltu prosessori. Se on olennainen komponentti tietokoneissa ja muissa laitteissa, joissa tarvitaan nopeaa ja tehokasta grafiikan renderöintiä. Sitä käytetään laajasti esimerkiksi videopelien, elokuvien ja 3D-mallinnuksen parissa. GPU:n etuna on sen kyky käsitellä suuria tietomääriä rinnakkain. Tämä tarkoittaa, että se pystyy suorittamaan saman laskentatehtävän useille tietoelementeille samanaikaisesti, mikä tekee siitä erittäin tehokkaan esimerkiksi grafiikan renderöinnissä, koneoppimisessa ja videoeditoinnissa. Nykypäivänä GPU:t sisältävät useita moniprosessoreita, joilla on omat jaetut muistilohkonsa sekä joukko prosessoreita ja rekistereitä. Tämä rakenteellinen suunnittelu mahdollistaa tehokkaan työnjaon ja resurssien käytön, mikä taas parantaa suorituskykyä laskentatehoa vaativissa tehtävissä. Grafiikkasuorittimeen sisältyy myös jatkuva muisti, joka mahdollistaa nopean pääsyn tarvittaviin tietoihin. (AWS Amazon, n.d.)

GPU:lla on tänä päivänä suuri merkitys myös chatbottien kehityksessä. Syvistä oppimisalgoritmeista, kuten neuroverkoista on tullut tehokkaita chatbotien taustalla. Koulutusprosessissa tarvitaan valtavasti laskentatehoa, ja GPU:n avulla voidaan suorittaa monia laskutoimituksia samanaikaisesti, mikä nopeuttaa mallin oppimista. Kun chatbot on koulutettu, se tarvitsee edelleen laskentatehoa vastatakseen kysymyksiin ja tuottaakseen tekstiä reaaliajassa. GPU voi parantaa botin vastausnopeutta sekä vastauksen laatua. Viimeisenä merkittävänä asiana voidaan mainita suorituskyky. GPU:n avulla voidaan käsitellä suuria tietomääriä tehokkaasti, mikä on erityisen tärkeää silloin, kun chatbotit ovat vuorovaikutuksessa useiden käyttäjien kanssa samanaikaisesti. (AWS Amazon, n.d.)

2.1.5 Myöhempi kehitys

Chatbot-teknologia jatkoi yleistymistään 2000- ja 2010-luvuilla. Vuonna 2001 tekoäly-startup ActiveBuddyn kehitystiimi julkaisi chatbotin nimeltä SmarterChild. SmarterChild julkaistiin käytettäväksi AOL Instant

Messengerissä (AIM). AIM oli tuohon aikaan suosittu chathuone ja pikaviestisovellus. AIMin käyttäjät pystyivät lisäämään SmarterChildin ystäväluettelonsa ja siten keskustella sen kanssa. ActiveBuddy julkaisi myöhemmin myös toisen chatbotin AIM:n kautta ja tämä oli nimeltään GooglyMinotaur. (Gisonna, 2024.)

Marraskuussa 2022 tekoälytutkimusta tekevä tutkimuskeskus OpenAI julkaisi tällä hetkellä tunnetuimman chatbotin nimeltä ChatGPT. ChatGPT on tekoälysovellus, joka käyttää luonnollisen kielen prosessointia keskustelun luomiseen. Luonnollisen kielen prosessoinnista käytetään nimitystä NLP (natural language processing). NLP tarkoittaa sitä, että tietokoneohjelmalla on kyky ymmärtää ihmisen puhumaa kieltä. Tällä kielimallilla on kyky vastata kysymyksiin sekä laatia erilaisia kirjallisia sisältöjä, kuten artikkeleita, esseitä, sosiaalisen median julkaisuja tai koodia. (Hetler, 2024.)

ChatGPT toimii siten, että se jakaa tekstisisällön pieniin osiin, joka auttaa sen neuroverkkoa käsittelemään tietoa tehokkaammin. Kun ohjelma saa uuden kysymyksen tai kehotteen, se muuttaa sen ensin merkkijonoksi, eli tekstiksi, jota se pystyy analysoimaan. Analyysin aikana ohjelma etsii tekstistä kuvioita ja suhteita sekä vertaa niitä aikaisemmin harjoitustiedoissa oppimiinsa malleihin. ChatGPT:n toimivuus perustuu muuntaja-arkkitehtuuriin, joka on keskeinen osa monissa luonnollisen kielen käsittelymalleissa. Tämä arkkitehtuuri käyttää erityistä menetelmää nimeltä self-attention, joka auttaa ohjelmaa ymmärtämään, mitkä sanat tai merkit ovat tärkeitä ja miten ne liittyvät toisiinsa. Analyysin jälkeen tulokset yhdistetään ja muokataan siten, että ne tuottavat uuden tekstijakson. Tavoitteena on luoda uusia tekstijaksoja, jotka ovat tilastollisesti samanlaisia kuin ne, jotka malli on jo oppinut, mutta eivät suoraan kopioi harjoitusdataa. Tämä koko prosessi tapahtuu erittäin nopeasti, joten ChatGPT pystyy tuottamaan lyhyitä ja merkityksellisiä vastauksia vain muutamissa millisekunneissa. (Rouse, 2024.)

2.2 Chatbotit Suomessa

Suomessa chatbotteja käyttävät useat yritykset ja organisaatiot. Chatbotit ovat yleistyneet Suomessa erityisesti asiakaspalvelussa. Ne tarjoavat asiakkaille nopeaa apua ja informaatiota monilla eri aloilla, kuten pankki- ja terveystalveissa, sekä verkkokaupoissa. Esimerkiksi Kela käyttää chatbottia, jonka tarkoitus on hoitaa palveluntarjoajinsa nettisivuilla pieniä ja suhteellisen helppoja asiakaspalveluun liittyviä tehtäviä. Suomessa chatbotit ovat ylipäättään vielä melko yksinkertaisia, etukäteen ohjelmoituja kokonaisuuksia, jotka osaavat vastata vain yksinkertaisiin kysymyksiin. (Nieminen, 2022.)

Myös suomen kieli asettaa chatboteille oman haasteensa. Haasteita suomen kielessä ovat esimerkiksi sijamuodot sekä murteet. Koneoppimisen ja tekoälyn kehityksen myötä tulevaisuuden botit tulevat kuitenkin osaamaan suomen kieltä yhä paremmin. Myöskään botti-sanana merkitys ei ole suomen kielessä aivan yksiselitteinen. Sen sijaan käytössä onkin myös muita ilmaisuja, kuten "chattibotti". Robotti-sanana perusteella käyttäjän on helppo ymmärtää, että kyseinen keskustelukumppani ei ole oikea ihminen. (Nieminen, 2022.)

