

PedAg-projekti

Kartoitusraportti 1.0

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Koulutusteknologia

Saara Huhta

Viivi Laitinen

Juho Liedes

Maija Turpeinen

Tiivistelmä:

Tämä dokumentti on yksi Jyväskylän yliopiston informaatioteknologian tiedekunnan Koulutusteknologian projekti -kurssilla toteutettavan PedAg-projektin tuotoksista. Raportissa kartoitetaan keskustelevien agenttien ja niiden kehityksessä käytettyjen alustojen nykytilannetta. Raportti käsittelee alustojen sekä agenttien teknisiä ominaisuuksia, agenttien kouluttamista, tulevaisuuden teknologioita sekä sitä, miten alustoja hankitaan markkinoilta. Kartoitusraportti sisältää listauksen kehitysalustoista, perusteluita alustan valikoitumisesta tarkempaan tutustumiseen tai puolestaan sen hylkäykseen sekä ehdotuksen projektin arviointimatriisissa käsiteltävistä alustoista.

Versio	Päivämäärä	Tekijät	Muutokset
0.1	20.2.2023	Huhta, Laitinen, Liedes & Turpeinen	Ensimmäinen kommentoitava luonnos
0.2	1.3.2023	Huhta, Laitinen, Liedes & Turpeinen	Ensimmäinen projektikokouksessa käsiteltävä versio
0.3	10.3.2023	Huhta, Laitinen, Liedes & Turpeinen	Kolmas kommentoitava versio, muutoksia ja lisäyksiä palautteen pohjalta
0.4	17.3.2023	Huhta & Turpeinen	Ensimmäinen kokouksessa hyväksyttävä versio
0.5	24.3.2023	Huhta	Muutokset: 2.1: syväoppimisen käsite, 2.7: Moodle-integraatioiden lähteet, 4.2: johdantovirkkeet alustoista, Liite 1: Microsoft Bot Frameworkin hylkäysperustelut, ChatGPT:n muutosarake
1.0	31.3.2023	Turpeinen	Dokumentin muuntaminen pdf-muotoon. Ei muutoksia sisältöön.

Taulukko 1. Versiohistoria

Sisällysluettelo

1	Johdanto	5
2	Teknologiat agenttien taustalla	6
2.1	Käsitteitä	6
2.2	Tekoäly	7
2.3	Agentin ilmenemistavat	9
2.4	Käyttöönotto	10
2.5	Erilaiset alustat	10
2.6	Kouluttaminen	12
2.7	Integroitavuus	14
2.8	Jyväskylän yliopiston Minerva-chatbot	15
2.9	Tulevaisuuden teknologiat	16
3	Teknologioiden hankinta	18
4	Johtopäätökset	19
4.1	Kartoitettujen alustojen yhteenveto	19
4.2	Esimerkkejä alustoista	20
4.2.1	Rasa	20
4.2.2	Botpress	21
	Lähteet	23
	Liitteet	26

1 Johdanto

Tämän raportin tarkoituksena on luoda ajantasainen kartoitus pedagogisten agenttien tilanteesta. Raportin tavoitteena on kartoittaa, millaisia agenteja sekä agenttien kehitysalustoja on olemassa, ja erityisesti sitä, mitä teknologioita on käytetty. Raportin avulla luodaan eri alustoista arviointimatriisi, jota hyödynnetään projektin pilotointivaiheessa käytettävän alustan valikoinnissa. Kartoituksen painopisteenä ovat agenttien kehittämisen mahdollistavat alustat ja niiden ominaisuuksien tarkastelu sen sijaan, että keskityttäisiin ainoastaan valmiisiin agenteihin. Kartoituksessa on käytetty JYKDOK-, Google Scholar- ja Gartner-tietokantoja tieteellisen materiaalin hankintaan sekä Googlen hakukonetta kaupallisten ja muiden lähteiden osalta. Kartoitusta tehdessä painotus on ollut uusimmissa ja ajankohtaisimmissa lähteissä, ja merkittävä osa löydetyistä materiaalista on tullut muualta kuin kasvatusalalta.

Aiemman tutkimuksen pohjalta on todettu, että pedagogisella agentilla voi olla monia rooleja. Agentti voi esimerkiksi olla motivaattori, joka kehittää opiskelijan minäpystyvyyttä. Se voi myös olla asiantuntija, joka auttaa oppimistavoitteiden saavuttamisessa, tai vaihtoehtoisesti mentori, joka nostattaa opiskelijan motivaatiota ja kehittää oppimista. Voidaan kuitenkin ajatella, että pedagoginen agentti voi nykyään edustaa kaikkia näitä rooleja yhtäaikaisesti. Tutkimukset ovat osoittaneet vastakkaisia ja ristiriitaisia näkökulmia pedagogisiin agenteihin liittyen, ja eriävät löydökset pohjautuvat esimerkiksi agentin identiteettiin, opiskelijan tasoon sekä agentin toimintoon ja rooliin oppimistilanteessa. (Dai ym., 2022)

Revang ym. (2022) totesivat tutkimuksensa perusteella, että tehokkaimpaan agentin kehitykseen ylletään yhdistämällä määriteltyjä kielimalleja, mallinnettua vuoropuhelua sekä sisäänrakennettuja, toimialakohtaisiin käyttötapauksiin tehtyjä integraatioita. Tässä raportissa käsitellään agenttien kehitykseen tarvittavia alustoja ja näissä käytettyä teknologiaa, joka osaltaan vaikuttaa kehityksen tehokkuuteen. Lisäksi raportissa käsitellään tulevaisuuden teknologiaa, joka tarjoaa mahdollisuuden viedä keskustelevien agenttien kehitystä eteenpäin. Teknologian lisäksi raportissa on tietoa siitä, minkälaisia keskustelevia agenteja eri aloilla ja palveluissa käytetään sekä miten agenttien toteutukseen käytettyä teknologiaa voidaan hankkia markkinoilta.