Ensimmäinen suomenkielinen chatbot on todennäköisesti ollut vuonna 2016 Autotien julkaisema Autokauppias. Autokauppias-chatbotin tarkoituksena oli neuvoa asiakkaita auton ostamisessa. (Halmeaho, 2022)

2.3 Satakunnan ammattikorkeakoulun chatbot

Satakunnan ammattikorkeakoulussa (SAMK) otettiin käyttöön chatbot, joka perustuu keskustelevaan tekoälyyn. Termi keskusteleva tekoäly kuvaa chatbot-ohjelmiston takana olevaa teknologiaa, joka kykenee olemaan vuorovaikutuksessa asiakkaan kanssa luonnollisella, ihmismäisellä tavalla. Keskusteleva tekoäly kehittyy paremmaksi vuorovaikutuksesta ihmisen kanssa ja tällä tavalla se voi tarjota entistä älykkäämpiä kokemuksia tulevaisuudessa. Keskusteleva tekoäly on älykäs ohjelmisto, joka käyttää

teknologioita, kuten luonnollisen kielen käsittelyä (NLP) sekä koneoppimista (ML). (Galic, 2024.) Luonnollisen kielen käsittely tarkoittaa sitä, että tietokone kykenee ymmärtämään, tulkitsemaan ja tuottamaan ihmisen kieltä luonnollisella tavalla (Numminen, 2023).

Koneoppiminen on tekoälyn alaluokka, eli osa tekoälyä. Se tarkoittaa tekniikkaa, jossa tietokoneet opetetaan oppimaan itsenäisesti, ilman ihmisen ohjausta. Koneoppiminen on prosessi, jonka tekoäly käy läpi suorittaessaan tehtäviään. Se käyttää oppimiaan tietoja ja kaavioita parantaakseen suoritustaan. Koneoppimisalgoritmit analysoivat suuria tietomääriä ja etsivät näistä tiedoista erilaisia kaavoja, sääntöjä tai suhteita. Saadessaan lisää tietoa, tietokone pystyy oppimaan ja kehittymään. Käytännössä tämä tarkoittaa sitä, että sen kyky ennustaa tai tehdä päätöksiä paranee, jonka ansiosta se voi toimia entistä tarkemmin ja tehokkaammin. (McFarland, 2022.)

Tämä SAMKin chatbot ymmärtää parhaiten englannin kieltä, sillä sen koko järjestelmä on alun perin luotu englanniksi. Teimme lähes kaiken sisällön tähän bottiin sekä suomeksi että englanniksi, sillä ne ovat kieliä, joilla SAMKissa voi opiskella tutkinnon. Botissa ei ole ainakaan vielä käytössä automaattista kääntämistä, joten kieliversiot täytyi luoda itse käsin. Botin tavoitteena on sujuvoittaa erityisesti opiskelijoiden ja hakijoiden palvelua, jonka vuoksi oli tärkeää, että sen kanssa voi kommunikoida suomen ja englannin kielellä. Saman tyyppinen botti on käytössä myös muissa oppilaitoksissa Suomessa; esimerkiksi Hämeen ammattikorkeakoulussa, Kaakkois-suomen ammattikorkeakoulussa sekä Vaasan ammattikorkeakoulussa. Oppilaitokset pystyvät käyttämään muiden aiemmin luomia bottivastauksia ja muokata niitä sopivaksi omalle organisaatiolle, eli kehittämistä ei tarvinnut ainakaan tässä SAMKin projektissa aloittaa nollassa, kun aiemmin luotuja vastauksia oli jo suhteellisen paljon. Asiakkaat käyttävät ja kouluttavat bottia itse, toki palveluntarjoajan avustuksella. Botin kouluttaminen on hyvin helppoa asiakkaan näkökulmasta, sillä kehittäminen ei vaadi esimerkiksi ohjelmointitaitoa, vaan siihen voi osallistua ilman aiempaa kokemusta.

3 PALVELUNTARJOAJA

Tämä Boost.ai:n kehittämä ja Front AI:n jälleenmyymä chatbot valikoitui Satakunnan Ammattikorkeakouluun, sillä se on otettu aiemmin käyttöön muutamassa muussakin Suomalaisessa korkeakoulussa. Chatbotin kehitysalusta on suunniteltu käyttäjäystävälliseksi, mikä tarkoittaa, että sitä on helppo käyttää. Kehitysalusta tarjoaa valmiita aikeita, eli ennakoituja käyttäjien kysymyksiä ja toimintoja, joiden avulla chatbot osaa vastata erilaisiin tilanteisiin. Koska nämä aiheet ovat jo valmiiksi ohjelmoitu ja chatbot on esikoulutettu tekoälymallilla, ei kehitystä tarvitse aloittaa täysin alusta. Tämä tarkoittaa, että botin käyttöönotto on hyvin nopeaa ja helppoa. Chatbot on valmiiksi opetettu vastaamaan tiettyihin kysymyksiin ja suorittamaan tiettyjä toimintoja. Matala osallistumiskynnys mahdollistaa sen, että useampi henkilö voi olla mukana kehitysprosessissa ilman erityistä teknistä osaamista. Tämä tekee botin käyttöönotosta sekä kehittämisestä joustavampaa ja nopeampaa, joka on aina etu organisaatiolle. Suurimmat syyt tämän botin valintaan olivat siis käyttöönoton helppous ja nopeus sekä muiden korkeakoulujen kokemukset.

3.1 Boost.ai

Satakunnan Ammattikorkeakouluun käyttöönotetun chatbotin palveluntarjoajana toimii norjalainen chatbottien teknologiantoimittaja Boost.ai. Boost.ai on perustettu vuonna 2016 ja se on erikoistunut yritystasoisiin keskusteleviin tekoälyihin. Yritys keskittyy tekoälypohjaisten chattibottien ja automaattioratkaisujen kehittämiseen asiakaspalvelun tueksi. Sen tarjoamat työkalut mahdollistavat automaattisen vuorovaikutuksen asiakkaiden kanssa eri kanavissa, kuten verkkosivustoilla ja sovelluksissa. Boost.ai:n tavoitteena on parantaa asiakaskokemusta, vähentää kustannuksia sekä vapauttaa resursseja, jotta ihmiset voisivat keskittyä vaativimpiin työtehtäviin, samalla, kun chatbotit hoitaisivat yksinkertaiset asiakaspalvelun tehtävät. (Boost.ai, n.d.)

Boost.ai on saanut alkunsa, kun muutama kunnianhimoinen norjalainen ohjelmistokehittäjä sai vision automatisoida paikallisen pankin asiakaspalvelutoimintoja. Vuosien varrella tähän ongelmaan kehitetty ratkaisu on kehittynyt alustaksi, joka tänä päivänä automatisoi vuosittain miljoonia vuorovaikutuksellisia toimintoja, sadoille johtaville organisaatioille ympäri maailmaa. Boost.ai toimii muun muassa rahoituspalveluiden, televiestinnän sekä julkisen sektorin parissa. Yrityksen maailmanlaajuinen kattavuus ulottuu Yhdysvaltoihin ja Eurooppaan, ja sillä on toimistot kuudella strategisella markkina-alueella sekä laaja verkosto alan tunnetuimpia brändejä, kumppaneita sekä jälleenmyyjiä. Boost.ai on tällä hetkellä yksi maailman johtavia keskustelelevan tekoälyn yrityksiä. Yrityksellä on tällä hetkellä yli 550 virtuaaliagenttia, eli verkkosivuilla tai sovelluksissa toimivaa chatbottia sekä yli 4000 sertifioitua AI-kouluttajaa. Boost.ai:n palvelut tuottavat vuosittain noin 60 miljoonaa automatisoitua keskustelua. (Boost.ai, n.d.)

3.2 Front AI

Suomalainen ohjelmistoyritys Front AI toimii Boost.ai:n jälleenmyyjänä. Front AI:n erityisosaamista ovat tekoälypohjaiset virtuaaliassistentit ja se jälleenmyy Boost.ai:n omistamaa tekoälyalustaa Suomessa sekä toimii alustaa käyttävien yritysten suorana yhteyshenkilönä Boost.ai:n sijaan. Front AI auttaa asiakkaitaan botin käyttöönotossa ja kehityksessä, sekä tarjoaa asiakkailleen tukipalveluita. Boost.ai ei siis ole suoraan tekemisissä suomalaisten asiakkaidensa kanssa, vaan kaikki yhteistyö yrityksen ja palveluntarjoajan välillä kulkee Front AI:n kautta. (Front AI, n.d.)

Yritys on perustettu vuonna 2019 ja sillä on pääkonttori Helsingissä, sekä konttorit myös Tukholmassa ja Kööpenhaminassa. Front AI:lla on yli 50 asiakasta ja sen palvelut tuottaa yli kaksi miljoonaa automatisoitua keskustelua vuodessa. Suomessa Front AI:n asiakkaita ovat muun muassa useat pankit ja oppilaitokset, Kela sekä Punainen Risti. Front AI:n perustajat uskoivat vahvasti jo vuonna 2019, että keskusteleva tekoäly tulee olemaan yksi tämän vuosikymmenen hienoimmista innovaatioista asiakaspalvelussa. Front AI:n tavoitteena on yksinkertaisesti muuttaa asiakaspalvelua parempaan

suuntaan. Saavuttaakseen tämän, he auttavat asiakkaitaan toteuttamaan asiakaspalvelukanavastrategiaansa käytännössä. Yritys hyödyntää seuraavan sukupolven tekoälyyn perustuvia älykkäitä chatbotteja, jotka kykenevät ymmärtämään eri kieliä ja murteita sekä sietämään kirjoitusvirheitä. Näiden ratkaisujen avulla asiakkaille tarjotaan mahdollisuus automatisoida toistuvia asiakaspalvelutehtäviä ilman ihmisen väliintuloa, mikä tehostaa palveluprosesseja sekä parantaa asiakaskokemusta. (Front AI, n.d.)

Kuten sanottu, Front AI:n tarjoamat ratkaisut keskittyvät asiakaspalvelun ja myynnin tehostamiseen tekoälyn avulla. Yrityksen tavoitteena on yksinkertaisesti auttaa asiakasorganisaatioitaan parantamaan tehokkuuttaan ja asiakastyytyväisyyttään, hyödyntämällä moderneja teknologiaratkaisuja.

Front AI:n keskeisiä piirteitä ja toimintoja ovat:

- **Chatbotit ja virtuaaliagentit:** Yritys kehittää chatbotteja, jotka voivat vastata asiakkaiden kysymyksiin reaaliaikaisesti eri viestintäkanavissa. Ne voivat myös opettaa itsensä käyttäjien vuorovaikutuksen perusteella, mikä taas ajan myötä parantaa niiden antamia vastauksia.
- **Analytiikka ja raportointi:** Järjestelmät tarjoavat syvällistä analytiikkaa asiakaskäyttäytymisestä, mikä auttaa organisaatioita ymmärtämään asiakkaiden tarpeita ja tämän avulla ne pystyvät parantamaan palvelua entisestään.
- **Monikanavainen tuki:** Ratkaisut voivat integroitua eri kanaviin, kuten verkkosivuihin, sosiaaliseen mediaan tai sovelluksiin, mikä mahdollistaa yhtenäisen asiakaskokemuksen.
- **Automaatio:** Tekoäly voi automatisoida toistuvia tehtäviä, kuten asiakaspalvelupyyntöjen käsittelyä, vapauttaen työntekijöiden aikaa vaativampiin tehtäviin. (Front AI, n.d.)

4 PROJEKTIN ALOITUS

Tämä bottiprojekti alkoi siten, että kaikki osallistujat suorittivat itsenäisesti palveluntarjoaja Boost.ai:n tarjoaman *AI Trainer* –sertifikaatin. Sertifikaatin suorittamisen tarkoituksena oli saada projektiin osallistuville käsitys siitä, miten botin käyttöönotto ja kysymysten sekä vastausten luominen tulee tapahtumaan. Sertifikaatin harjoitustehtävät suoritettiin samanlaisella alustalla, jolla bottia kehitetään. Sertifikaatin suorittaminen oli tärkeä vaihe projektissa, sillä sen ansiosta kehitysalusta tuli jo etukäteen tutuksi ja jokainen osallistuja sai selkeämmän käsityksen siitä, miten helppoa tai haastavaa tämän projektin toteuttaminen tulee itselle olemaan.

4.1 Kick-off

Seuraava vaihe sertifikaatin suorittamisen jälkeen oli Front AI:n edustajan pitämä Kick-off, eli projektin aloitustilaisuus. Siellä kävimme läpi projektin aikataulun ja tutustuimme lisää botin kehitysalustaan. Kick-off sujui jouhevasti, sillä kaikilla projektiin osallistuvilla oli jo jokin käsitys kehitysalustasta, koska kaikki olivat suorittaneet etukäteen *AI Trainer* -sertifikaatin. Toki sertifikaattia suorittaessa käytiin läpi niin monta eri osa-aluetta ja vaihetta, että täysin kaikkea ei pystynyt mitenkään omaksumaan. Pääsimme kysymään Front AI:n edustajalta meitä askarruttaneita kysymyksiä ja hän auttoi meidät alkuun projektin kanssa. Front AI:n edustaja oli projektissa tukihenkilönämme koko käyttöönottoprojektin ajan. Hän piti meille koulutuksia säännöllisin väliajoin projektin edetessä ja oli tavoitettavissa viestitse, mikäli tulisi ongelmatilanteita tai jotakin kysyttävää.

4.2 Botin kehitysalusta

Chatbottia kehitettiin palveluntarjoaja Boost.ai:n tarjoamalla alustalla nimeltä Academy Boost.ai. Kaikki aikeet (intents) on järjestetty alustalla aiehierarkiaosioon, joka löytyy alustan hallintapaneelin vasemmasta reunasta.

Aiehierarkia on havainnollistettu kuvassa 2. ja aikeista tulen kertomaan tarkemmin kappaleessa 5.1.