2 Teknologiat agenttien taustalla

Agenttien ja niiden kehittämisen taustalla oleva teknologia on jatkuvan kehityksen alla. Erityisesti tekoälyllä on merkittävä rooli agenttien kehityksessä, kouluttamisessa sekä käytössä, ja sen merkitys vahvistuu entisestään teknologian kehityksen sekä tämän myötä lisääntyneiden vaatimusten ja odotusten vuoksi. Tässä kartoituksessa on pyritty tunnistamaan ominaisuuksiltaan edistyneempiä sekä tekoälyä hyödyntäviä ratkaisuja, joiden teknologioita kuvataan tarkemmin tässä luvussa.

2.1 Käsitteitä

Monimutkaisempien agenttien taustalla vaikuttaa erilaisia tekoälyyn pohjautuvia teknologioita. Jokainen agentti on omanlaisensa ja näin ollen hyödyntää eri teknologioita käyttötarkoituksensa mukaan. Tässä luvussa näitä tekoälyyn perustuvia teknologioita listataan, ja määritellään lyhyesti, mitä ne ovat.

Tekoäly on käsite, jolla kuvataan ihmisiä matkivia koneita tai teknologioita. Tekoälyn tarkoituksena on helpottaa ihmisten elämää: sille voidaan delegoida aiemmin ihmisen vastuulla olleita tehtäviä, se voi nopeuttaa ja optimoida työntekoa – teknologia kykenee oppimaan sille syötetystä datasta ja hyödyntää näitä oppeja sille määriteltyjen tehtävien suorittamiseksi. Tekoäly on monisyinen käsite, jonka voi määritellä kahdella tavalla: ensimmäinen tapa on määritellä tekoäly sen evoluutioasteiden kautta, jotka ovat heikko/kapea tekoäly, perustason tekoäly ja supertekoäly. Toinen tapa on luokitella erityyppiset tekoälyjärjestelmät analyttiseksi tekoälyksi, ihmisvaikutteiseksi tekoälyksi ja inhimillistetyksi tekoälyksi. Tekoälyä hyödyntäviin teknologioihin kuuluvat muun muassa käsitteet koneoppiminen, neuroverkko ja syväoppiminen. (IBM, 2023; Kaplan ja Haenlein, 2018)

Koneoppiminen on kattokäsite neuroverkolle ja syväoppimiselle sekä tekoälyn alakäsite. Koneoppiminen matkii ihmisen oppimisprosessia hyödyntämällä dataa ja algoritmeja. Klassinen koneoppiminen (ei ”syvä”) on laajalti riippuvainen ihmisen esivalmisteluista: jotta koneoppiminen toimii, on ihmisen siis tarpeen strukturoida eli luokitella dataa jo ennakkoon, jotta kone osaa toimia oikein. (Kavlakoglu, 2020a)

Neuroverkko matkii ihmisaivoja tiettyjen algoritmien avulla. Se muodostuu useista kerroksista, joista jokaisella on oma vastuualueensa. Neuroverkko voi avustaa vaikkapa siinä, onko pitsan tilaaminen järkevää, kunhan se vain saa tarpeeksi dataa tätä päättelyprosessia varten. (Kavlakoglu, 2020a)

Syväoppiminen ymmärtää ja osaa luokitella strukturoimatonta dataa itsenäisesti eli ihmisen puuttuminen prosessiin ei ole välttämätöntä. Suuremman autonomisuuden ansiosta koneelle syötettävä tietoaaineisto voi olla suurempi kuin koneoppimisessa. Syväoppiminen on läheistä sukua neuroverkko-käsitteelle, toisin sanoen sen alakäsite: jos neuroverkossa on enemmän kuin kolme kerrosta, olisi syytä puhua enemmän syväoppimisalgoritmista tai syvästä neuroverkosta. Niin yksinkertaiset neuroverkot kuin monimutkaisemmat syvät

neuroverkot ovat muun muassa NLP:n ja puheentunnistuksen peruspilareita. (Kavlakoglu, 2020a; IBM, 2023.) Syväoppiminen sekä syvät neuroverkot ovat jatkuvan kehityksen alla ja jotta näitä voidaan ymmärtää paremmin, on niiden luonnetta analysoitava systemaattisemmin (Kärkkäinen & Hänninen, 2022).

Luonnollisen kielen prosessointi (*Natural Language Processing, NLP*) on tieteenala, joka tutkii luonnollista kieltä prosessoivaa teknologiaa. Toisin sanoen sen avulla tietokoneet ymmärtävät ja tulkitsevat ihmiskielellä kirjoitettua tekstiä tai puhuttua puhetta. NLP vaatii toimiakseen muun muassa matematiikkaa, lingvistiikkaa ja tietotekniikkaa, kuten tekoälyä, sekä kone- ja syväoppimisen malleja. (Chowdhury, 2003; IBM, 2023.)

Luonnollisen kielen ymmärtäminen (*Natural Language Understanding, NLU*) kuuluu NLP-kattokäsitteen alle ja se analysoi tekstiä ja puhetta syntaktisesti (lauseen kieliopillinen rakenne) ja semanttisesti (sanojen ja ilmaisujen merkitys). Näiden analyysien avulla kone ymmärtää erilaisten tekstien viestin ja kontekstin. (Kavlakoglu, 2020b)

Luonnollisen kielen generointi (*Natural Language Generation, NLG*) kuuluu sekin NLP-kattokäsitteen alle. NLU:n keskittyessä luetun ymmärtämiseen NLG mahdollistaa koneen tekstin tuottamisen. NLG:n avulla kone osaa vastata tekstimuodossa vastauksena saamaansa dataan, kuten käyttäjän palautteeseen. NLU:n tavoin NLG:n täytyy ymmärtää kielen sääntöjä ja nyansseja. (Kavlakoglu, 2020b)

Puheentunnistus (*speech-to-text*) tarkoittaa nimensä mukaisesti sitä, että tietokone ymmärtää ihmisen ihmiskielellä puhuttua puhetta ja osaa muuntaa sen tekstiksi. Kaikki äänikomentoja ymmärtävät sovellukset vaativat tämän ominaisuuden mukanaoloa. Puheentunnistus voi olla mukana esimerkiksi navigaattoreissa ja älypuhelimissa. (Kavlakoglu, 2020b)

Teksti puheeksi (*text-to-speech*) tarkoittaa sitä, että tietokone osaa toistaa ihmisen kirjoittaman tekstin puheena.