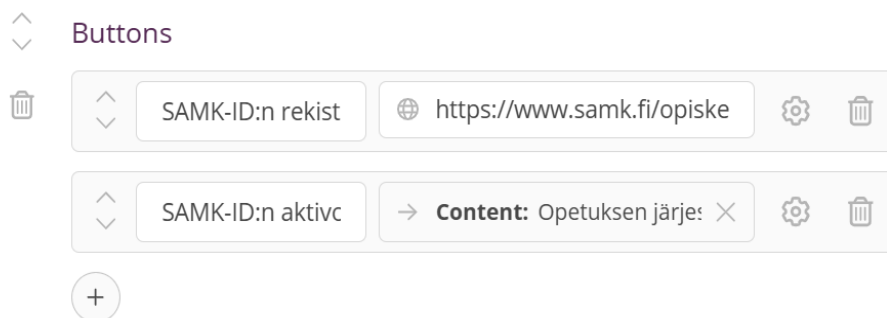


Kuva 2. Aiehierarkia (Boost.ai)

Alustalla on helppo tehdä vahingossa virheitä, sillä muiden alustaa käyttävien oppilaitosten sisällöt ovat kaikkien saatavilla ja niitä pystyy muokkaamaan. Muiden sisältöihin ei kuitenkaan saa tehdä muutoksia. Myöskään uusia aikeita ei voida luoda ilman muiden oppilaitosten lupaa, sillä aiheet ovat oppilaitosten yhteisiä ja niiden lisäämisestä tulee neuvotella yhdessä muiden kanssa.

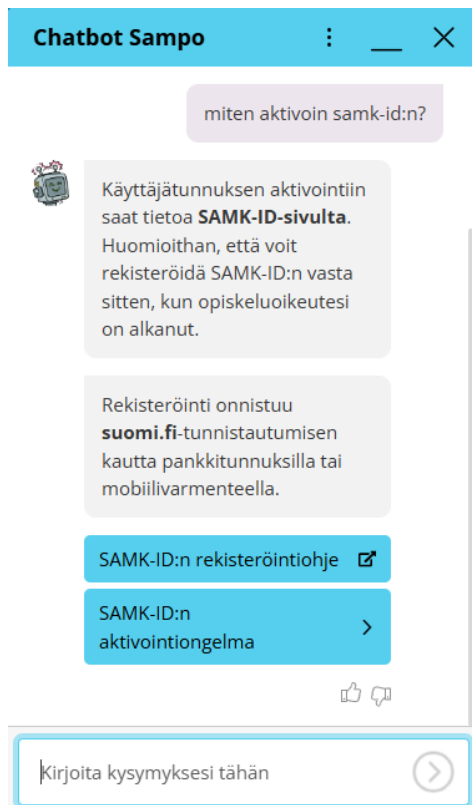
Kehitysalustalla pystyi vastauksia kirjoittaessa katsomaan, miltä teksti näyttäisi chatissa. Tällä tavoin oli helppo skaalata esimerkiksi tekstin pituutta, sillä liian pitkä teksti ei toimi bottivastauksissa. Linkkien nimet eivät myöskään voi olla liian pitkiä, vaan koko nimi tulee näkyä käyttäjälle, joten chatin esikatselutoiminto oli tässäkin kohtaa tarpeellinen. Ennen botin julkaisua pystyimme testaamaan luomiemme sisältöjen toimivuutta erillisessä testichatissa. Testichatilla oli helppo kokeilla, miten botti osaa vastata, jos kysyy esimerkiksi samaa asiaa eri tavoin tai laittaa kysymykseen kirjoitusvirheitä.

Jokainen sisältökomponentti on itsenäinen olio. Tällä tarkoitetaan sitä, että kaikki sisältö koostuu erillisistä ja itsenäisistä palasista, jotka voivat olla osa suurempaa kokonaisuutta ilman, että niitä on sidottu muihin osiin. Tällöin komponentit voivat olla käytettävissä sekä hallittavissa itsenäisesti ilman, että ne vaikuttavat liikaa muihin järjestelmän osiin. Kehitysalustalla voidaan linkittää sisältökomponentista toiseen painikkeilla (buttons) ja tämä on helppo tehdä hyödyntämällä sisältöjen omia ID:itä. Jokaisella sisällöllä on oma ID-numero, joka voidaan liittää painikkeisiin, kun halutaan, että käyttäjä voi siirtyä chatissa sisällöstä toiseen. Painikkeisiin voidaan lisätä myös linkkejä ulkoisille verkkosivuille. Kuvassa 3. on havainnollistettu, miltä nämä näyttävät kehitysalustalla. Ylimpänä oleva painike, jossa on linkin edessä selain-kuvake, tarkoittaa, että valitsemalla tämän painikkeen siirrytään ulkoiselle verkkosivulle, pois chatista. Alempi "content"-painike taas tarkoittaa sitä, että valitsemalla kyseisen painikkeen, siirrytään sisältökomponentista toiseen ja käyttäjä saa chatiin lisää vastauksia tekstinä.



Kuva 3. Painikkeet (Boost.ai)

Kuvassa 4. näkyy, miten erilaiset painikkeet näyttäytyvät kysyjälle keskustelussa botin kanssa. Ylemmässä painikkeessa "SAMK-ID:n rekisteröintiohje" -tekstin perässä on kuvake, joka viittaa sisällön avautuvan toiselle sivulle. Alemmassa painikkeessa "SAMK-ID:n aktivointiongelma" -tekstin perässä on nuoli, joka viittaa siihen, että valitsemalla tämän painikkeen, botti syöttää lisää vastauksia samaan keskusteluun.



Kuva 4. Keskustelunäkymä käyttäjälle (samk.fi)

4.3 Edistymisen seuranta

Seurasimme projektin edistymistä yhteisen Excel-taulukon avulla. Excel-taulukkoon oli merkittynä alustalla valmiina olevat aiheet ja näistä lähdettiin sitten valitsemaan, kuka osallistujista hoitaa minkäkin osuuden. Exceliin merkittiin oma nimi sen aiheen kohdalle, jonka aikoo toteuttaa. Työskentelin itse projektin aikana SAMKin hakijapalveluissa, joten luonnollisesti lähdin tekemään hakijapalveluihin liittyviä sisältöjä. Merkitsimme Exceliin oman aikeemme kohdalle edistymisprosentit siten, että 50 % tarkoitti aiheen sisällön olevan valmis toisella kielellä (suomi tai englanti) ja 100 % tarkoitti aiheen sisältöjen olevan valmis molemmilla kielillä. Tällä tavoin kaikkien oli selkeä seurata projektin edistymistä sekä pystyimme pitämään huolen siitä, että vain yksi ihminen lähti toteuttamaan tiettyä aietta. Front AI:n edustaja, joka toimi tukihenkilönämme, pääsi myös seuraamaan Excel-taulukosta projektin edistymistä ja pystyi tällä tavoin olemaan perillä siitä, missä vaiheessa

projektissa mennään. Tähän samaan seuranta Exceeliin lisättiin koulukohtaiset termit, jotka tulee lisätä synonyymeihin, sekä uudet aie-ehdotukset.

5 KÄYTTÖÖNOTTO

Ennen chatbotin käyttöönottoa organisaatiossa, on tärkeää miettiä ainakin seuraavia keskeisiä asioita:

- **Yleisimmät kysymykset:** Mitkä kysymykset toistuvat organisaation eri osastoilla usein, millaisiin kysymyksiin vastataan lähes päivittäin?
- **Selkeät vastaukset:** Onko näihin yleisimpiin kysymyksiin olemassa niin selkeitä ja yksinkertaisia vastauksia, että niihin vastaaminen voitaisiin luottaa chatbotin hoidettavaksi?
- **Asiakaskokemus:** Miten botti voisi parantaa asiakaskokemusta? Voisiko botti helpottaa ja nopeuttaa palvelua, esimerkiksi tarjoamalla nopeasti yksinkertaisia vastauksia tai ohjaamalla asiakkaan sivulle, josta asiakas löytää haluamansa informaation helposti?

Näiden kysymysten pohtiminen auttaa varmistamaan, että chatbot toimii mahdollisimman tehokkaasti ja tuo lisäarvoa sekä organisaatiolle että asiakkaille.

5.1 Aikeet

Projektissa oli alkuun valittuna tietyt aikeet, joihin luotiin vastaukset, kysymykset ja synonyymit. Aikeet tarkoittavat erilaisia aiheita, joista käyttäjillä saattaa olla kysymyksiä ja joihin chatbotin halutaan vastaavan. Kaikki aikeet löytyvät kehitysalustalla vasemmassa reunassa olevasta aiehierarkiasta. Aiehierarkiassa aikeet ovat niin sanotusti pääotsikkoja, joiden alle lisätään sisältöä (content). Jokaisella aikeella voi olla koulutusdataa, joka edustaa erilaisia tapoja, joilla käyttäjä saattaa kysyä aikeeseen liittyen. Tarkennettu malli käyttää koulutusdataa auttaakseen virtuaalista avustajaa, eli chatbottia,

ymmärtämään, mitä käyttäjä kysyy sekä tekemään ennusteen. Tarkennetulla mallilla tarkoitetaan chatbottia, joka on rakennettu käyttämällä mallia, joka ei ole erikoistunut mihinkään tiettyyn tehtävään, mutta jota on tarkennettu tietyllä koulutusdatalla. Few-shot-malli voi myös käyttää koulutusdataa ymmärtääkseen, mitä käyttäjä kysyy, mutta se voi sen lisäksi tehdä ennustuksia pelkästään aikeen nimen perusteella, jos nimi on tarpeeksi kuvaileva. (Boost.ai) Few-shot-malli tarkoittaa koneoppimiskehystä, jossa tekoälymalli oppii tekemään tarkkoja ennusteita harjoittelemalla pientä määrää merkittyjä esimerkkejä, ja sitä käytetään tyypillisesti mallien kouluttamiseen sekä luokittelutehtäviin, kun sopivaa harjoitustietoa on vähän (IBM n.d.).

Tässä projektissa oli valmiina kuuden muun korkeakoulun luomia aikeita ja niiden sisältöjä, joten meidän ei tarvinnut luoda kaikkea alusta asti itse. Toki muokkasimme aikeiden sisällöt omalle organisaatiollemme sopiviksi. Teimme lähes kaikki aikeet sekä suomeksi että englanniksi. Teimme molemmat kieliversiot käsin, sillä botissa ei ole ainakaan vielä käytössä automaattista kääntämistä. Tähän käyttöönottoprojektiin oli valittu 150 aietta, jotka piti luoda botin julkaisuun mennessä. Nämä ensimmäiset aikeet oli valittu sen mukaan, mitkä aihealueet olisivat opiskelijan tai hakijan kannalta tärkeimmät ja kysytyimmät. Käyttöönotto oli niin nopea prosessi, ettei kaikkea mahdollista sisältöä olisi voitu luoda kerralla, joten tämä rajaus oli ainakin omasta mielestäni erittäin toimiva. Kuten aiemmin mainittu, meitä botin kehittäjiä helpotti jo valmiina olevat aikeet, joita muissa korkeakouluissa on luotu ja käytetty. Aikeita pystyi vaivattomasti kopioimaan ja muokkaamaan omalle oppilaitokselle sopiviksi. Tulevaisuudessa bottiin tullaan lisäämään tietysti enemmän aikeita ja sisältöjä. Botti kerää dataa siitä, mitä siltä kysytään ja tämän datan perusteella voidaan luoda uusia aikeita ja sisältöjä, jotta botti palvelisi sen käyttäjäkuntaa mahdollisimman laajasti ja tehokkaasti.

5.2 Synonyymit

Projektin alussa koulukohtaiset termit kirjattiin aiemmin mainittuun seuranta Exceliin, omalle termit-välilehdelle. Tämän taulukon perusteella pystyttiin

tarkastelemaan, onko paneelissa jo olemassa sana omassa organisaatiossa käytetyille termeille tai opetetaanko synonyymejä uusien aikeiden luomisen yhteydessä oman oppilaitoksen projektin ulkopuolella. Hallintapaneelissa koulukohtaiset termit lisätään jo valmiiksi olemassa oleviin vastaaviin aikeisiin, paneelin Language processing -puolelle. Tämä tehtiin yhdessä Front AI:n pitämällä tukitunnilla. Front AI:n tukihenkilö avusti meitä, jos lisättävillä synonyymeillä oli ristiriitoja olemassa olevien synonyymien kanssa. Uusina opetettavat aiheet ja niiden synonyymit tullaan opettamaan mallille vasta tämän käyttöönottoprojektin jälkeen.