Speech Emotion Recognition (SER) on syväoppimiseen perustuva teknologia, joka tunnistaa käyttäjän tunnetilan hänen äänensä perusteella riippumatta lausahduksen semanttisesta sisällöstä. SER sai alkunsa puheentunnistuksesta, jossa kone muuttaa käyttäjän puheen tekstiksi. Monimutkaisemmissa toteutuksissa konteksti ja käyttäjään samastuminen muuttuvat olennaisiksi. Kone tarvitsee avukseen käyttäjän tunnetilan tunnistamiseen matemaattisen mallin. (Hmida, 2022; ProjectPro, 2023; Abbaschian ym., 2021)

2.2 Tekoäly

Jotkut alustat hyödyntävät agentin kehityksessä ja koulutuksessa sekä valmiissa agentissa tekoälyä, joka ilmenee luonnollisen kielen prosessointina (*Natural Language Processing, NLP*) ja ymmärtämisenä (*Natural Language Understanding, NLU*). NLP tunnistaa synonyymeja, tulkitsee ja korjaa oikeinkirjoitusta ja jakaa käyttäjän pyynnön pienempiin, ymmärrettäviin osiin. NLU puolestaan hyödyntää syvä- ja koneoppimista, jolloin agentti

tunnistaa oikean aikomuksen (*intent*), sekä muun saatavilla olevan tiedon avulla käynnistää seuraavia askeleita, kuten entiteettien (*entity*) tunnistamista. NLU:lla kerättyä tietoa hyödyntäen agentti muodostaa vastauksen käyttäjälle. (Boost.ai, 2023.) NLP pyrkii muuntamaan strukturoimatonta kieleen liittyvää dataa strukturoituun muotoon, jonka avulla kone kykenee ymmärtämään puhetta ja tekstiä vastausten laatimista varten. NLU sekä luonnollisen kielen generointi (*Natural Language Generation, NLG*) puolestaan ovat NLP:n alakäsitteitä, joista NLU keskittyy koneen luetunymmärtämiseen tekstin oikean tarkoituksen tunnistamista varten. NLG sen sijaan keskittyy konelähtöiseen ja dataan perustuvaan tekstin tuottamiseen. (Kavlakoglu, 2020b.)

Tekoälyyn pohjautuvat agentit oppivat käydyistä keskusteluista ja kehittävät omaa vuorovaikutustaan ajan kanssa, esimerkiksi pyytämällä käyttäjältä keskustelun lopuksi palautetta. Mikäli agentissa ei ole käytetty edistynyttä tekoälyä, perustuu sen toiminta lähtökohtaisesti sääntöpohjaisuuteen (*rule-based*), jolloin keskustelun aiheiden on pysyttävä ennalta määritettyjen raamien sisällä, eikä vuorovaikutus ole yhtä joustavaa (Alburger, 2022).

Tekoälyä hyödyntävät keskustelevat agentit voidaan jakaa viiteen tasoon, joissa loppukäyttäjän vastuu vuorovaikutuksen onnistumisesta laskee taso tasolta (Nichol, 2020):

1. Ilmoittava

- Kaikki työ loppukäyttäjällä
- Esim. nettilomakkeet

2. Peruskäyttöiset chatbotit

- Kerää tietoa kysymyksillä
- Keskustelu kaatuu herkästi, jos poiketaan ennalta määritetyltä polulta

3. Kontekstuaalinen

- Käyttäjä voi pyytää selvennystä, muuttaa mieltään tai muokata omia sanomisiaan ilman, että agentti sekoaa
- Oltava tarkka päämäärä, mutta käyttäjän ei tarvitse tietää, miten agentti toimii, jotta vältyttäisiin kaatumiselta

4. Konsultoiva

- Käyttäjä voi ilmaista asiansa omin sanoin
- Agentin tehtävä on löytää ratkaisu ja keino auttaa

5. Adaptiivinen

- Muokkautuu käyttäjän tarpeen mukaan monipuolisemmaksi tai yksinkertaisemmaksi
- Tunnistaa vihjeet ja merkit ja mukautuu niiden perusteella

Esimerkiksi Rasa (ks. liite 1) luokitellaan 3. tason agentteja tuottavaksi alustaksi, jolloin alustalla luotu chatbot ymmärtää kontekstin, eikä käyttäjän syötteen tarvitse olla täysin ennalta määritetty (Karajgi, 2021).

Syväoppimista hyödyntävät agentit luodaan koneoppimisen algoritmien avulla, mutta pelkkää koneoppimista käyttävistä agenteista poiketen nämä agentit vaativat vähemmän ihmisen väliintuloa. Algoritmeista luodaan useita kerroksia, jotka muodostavat keinotekoisia neuraalisia verkkoja. Näiden verkkojen avulla agentit voivat tehdä päätöksiä ja toimia strukturoidun datan mukaisesti ihmisten välisen vuoropuhelun tarjoaman datan perusteella. (Luchanka, 2022.) Agenttien koulutuksessa käytetään erilaisia syväoppimisen tekniikoita ja menetelmiä, kuten syvää neuroverkostoa (*Deep Neural Network, DNN*), konvolutiivista neuroverkkoa (*Convolutional Neural Network, CNN*) sekä syvää vahvistusoppimista (*Deep Reinforcement Learning, DRL*). (Dhyani & Kumar, 2021.)

Google on luonut Tensorflow-nimisen syväoppimisen kirjaston, joka on avoimen lähdekoodin myötä vapaasti saatavilla. Kirjasto on Python- ja JavaScript-yhteensopiva ja täynnä kone- ja syväoppimisen malleja sekä algoritmeja (ts. neuroverkkoja) (Dhyani & Kumar, 2021). Tensorflow kouluttaa ja ajaa syviä neuroverkkoja muun muassa kuvantunnistamiseen, luonnollisen kielen prosessointiin ja sanojen upottamiseen, jossa sanat esitetään vektoreina (Yegulalp, 2022). Monilla agenttien kehitysalustoilla, kuten Rasassa ja Botonicissa, hyödynnetään Tensorflow-kirjastoa.

2.3 Agentin ilmenemistavat

Agentti voi ilmentyä monella tavalla. Se voi olla Suomen koululaitoksissa käytetyn ViLLE-oppimisjärjestelmän kaltainen taustatoimija, joka ei selkeästi näy käyttäjälle tai selvästi esillä oleva, persoonallinen ja virtuaalinen avatar tai muu kaksi- tai kolmiulotteinen hahmo, joka on vuorovaikutuksessa käyttäjän kanssa puheen tai tekstin kautta. Agentti voi olla myös fyysinen tosimaailman esine, kuten Jyväskylän yliopiston Pepper-robotti. Agentti voi ilmentyä myös chatbottina, joita on monenlaisia, kuten Engati (2023) on listannut:

- **Painikepohjainen chatbot** (*menu/button-based chatbot*), joka perustuu siihen, että käyttäjälle annetaan aina vain uusia vaihtoehtoryppäitä, jonka vaihtoehtoista hänen täytyy valita osuvin ongelmansa ratkaisemiseksi asiakaspalvelupuhelinten tyyliin. Tämä sopii usein kysytyt kysymykset -tyylisiin yhteydenottoihin.
- **Sääntöpohjainen chatbot** (*linguistic based/rule-based chatbot*) perustuu siihen, että chatbotille on opetettu sanastoa ja kielioppia, joiden voimin se kykenee vastaamaan käyttäjän kysymyksiin, mikäli se tunnistaa avainsanoja ja -ilmaisuja käyttäjän yhteydenotossa. Tällaisen chatbotin kehittäminen toimivaksi ja kattavaksi voi olla hidasta.
- **Avainsanoja tunnistava chatbot** (*keyword recognition-based chatbot*) perustuu NLP:hen, jonka avulla se päättää, miten käyttäjän yhteydenottoon pitäisi vastata. Jos tällainen chatbot saa liikaa

samankaltaisia kysymyksiä, sen avainsanatunnistus voi hämääntyä, jolloin chatbot ei enää ole yhtä hyödyllinen.

- **Koneoppimischatbot** (*machine learning chatbot*) perustuu nimensä mukaisesti koneoppimiseen ja myös tekoälyyn. Se muistaa aiemmin käymänsä keskustelut, ja oppii niistä pikkuhiljaa. Tämä chatbot tunnistaa käyttäjän yhteydenotosta nyansseja.
- **Hybridimallin chatbot** (*the hybrid model*) hyödyntää sekä sääntöpohjaisen chatbotin että tekoälychatbotin ominaisuuksia.
- **Äänichatbot** (*voice bot*) on puheen tunnistamiseen perustuva chatbot, kuten Applen Siri ja Amazonin Alexa.

2.4 Käyttöönotto

Chatbotin rakentamisen kaksi päävaihetta ovat keskustelun suunnittelu ja itse botin rakentaminen valitulla alustalla. Ensimmäisessä vaiheessa käytetään työkaluja ja kartoitetaan vuorovaikutustilanteet, joihin chatbotin pitäisi pystyä osallistumaan. Toisessa vaiheessa käytetään valittua alustaa itse chatbotin rakentamiseen. (Henderson, 2019)

Alustojen käyttöönotto voi olla helppoa tai vaikeaa riippuen alustan käyttäjästä, hänen osaamisestaan sekä valitusta alustasta. Alustoja on tarjolla hyvin erilaisia, ja niitä voidaan luokitella usealla eri tavalla. Yksi mahdollisuus on luokitella alustat ryhmiin sen mukaan, millaisia taitoja niiden käyttöön ottaminen vaatii. Alustat, jotka edellyttävät ohjelmointiosaamista ja koodikielten hallitsemista ovat niin kutsuttuja Framework-alustoja. Yleisimmät ohjelmointikielien eri alustoilla ovat Python, JavaScript, TypeScript ja Java. Stephan Millerin (2022) kuuden suosituimman koodikielen listalta löytyvät myös Ruby, C++, PHP ja Clojure. Platform-alustat tarjoavat enemmän valmiita elementtejä ja graafisen käyttöliittymän. Osa alustoista toimii selaimessa, mutta jotkin edellyttävät alustan koneelle lataamista.

Alustoilla rakennetut agentit, kuten chatbotit, voidaan loppusijoittaa useampaan paikkaan. Agentti voi sijaita pilvipalvelussa tai se voidaan upottaa suoraan esimerkiksi yrityksen verkkosivuille.

2.5 Erilaiset alustat

Framework-alusta sisältää yleensä rakennustyökalun, jolla voi rakentaa erilaisia agenteja. Frameworkiin sisältyy myös bottiyhdistin (*bot connector*), bottihakemisto (*bot directory*) sekä kehittäjäportaali (*developer portal*). (Stefanowicz, 2023.) Esimerkiksi Microsoft Bot Framework -alustalla bottiyhdistimen avulla chatbot voi vaihtaa viestejä käyttäjien kanssa kanavissa, jotka on määritetty Bot Frameworkissa (Python Software Foundation, 2023). Framework-alustoja on tarjolla useilla eri koodauskielillä, ja niitä käyttävätkin henkilöt,

joilla on hyvät ohjelmointitaidot. Framework-alustalla agentin luominen aloitetaan usein niin sanotusti tyhjästä eli alusta asti itse koodaten. (Stefanowicz, 2023.)