Kielenkäsittely (language processing) yksinkertaistaa kieltä ryhmittelemällä sanoja, jotka tarkoittavat samaa asiaa. Käsittelyn jälkeen virtuaalinen assistentti, eli botti korvaa kaikki synonyymiryhmän sanat yhdellä niin sanotulla näyttösanalla. Sellaiset sanat, jotka eivät lisää merkitystä käyttäjän viestiin, poistetaan. Ennen uuden synonyymiryhmän luomista tulee aina ensimmäiseksi etsiä vastaavia termejä paneelin language processing -puolelta. Esimerkiksi jonkin oppimisympäristön nimi voi olla jo valmiiksi olemassa synonyymikirjastossa ja omassa oppilaitoksessa käytettävä nimi lisätään tähän samaan synonyymiin. Tämä tarkoittaa sisällönrakentamisen kannalta sitä, että olemassa olevaa aietta päästään hyödyntämään, eikä uutta oppimisympäristöä koskevaa aietta tarvitse kouluttaa alusta asti. Puuttuvia synonyymejä voi lisätä myös jälkikäteen, mutta helpompaa on kuitenkin tarkastella etukäteen, millaisia vaikutuksia esimerkiksi uutena luotavan aikeen treenidatalla olisi nykymallin kannalta. On myös tärkeää tarkastella, miten olemassa olevissa aikeissa on opetettu yhdyssanoja. Esimerkiksi kahden tai kolmen sanan pituiset yhdyssanat eivät esiinny synonyymikirjastossa kokonaisina termeinä, vaan yhdyssana stemmautuu sana sanalta aikeen treenidatan joukossa. Stemmautuminen tarkoittaa prosessia, jossa sana pelkistetään sen juurimuotoon. Kuvassa 5. on esimerkki sanojen stemmautumisesta. Juurimuoto ei välttämättä ole sana sinänsä, mutta sitä voidaan käyttää sanojen luomiseen yhdistämällä oikea päätte. Vain aidot termit opetetaan omana synonyyminä. Koska yhteismallissa kaikki vaikuttaa kaikkeen, tulee ensisijaisesti tutustua aikeiden tärkeisiin sanoihin ja tarkastella prosessoitua muotoa siitä lauseesta, jonka halutaan ennustuvan tiettyyn

aikeeseen. Tulee selvittää, miten se eroaa aikeen nykyisestä treenidatasta ja lauseiden prosessoiduista muodoista. Onko jo haettu tärkeäksi koetuilla sanoilla language processing -puolelta sekä haettu paneelin aikeiden nimistä vastaavuuksia etsitylle asialle? Vai onko sanojen/yhdyssanojen prosessoitu muoto niin sanotusti out of scope, eli kyseisen lauseen ei kuuluisikaan ennustua nykyisiin aikeisiin, koska vastaavaa aietta ei ole olemassa.

The screenshot shows a user interface for 'Language processing'. At the top, there is a search bar with the text 'Language processing' and a question mark icon. To the right of the search bar is a button labeled '+ New synonym group'. Below the search bar, there is a table with two columns: 'Display word' and 'Synonyms'. The table contains three rows of data:

Display word	Synonyms
ict problem	it issu (EN) it problem (EN) ict problem (EN)
ict support	it tue (FI) it tuk (FI) it tuke (FI) it tuki (FI) atk tuki (FI) +19 More
ict support request	tiket (FI) työpyyntö (FI) tukipyyntö (FI) palvelupyyntö (FI)

Kuva 5. Esimerkki synonyymeista ja sanojen stemmautumisesta Language processing -osiossa.

Jos aikeisiin halutaan lisätä treenidataa tai testidataa, tulee tutustua treenikierroksen jälkeen myös botin kehitysalustalla olevan analytiikkasivun *High-priority clean-up: Non synonyms* -sivuun. Jos sana ei prosessoidu oikein tai se näkyy suomeksi, se tulee opettaa language processing -puolelle. Poikkeuksena ovat ihmisten nimet tai mallin kannalta turhat sanat, joita ei opeteta. Oikein prosessoituvat synonyymit voidaan hyväksyä. Ennen hyväksymistä on hyvä tarkistaa, että non synonyms -listalla esiintyvä stemmattu (tai tarvittaessa exact match) muoto esiintyy myös language processing -synonyymikirjastossa. Se tulee lisätä sopivan *display wordin* alle, jos muotoa ei ole vielä opetettu mallille.

Synonyymia ei tarvitse etsiä sen "kokonaisvartalon" kanssa, vaan se tulee etsiä lyhyempänä eli stemmattuna versiona sanasta. AI trainereiden eli botin kouluttajien tulee etsiä sanoja vain synonyymien alkupään perusteella ja leikata synonyymien vartalot stemmerin mukaisesti uutta sanaa lisätessä. Jos synonyymi menee stemmatussa muodossaan sekaisin jonkin toisen synonyymien kanssa, se tulee opettaa stemmauksen sijaan muodossa *Exact*

match. Stemmeri ilmoittaa käyttäjälle, jos stemmattu muoto sanasta on jo olemassa. Jos botin käyttäjät kirjoittavat keskusteluissa ristiin esimerkiksi suomea ja englantia, englanninkielinen sana saattaa tarvita myös suomenkielisen stemmauksen lisäyksen tai toisinpäin. Tämä lisätään language processing -puolella. Jos lisätään sisältöä tai testi-/treenidataa muilla kielillä kuin suomeksi, se tulee tehdä oikean kielen puolella paneelissa. Kieltä pääsee vaihtamaan sivun vasemmasta yläaidasta. On tärkeää muistaa palata takaisin suomen kielen puolelle, kun halutaan taas lisätä suomen kielistä sisältöä tai testi-/treenidataa aikeeseen, tai halutaan muuten vaan tarkastella suomenkielistä puolta mallista. Jos aikeissa esiintyvistä tai aikeisiin lisättävistä synonyymeistä on epävarmuutta, on hyvä testata tärkeitä synonyymejä tai sanoja sisältäviä lauseita testichatin puolella. Lause kannattaa kokeilla kirjoittaa muutamalla eri tavalla ja kokeilla, miten botti vastaa niihin.

5.3 Optimointi

Kirjautuessa kehitysalustalle, aukeaa statussivu, joka näyttää selkeän tehtävälistan, jota botin kehittäjän tulee noudattaa. Useat statussivulla olevat indikaattorit näyttävät, mitä asioita tulisi korjata parantaakseen chatbottia. Jokainen paneeli näyttää tiedot viimeisen seitsemän päivän ajalta, ellei toisin mainita. Mallin optimointiin käytetään alustalla *high priority clean-up* -osiota. Optimointi tehdään jokaisen treenikierroksen jälkeen, erityisesti silloin, kun mallille opetetaan uusia aikeita. Optimointi kuuluu niiden AI Trainereiden eli botin kehittäjien työhön, joilla on näkyvyys sekä paneelin treeni- että testidataan sekä *high priority clean-up* -osioon. High priority clean-up auttaa huoltamaan mallia erityisesti uusia aikeita treenatessa ja jokaisen treenikierroksen jälkeen. Osiot, jotka paneelista kannattaa erityisesti priorisoida, ovat *Non synonyms* ja *Training data issues*. Klikkaamalla osiot auki paneelissa, pääsee tarkistamaan, millaisia korjauksia tarvitsee tehdä.

Non synonyms -osio tarkoittaa nimensä mukaisesti niitä synonyymejä, joita botti ei vielä osaa. Jokaisen treenikierroksen jälkeen on suositeltavaa tarkistaa *non synonyms* -lista. Onko listalla paljon synonyymejä, joita ei ole opetettu tai

onko siellä sellaisia synonyymeja, jotka saattavat prosessoitua väärin jonkin toisen synonyymien yhteydessä? Tällaiset synonyymit siivotaan pois opettamalla ne *language processing* -puolelle tai vaihtoehtoisesti hyväksytään ne, jos kyseessä on esimerkiksi henkilön tai paikan nimi, jota ei ole tarkoituskaan opettaa. Avaamalla *High priority clean up: Non synonyms* -listan, näkee todellisen tilanteen synonyymien prosessoitumisesta. Joskus, vaikka lukumäärä näyttäisi nolaa, saattaa sieltä löytyä muutama opettamaton synonyymi. Tämä osio kannattaa siis aina avata ja käydä tarkistamassa sisältö. Oikealla puolella sivua näkyy datan tyyppi; chattidata, treenidata ja testidata. Näistä on hyvä tarkistaa erityisesti treenidatan prosessoituminen, koska sen kuuluu prosessoitua oikein, jotta botit voi kehittyä.