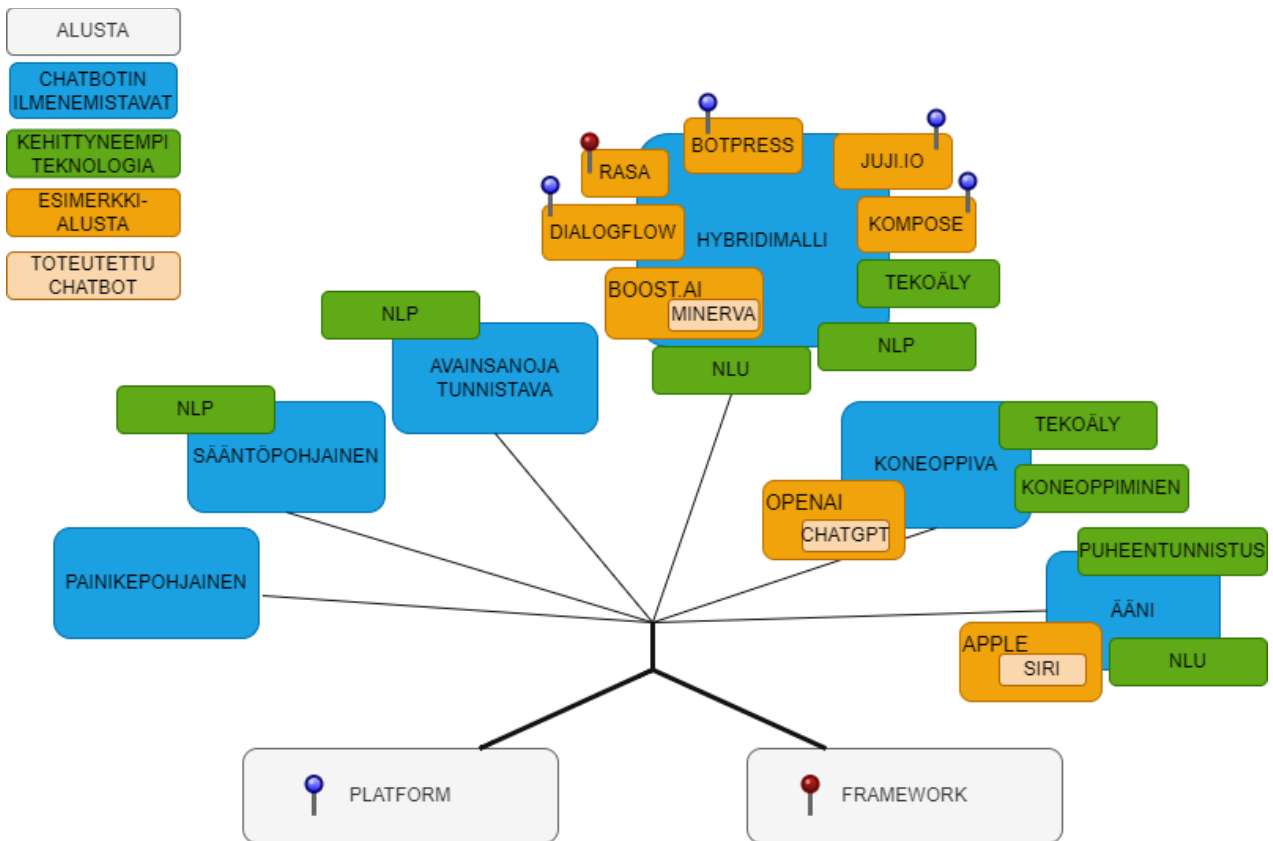
Framework-alustat tarjoavat usein avoimella lähdekoodilla valmiita koodeja, joita lataamalla ja muokkaamalla pystyy rakentamaan agentteja. Tämä nopeuttaa agenttien rakentamista alustoilla, mutta vaatii osaamista ja aikaa. Valmiita koodeja sisältäviä tiedostoja voi olla satoja, joten on huomioitava myös se, että ne vievät tilaa koneelta ja aikaa varsinaiselta agentin rakentamiselta latautuessaan. Tiedostojen lataaminen vaatii usein myös osaamista komentorivin käytöstä.

Avoimen lähdekoodin chatbot on ohjelmisto, jonka lähdekoodi on kaikkien saatavilla. Käyttäjillä on mahdollisuus muokata valmista koodia omien mieltymystensä mukaan. Lähdekoodit löytyvät verkkosivustoilta, esimerkiksi GitHubista. (Stefanowicz, 2023.)

Platform-alustan visuaalisuus helpottaa hahmottamaan sitä, mitä ja millä tavoilla on tarkoitus alustalle syöttää, ja mistä löytyy mitään agentin rakentamisen palikoita. Niiden käyttö ei vaadi ohjelmointitaitoja ja ne ovat sen vuoksi usein nopeampia ja helpompia käyttää kuin Framework-alustat.

Vaikka framework-alustan käyttö saattaa olla työläämpää ja haastavampaa kuin platform-alustan, antaa se eri tavoin mahdollisuuksia agenttien luomiseen: framework-alustat ovat joustavampia ja mahdollistavat usein monipuolisempien robottien luomisen platform-alustoihin verrattuna (Stefanowicz, 2023).

Havainnekuviossa alustoista (kuva 1) on visualisoitu, miten erilaisilla alustoilla voidaan rakentaa eri tavoin ilmeneviä chatbotteja. Kuviossa näkyy myös osa useimmin hyödynnetyistä kehittyneemmistä teknologioista, joita chatboteissa käytetään. Kartoituksessa on painotus kehittyneempiin ratkaisuihin. Kuvioon on tuotu esimerkkeinä muutama tunnettu chatbot havainnollistamaan, millaisesta chatbotin ilmenemistavasta ja toteutuksesta voi olla kyse.



Kuva 1 Havainnekuvio alustoista

2.6 Kouluttaminen

Alustoilla voidaan rakentaa ja kouluttaa erilaisia pedagogisia tai vuorovaikuttavia agenteja. Agenttien koulutus on kuin ennustamista, ja kouluttajan on pyrittävä tietämään, mitä käyttäjät aikovat sanoa tai toivovat saavansa koulutetulta agentilta. Pedagogisten agenttien kouluttamiseen liittyvät vahvasti muutamat käsitteet, joihin tullaan todennäköisesti törmäämään jo kouluttamisen alkuvaiheessa. (Alburger, 2021)

Käsite *utterances* voidaan suomentaa muotoon ilmaisut ja sillä tarkoitetaan jotain, jonka käyttäjä voi sanoa botille (Alburger, 2021).

Käsite *intent* eli aikomus on asia, jonka käyttäjä toivoo saavansa vastauksena botilta (Alburger, 2021). Jos käyttäjä sanoo chatbotille "Hei" tai "Moi", voidaan nämä ryhmitellä tervehdykset-aikomusten ryhmään. Aina kun chatbot vastaanottaa viestin, joka on samankaltainen kuin muut tervehdykset, luokittelee se viestin tervehdykset-aikomuksiin. (Karajgi, 2021)

Käsite *entities* eli suomennettuna entiteetit tarkoittavat avainsanoja, joilla on tarkoitus selvittää käyttäjän tarkoitusta (Alburger, 2021). Entiteetit ovat tietoja, joita poimitaan käyttäjien viesteistä. Esimerkiksi jos

chatbot kysyy käyttäjän yhteystietoja ja vastauksessa on nimi ja sähköposti, luokitellaan nimi ja sähköposti tietoina, jotka chatbot purkaa harjoitustietojen mukaan. (Karajgi, 2021)

Käsite *slots* tarkoittaa chatbottien muistia (Karajgi, 2021).

Käsite *responses* suomentuu luontevasti vastauksiksi ja vastaukset luonnollisesti ovat sitä, mitä chatbotit sanovat käyttäjille. Jotta chatbotit pystyvät antamaan asianmukaisen vastauksen, on niiden ymmärrettävä ensin, mitä käyttäjät sanovat heille. Kun esimerkiksi chatbot kysyy käyttäjän yhteystietoja, ja käyttäjä on ne antanut, voidaan olettaa, että chatbot on ohjelmoitu vastaamaan ”kiitos”. (Karajgi, 2021)

Kun agenttia lähtee kouluttamaan, on itsellä oltava käsitys siitä, millaisen agentin haluaa kouluttaa ja siitä, millaisia vastauksia esimerkiksi toivoo agentin antavan käyttäjälle. Kouluttaessa on mietittävä erilaisia skenaarioita, joita agentilla saattaa tulla vastaan sekä suunniteltava ja testattava etukäteen se, millä tavoin agentti reagoi käyttäjän toimiin. Iso osa esimerkiksi chatboteista on koulutettu sääntöjen mukaan vastaamaan tiettyihin kysymyksiin tietyllä tavalla. Mikäli alusta mahdollistaa tekoälyn käytön laajemmin, on mahdollista kouluttaa agenttia vuorovaikutteisemmaksi.

Agenttien kouluttamisessa ensimmäinen askel on keskustelun suunnittelu. Kun selvillä on se, mitä käyttäjät agentilta haluavat, kartoitetaan mahdollisia keskustelun polkuja. Tavoitteena on huomioida monipuolisesti useita käyttäjän vastauksia kuhunkin chatbotin tulosteeseen. Näistä vastauksista muotoutuu niin sanottu virtauskartta, joka toimii komentosarjojen perustana. Seuraavasta vaiheesta voidaan käyttää nimitystä skriptaus. Tämän vaiheen päätavoite on agentin persoonallisuuden kehittäminen ja suunnittelu. Näiden vaiheiden jälkeen päästään valitsemaan haluttu alusta, jolla varsinainen agentti luodaan. (Henderson, 2019)

Agentin kouluttamisen vaiheet
1. Määritä chatbotin erityiset käyttötapaukset.
2. Varmista, että aikomukset ovat erilaisia.
3. Varmista, että jokainen aikomus sisältää useita ilmaisuja.
4. Kasaa monipuolinen tiimi huolehtimaan botin koulutusprosessista.
5. Varmista, että entiteetit ovat tarkoituksenmukaisia.
6. Älä unohda lisätä persoonallisuutta.
7. Älä luota vain tekstiin.
8. Älä lopeta harjoittelua.

Taulukko 2. Agentin kouluttamisen vaiheet. (Alburger, 2021)

Tekoälyä koulutetaan esimerkkien avulla. Käytännössä se tarkoittaa sitä, että ihminen on antanut merkitykset laajalle joukolle esimerkkilauseita, jotka on syötetty alustalle. Pelkät esimerkkilauseet eivät kuitenkaan riitä vuorovaikutteisen agentin luomiseen, vaan tarvitaan matemaattisempia menetelmiä, eli koneoppimista. Kun esimerkiksi chatbotin neuroverkkoa lähdetään kouluttamaan, ei se aluksi ymmärrä mitään, vaan arvaa täysin satunnaisesti. Jotta tekoäly oppii lauseista merkityksiä, tarvitaan miljoonia toistoja. Kun agentti vastaa väärin, se saa käyttäjältä tiedon oikeasta vastauksesta. Koulutusalgoritmi lähettää korjaussignaalin etenemään neuroverkossa, mikä korjaa hermosolujen käyttäytymistä ja vähitellen koko käyttäytyminen korjautuu oikeampaan suuntaan. Tuhansia kertoja toistettu yritys-erehdys-korjaus-ketju vaikuttaa neuroverkkoihin, minkä seurauksena neuroverkko alkaa muistamaan tiettyjä rakenteita ja reagoimaan niihin tietyillä tavoilla. (Onera, n.d.)

Tavoitteena chatbottien kouluttamisessa on syöttää alustalle mahdollisimman monipuolista tietoa, jotta chatbotin on mahdollista ymmärtää muun muassa sanojen välisiä assosiaatioita. Mahdotonta on kattaa kaikkia mahdollisia vaihtoehtoja, mutta yleisimmät ja todennäköisimmät sanat ja lauseet opettamalla voidaan varmistaa, että chatbot ymmärtää sille osoitettuja viestejä mahdollisimman hyvin. (Karajgi, 2021)