Tyypillisiä korjattavia asioita *training data issues* -osiossa ovat erityisesti duplikaatit. Duplikaatti tarkoittaa sitä, että sama treenilause esiintyy kahdessa eri aikeessa, eli prosessoitu muoto lauseelle on sama kahdessa eri aikeessa. Sama lause saa sijaita vain yhdessä aikeessa, joten tällaisissa tapauksissa duplikaatti tulee poistaa siitä aikeesta, joka on lauseelle epäsovelias. Tärkeiden sanojen kuuluu erota toisistaan aikeessa. Jos ne eivät eroa toisistaan tai tärkeitä sanoja esiintyy aikeessa useita kappaleita, aikeen treenidata ei välttämättä ole tarpeeksi laadukasta. Tärkeiden sanojen määrä aikeessa riippuu mallista, merkittävintä on tärkeäksi sanaksi nousseen sanan olennaisuus juuri siinä aikeessa. Turhia täytesanoja tulee karsia pois treenilauseista, jos niitä nousee tärkeiden sanojen joukkoon treenikerroksen jälkeen. Treenidata tulee pitää puhtaana turhista synonyymeista ja epäolennaisilla synonyymeilla varustettu lause saattaa kuulua enemmän testidatan puolelle. Treenilause saattaa ennustua väärään aikeeseen myös lauserakenteensa takia. Siitä kannattaa tarkastella aina vain prosessoitua muotoa, sillä se paljastaa esimerkiksi lauseen synonyymit, mahdollisesti väärin merkatun kielen tai opettamattomat synonyymit.

6 BOTIN JULKAISU

Chatbot julkaistiin Satakunnan Ammattikorkeakoulun verkkosivuilla 11. kesäkuuta 2024 ja se sai nimekseen Sampo. Nimi päätettiin henkilökunnalle järjestetyn nimikilpailun perusteella. Chatbot on suunniteltu palvelemaan ensisijaisesti SAMKin opiskelijoita sekä SAMKiin opiskelijaksi hakevia. Chatbotin tarkoituksena on tarjota kysyjälle reaaliaikaista tukea ja antaa vastauksia yleisimpiin ja yksinkertaisimpiin kysymyksiin, liittyen esimerkiksi hakemiseen tai opintojen etenemiseen. Tämän käyttöönoton tavoitteena on tarjota jatkossa sujuvampi asiointi ja nopeuttaa tiedonsaantia.

Kesäkuun 2024 ja lokakuun 2024 välillä botti on käynyt yhteensä 1127 erillistä keskustelua. Keskusteluista 435 on käyty englanniksi ja 756 suomeksi. Yhteensä näistä tulee 1191 keskustelua, mikä tarkoittaa sitä, että osassa keskusteluista on käytetty sekä suomea että englantia. Botin käyttötilastot kuukausittain ovat seuraavat: kesäkuu 164 keskustelua, heinäkuu 183 keskustelua, elokuu 311 keskustelua, syyskuu 260 keskustelua ja lokakuu 209 keskustelua. Vähiten bottia on käytetty siis kesäkuussa, joka oli sen julkaisukuukausi, ja eniten käyttöä on ollut elokuussa.

Mitä näistä luvuista voi sitten päätellä? Botti julkaistiin vähän ennen kesäkuun puoliväliä, jolloin on ihan ymmärrettävää, että bottia on käytetty silloin vähiten, kun lukuja tarkistellaan kuukausitasolla. Heinäkuussa sitä on käytetty jo vähän enemmän, mutta varmasti koulujen kesäloma on vaikuttanut siihen, että luku ei ole noussut kovin paljoa. Toki heinäkuussa on tehty opiskelijavalintoja, joka on varmasti tuonut jonkin verran liikennettä SAMKin verkkosivuille. Elokuussa keskustelujen määrän nousu on kuitenkin huomattava. Elokuun lopulla uudet opiskelijat aloittivat koulun, mikä on ymmärrettävästi lisännyt vielä enemmän SAMKin verkkosivuilla käyneiden ihmisten määrää ja tämän myötä myös bottia hyödyntäneiden määrää. Syyskuun alussa jatkavat opiskelijat aloittivat syyslukukauden, ja kysymysten määrä on tässä kohtaa ollut todennäköisesti vielä aika korkealla. Syyskuun lopulla - lokakuussa botti on käynyt keskusteluja ihan kiitettävän määrän ja määrä on pysynyt aika tasaisena, toki hieman

laskenut elo- syyskuuhun nähden. On positiivista, että botti on löydetty ja sitä on hyödynnetty. Sitä on tässä kohtaa vielä vaikea sanoa, onko botti todellisuudessa vähentänyt puhelimitse tai sähköpostitse tullutta kysymysmäärää. Bottia kehitetään, jotta se osaisi vastata vielä useamman aihealueen kysymyksiin ja se pystyisi vastaamaan käyttäjälle mahdollisimman informatiivisesti. Tulevaisuudessa nähdään, toimiiko tämä oikeasti osana sujuvaa asiakaspalvelua.

7 HAASTEET PROJEKTISSA

Kuten missä tahansa projektissa, tässäkin oli omat haasteensa. Omasta näkökulmastani projektin yksi suurimmista haasteista oli aikataulu. Aikataulu tuntui paikoittain hyvin nopealta, sillä bottia piti ehtiä kehittämään muiden töiden ohella ja välivaiheita projektin aloituksesta botin julkaisuun oli yllättävän useita.

Alkuun kehitysalustan ymmärtäminen toi myös hieman haasteita, mutta kun sitä oppi käyttämään, alkoi kysymysten ja vastausten luominen sujua tehokkaammin. Alussa suoritettu sertifiointi oli hyödyllinen, sillä sitä suorittaessa oppi hieman käyttämään kehitysalustaa, joka helpotti omaa tekemistä tässä projektissa.

7.1 Haasteet botin vastausten luomisessa

Yksi suuri haaste oli se, että Bottiin ei voitu luoda vastauksia, joissa olisi kerrottu esimerkiksi yhteystietoja tai muuta sisältöä, joka saattaa muuttua. Tällaisessa tilanteessa vastaukseen piti lisätä linkki sivulle, josta nämä tiedot löytyvät. Tämä tehtiin näin sen vuoksi, että tällaisia muuttuvia tietoja ei todennäköisesti muisteta korjata bottiin, vaan ne muokataan vain verkkosivuille. Tämä toi omanlaisensa haasteen botin vastauksia miettiessä. Jokaista vastausta luodessa täytyi pohtia, onko tämä tieto niin pysyvä, että sen

voi kirjoittaa suoraan botin vastaukseksi, vai pitääkö tätä varten olla linkki tiedonlähteeseen.