2.7 Integroitavuus

Agentit voivat olla alustasta riippuen integroitavissa sosiaaliseen mediaan, viestintäkanaviin tai älykkäisiin assistentteihin, kuten Applen Siriin tai Amazonin Alexaan. Agentti voi myös olla oma sovelluksensa tai verkkopohjaisesti toimiva kokonaisuutensa, jolloin se toimii serverillä, alustalla tai itsenäisesti sovelluksessa. Eri alustoilla on erilaiset toimintamallit sille, miten alustalla tehdyn agentin voi viedä osaksi toista palvelua tai nettisivustoa. Osa alustoista konfiguroi skriptin valmiiksi, joka voidaan upottaa kohteen HTML:ään. Moni alusta mahdollistaa integroinnin eri viestintäkanaviin, kuten Facebook Messengeriin, Teamsiin tai Slackiin, ja jokaiselle kanavalle on omat ohjeensa. Jotkin alustat vaativat konttitekniikan hyödyntämistä, jotta koodit ja

riippuvuudet saadaan kokonaisuutena serverille, pilveen tai muuhun kohteeseen. Lisäksi monet integroinnit hyödyntävät webhookeja, eli uuden tietosisällön laukaisemia ja sovelluksen lähettämiä automatisoituja viestejä, joiden avulla voidaan välittää ja tallentaa chatbotin ja käyttäjän toimintaan liittyvää tietoa.

Kartoituksessa on pyritty huomioimaan erityisesti niitä alustoja, joilla voidaan integroida agentti osaksi Moodlea ja Thinglinkiä. Moodleen on onnistuneesti integroitu esimerkiksi DialogFlow'lla (Bandita & Bändeali, n.d), Tidio.io:lla (Less than 5 minutes, 2022), Landbotilla (Integrately, 2022) ja Komposella (Avinash, 2021) tehtyjä agentteja. ThingLinkin kaltaisille alustoille voi olla haastava integroida vuorovaikutteista agenttia, jossa teknologia mahdollistaa mukautuvan oppimisen.

2.8 Jyväskylän yliopiston Minerva-chatbot

Kartoituksen aikana lähestyttiin Jyväskylän yliopiston digipalveluita yliopiston oman Minerva-chatbotin kehityksen tiimoilta. Agentin kehitysalusta, boost.ai (ks. liite 1), on maksullinen ja kokeiluunkin on pyydettävä tarjous, jonka vuoksi alustaa ei ole kokeiltu tässä vaiheessa projektia. Minervan pääkäyttäjä kuitenkin järjesti chatbotista demotilaisuuden, jossa havainnollistettiin boost.ai:n ja Minervan ominaisuuksia.

Boost.ai on digipalveluiden tiedon perusteella laadukas ja hintansa arvoinen alusta, jolla voidaan luoda monipuolisia chatbotteja ilman koodausta tai vahvaa teknistä osaamista. Alustalla luotujen chatbottien ylläpitäjille järjestetään ilmaisia koulutuksia, joiden avulla kuka tahansa voi kerryttää tarvittavat tiedot chatbotin ylläpitoa ja koulutusta varten. Alustojen kartoitus- ja kilpailutusvaiheessa digipalvelut asettivat muutamia kriteerejä, joiden mukaan boost.ai valikoitui Minervan kehitysalustaksi. Haluttiin hyödyntää erilaisia tekoälyratkaisuja, jotta välttyttäisiin mahdollisimman paljon manuaaliselta työltä chatbotin kehitykseen ja koulutukseen liittyen. Lisäksi toivottiin, että tuleva chatbot voisi tunnistaa hyvin laajasti kielellisiä elementtejä ja pystyä keskustelemaan käyttäjän kanssa. Chatbotin toimittamisessa yliopistolle hyödynnettiin Front.ai-nimistä yritystä, joka tarjoaa korkeakouluille chatbotteja.

Laurean ammattikorkeakoulu oli ottanut aiemmin käyttöön chatbotin samoja palveluita ja alustaa hyödyntäen, joten Jyväskylän yliopisto sai heiltä lisätietoa aiheeseen liittyen. Minerva-chatbotin käyttöönottoa helpottivat Laurealta saadut esimerkit ja valmis pohja, jota hyödynnettiin Minervan kehityksessä. Sisällön suunnitteluun ja tuottamiseen meni syksyllä 2021 pari kuukautta, eikä sisällön tarvinnut olla lanseerausta varten täysin hiottu, sillä suosituksena on julkaista chatbot mahdollisimman aikaisessa vaiheessa turhan työn välttämiseksi. Testausprosessissa oli koulutuspalveluilta koottu ryhmä mukana testaamassa sisältöä ja Minervan ominaisuuksia, mutta sisällön luonnin ja lanseerauksen jälkeen resurssien tarve on pienentynyt. Minervan pääkäyttäjä pääsee muokkaamaan kaikkia osa-alueita ja kohdentamaan eri osa-alueiden muutamasta henkilöstä koostuville ylläpitäjäryhmille muokausehdotuksia. Tällä hetkellä ylläpitäjiä on kaiken kaikkiaan yli 10.

Teknisestä näkökulmasta Minervan kouluttamista ja kehittämistä voidaan pitää melko yksinkertaisena, sillä alusta ei vaadi koodin kirjoittamista. Toiminnot ovat kuitenkin hyvin monipuolisia ja pitkälti automatisoituja, ja ylläpitäjät saavat keskustelujen pohjalta ehdotuksia lisättävistä aihealueista, aikeista, sanoista ja synonyymeista. Chatbot ei kuitenkaan kouluta itse itseään, vaan jonkun on käytävä hyväksymässä tai hylkäämässä ehdotukset manuaalisesti. Omaa dataa esimerkiksi koulutukseen liittyen voidaan tuoda Excel-tiedostosta, ja chatbotin vastauksessa voidaan hyödyntää kuvia, videoita, tekstiä, linkkejä sekä JSON-objekteja. Minerva on toistaiseksi integroitu eri nettisivuille HTML-skriptien avulla, jotka Front.ai on tarjonnut, sillä chatbot sijaitsee heidän palvelimellaan. HTML-skriptin kautta voidaan tuoda boost.ai:lla tehty chatbot myös Moodleen, mutta Minervan kohdalla tällaista integraatiota ei ole kokeiltu.

2.9 Tulevaisuuden teknologiat

Keskusteleuvia agenteja, kuten chatbotteja, kehitetään jatkuvasti. Uusia teknologioita ja innovaatioita pyritään hyödyntämään mahdollisimman monipuolisesti, jotta agenteista saadaan mahdollisimman tehokkaita ja käyttäjäystävällisiä.