Toinen haaste botin vastauksia luodessa oli tietysti myös se, että botin vastaukset tulisivat olla mahdollisimman lyhyitä ja yksinkertaisia. Botin-ikkuna on verkkosivuilla aika pieni ja vastaus tulisi näkyä kokonaan siinä ikkunassa. Toki aiemmin mainittu haaste siitä, että bottiin piti linkittää muuttuvaa informaatiota sisältäviä sivuja, auttoi tässä kohtaa. Botin ei mitenkään pystyisi antamaan kaikkiin kysymyksiin tarpeeksi selkeitä ja lyhyitä, mutta kattavia vastauksia ilman linkkejä sivuille, josta löytyy kysymykseen laajempi vastaus tai ohjeistus.

8 YHTEENVETO

Tämä chatbot-projekti oli kokonaisuudessaan mielenkiintoinen ja koen, että pääsin oppimaan paljon sellaista asiaa, jota tuskin pääsisin opettelemaan missään muussa yhteydessä. Aikataulullisesti projekti oli hyvin nopea ja välillä tahti tuntui vähän liiankin ripeältä. Chatbotti saatiin kuitenkin julkaistua tavoiteajassa ja sellaisena, kuin alun perin oli suunniteltukin. Aikeet, jotka oli suunniteltu toteutettavan tämän käyttöönoton aikana, saatiin kaikki luotua julkaisuun mennessä.

Eniten itseäni mietityttää chatbotin kannalta se, kuinka kauan tällainen botti pysyy relevanttina. Voiko täysin tekoälyä käyttävä botti korvata tämän nykyisen, ja kuinka pian? Tämän botin hyvä puoli on toki se, että olemme itse päässeet luomaan kysymyksiä ja vastauksia sellaisiin aihealueisiin, jotka koskevat erityisesti Satakunnan ammattikorkeakoulun opiskelijoita, sekä tänne opiskelijaksi hakevia. Osaisiko tekoäly tuottaa tarpeeksi tarkkoja vastauksia, jotka palvelisivat tätä tiettyä kohderyhmää?

Jatkoin chatbotin kehittämistä työni ohessa myös syksyllä 2024. Sain botin pääkäyttäjältä tietooni kysymyksiä, joihin botti ei ollut osannut vielä vastata ja lähdin niiden perusteella luomaan botille uusia vastauksia. Bottiin ei oltu luotu käyttöönottovaiheessa oman yksikköni, eli SAMKin opetuspalvelujen tai eSAMK tuen osalta ollenkaan sisältöjä, joten lähdin työstämään niitä, jotta botti osaisi vastata entistä laajemmin eri aiheisiin liittyviin kysymyksiin. Kysymyksiä tulee eSAMKiin päivittäin paljon ja suuri osa niistä on saman tyyppisiä tai samoihin aihealueisiin liittyviä, joten on mielestäni tärkeää, että näitä vastauksia pystyttiin antamaan botille. Toivon, että tulevaisuudessa botti helpottaisi työmäärää eSAMK tuessa oikeasti, ja yksinkertaisimmat sekä päivittäin toistuvat kysymykset, esimerkiksi järjestelmiin kirjautumiseen liittyen, osattaisiin kysyä siltä.

LÄHTEET

AWS Amazon. What is a GPU <https://aws.amazon.com/what-is/gpu/> Viitattu 25.10.2024

Boost AI. Haettu osoitteesta <https://boost.ai/>

Dey A. 2019. Story of Cleverbot and Rollo Carpenter. Haettu 7.10.2024 osoitteesta <https://www.yourtechstory.com/2019/11/13/success-story-of-cleverbot/>

Front AI. Haettu 4.10.2024 osoitteesta <https://front.ai/about-front-ai/>

Galic D. 2024 What is conversational AI? Haettu 3.10.2024 osoitteesta <https://www.zendesk.com/blog/customers-really-feel-conversational-ai/>

Gandhi M. 2024 Eliza: The pioneering chatbot from the 1960s. Haettu 4.10.2024 osoitteesta <https://medium.com/@meetgandhi586/eliza-the-pioneering-chatbot-from-the-1960s-b19e7b90fd66>

Gisonna N. 2024. Chatbot. Haettu 29.9.2024 osoitteesta <https://www.britannica.com/topic/chatbot>

Halmeaho V. 2022. Chatbot FAQ – kaikki mitä sinun tulee tietää chatboteista. Haettu 14.10.2024 osoitteesta <https://www.salesforce.com/fi/blog/2018/chatbot-usein-kysytyt-kysymykset.html>

Hetler A. 2024. What is ChatGPT? Haettu 27.9.2024 osoitteesta <https://www.techtarget.com/whatis/definition/ChatGPT>

IBM. What is few-shot learning? Haettu 7.11.2024 osoitteesta <https://www.ibm.com/topics/few-shot-learning>

Lindström S. 2020. Tekoälytutkimuksen klassikko Turingin testi mittaa tietokoneen älykkyyttä – sen kehittäjä Alan Turing koki synkän kohtalon.

Haettu 20.9.2024 osoitteesta

<https://www.itewiki.fi/blog/2020/08/tekoalytutkimuksen-klassikko-turingin-testi-mittaa-tietokoneen-alykkyytta-sen-kehittaja-alan-turing-koki-synkan-kohtalon/>

McFarland A. 2022. Koneoppiminen vs tekoäly: keskeiset erot. Haettu

3.10.2024 osoitteesta <https://www.unite.ai/fi/machine-learning-vs-artificial-intelligence-key-differences/>

Nieminen K. 2022. Mikä on chatbot. Haettu 9.9.2024 osoitteesta

<https://markkinointitrendit.fi/mika-on-chatbot/>

Numminen L. 2023. Mitä on luonnollisten kielten käsittely, eli NLP? Haettu

3.10.2024 osoitteesta <https://www.finnishup.com/mita-on-luonnollisten-kielten-kasittely-eli-nlp/>

Rouse M. 2024. ChatGPT. Haettu 29.9.2024 osoitteesta

<https://www.techopedia.com/fi/sanasto/chatgpt>

Sachdeva N. 2024. The history and evolution of chatbots. Haettu 27.9.2024

osoitteesta <https://insights.daffodilsw.com/blog/the-history-and-evolution-of-chatbots>

Tyyskä I. 2022. Parhaat chatbotit. Haettu 19.9.2024 osoitteesta

<https://www.raikasdigital.fi/parhaat-chatbotit/>

Visualwebz 2024. What is a turing test. Haettu 27.9.2024 osoitteesta

<https://seattlewebsitedesign.medium.com/what-is-a-turing-test-d379477fed9b>