Yrityksillä on pyrkimyksenä madaltaa omien chatbottien markkinoilletuontiaikaa esimerkiksi siirtymällä pois kehitysalustoista ja kohti valmiiden, muokattavien ja vähemmän koodausta vaativien teknologioiden hyödyntämistä. On ennustettu, että siirtyminen näihin dynaamisiin agenteihin voi vähentää markkinoilletuontiaikaa 50 prosentilla. (Ravinutala, 2022.) Tulevaisuudessa tullaan näkemään useamman eri chatbotin hyödyntämistä samassa ympäristössä, kun yksittäisen chatbotin tarkoitusta rajataan ja tarkennetaan. Tällöin käyttäjä kohtaa päächatbotin, joka ohjaa kyselyiden ja käyttäjän vaatimusten perusteella käyttäjän seuraavalle chatbotille. (yellow.ai, 2022.)

Large Language Models (LLM) -mallien tutkimusta edistetään jatkuvasti, ja niiden on suunniteltu helpottavan muun muassa keskustelun autogenerointia. LLM kokonaisuudessaan vaatii vielä paljon kehittämistä, mutta sen odotetaan auttavan luonnollisen kielen generoinnissa, esimerkiksi rakentamalla NLU-malleja. LLM voi olla hyödyksi muun muassa silloin, kun käyttäjän ilmaisu (*utterance*) ei täsmää minkään aikomuksen (*intent*) kanssa, mutta virheilmoituksen sijaan LLM auttaakin laajentamaan käyttäjän ilmaisua ja sitä kautta löytämään täsmävän aikomuksen. (Merritt, 2022.) LLM-malleja on kehitetty vuosikymmenten ajan, ja tällä hetkellä pinnalla oleva ChatGPT-teknologia on esimerkki kehityksen edistymisestä. Alan asiantuntija Adi Andrein mukaan LLM-mallit ovat aiemmin tarjonneet noin 65 prosentin tarkkuuslukeman, mutta nykyään luku lähentelee jo 85–90 prosenttia, ja tästä syystä malleja nähdäänkin suurten käyttötapausmäärien konteksteissa, kuten ChatGPT-sovelluksessa. (Drenik, 2023.) Odotetaan, että tulevaisuudessa LLM-mallit voivat generoida omaan koulutukseensa käytettävää materiaalia ja dataa sekä tarkistaa faktansa omatoimisesti (Toews, 2023).

On ennustettu, että multimodaalisten, eli useampaa vuorovaikutuskanavaa ja -tapaa hyödyntävien keskusteluvien agenttien kehitys lisääntyy, jolloin esimerkiksi chatbot voi reagoida käyttäjän puheeseen ja tekstimuotoiseen viestintään yhtä lailla (Dilmegani, 2023a). Käyttäjän tunnistautumista vaativissa ympäristöissä, kuten terveydenhoitoalan palveluissa, voidaan odottaa puheentunnistamista ja biometriikkaa hyödyntävää teknologiaa. Koetaan, että laadukas puheen ominaisuuksiin pohjautuva teknologia on jopa turvallisempi kuin perinteisen salasanan käyttäminen. Vuorovaikutuksen laadun odotetaan jatkuvasti kehittyvän, kun NLP-mallit paranevat ja NLU tehostuu entisestään. (Dilmegani, 2023b.) Lisäksi odotetaan, että agenteissa käytetty tekoäly mahdollistaisi entistä paremmin kuvien/dokumenttien tunnistusta ja näistä olennaisten tietojen keräystä, jotta käyttäjän ei enää tarvitsisi kirjoittaa kaikkea tekstimuotoon (Merritt, 2022).

3 Teknologioiden hankinta

Käytettävän alustan valintaan vaikuttaa merkittävästi se, miten teknologia on hankittavissa. Erilaisia alustoja on markkinoilla kymmenittäin. Alustoja on saatavilla ilmaiseksi sekä maksua vastaan. Alustat toimivat joko ladattavina ohjelmina tai selaimessa. Ilmaiseksi saatavilla olevan alustan käyttö vaatii pääsääntöisesti rekisteröitymistä ja monessa tapauksessa useiden tietojen ja oikeuksien luovuttamista tai niiden luovuttamisen hyväksymistä. Useat alustat ovat suunniteltu sosiaalisen median viestintäkanaviin, joten eräskin (wit.ai, liite 1) alusta hyväksyy vain Facebook-kirjautumisen. Monet ilmaiset alustat ovat ladattavissa avoimella lähdekoodilla GitHubista, ja ladattavia tiedostoja voi olla satoja.

On tärkeää olla tarkkaavainen, minkä kehittäjän alustaa on hankkimassa ja lataamassa. On hyvä tutustua alustan kehittäjän ja julkaisijan taustaan. Kuinka vanha tai uusi alusta on? Onko taustalla tunnettu toimija? Paljonko alustalla on käyttäjiä? Mikä taustarahoitus ja käyttötarkoitus alustalla on? Taustaa selvittämällä voi myös arvioida, tuleeko alusta olemaan kestävä tai kehittyvä tulevaisuudessa. Onko alustan kehittäjiltä tai muilta käyttäjiltä saatavissa ohjeita, valmiita koodeja tai muuta apua? Alustan ja agentin käyttö vaatii muistitilaa ja pilvipalveluita, joten on tärkeää perehtyä siihen, mikä on näiden sijainti, yksityisyys ja toimintavarmuus. (Carpentier, n.d.)

Avoimen lähdekoodin alustat eivät tarkoita aina ilmaisia alustoja. Ilmaisen, avoimen lähdekoodin alustan käyttö vaatii kehittyneitä ohjelmointitaitoja, kuten luvussa 2.4 on esitelty. Paljon käytetyille avoimen lähdekoodin alustoille voi olla paljonkin valmista koodia saatavilla. Useasta ilmaisesta tai avoimen lähdekoodin alustasta on myös olemassa maksullisia versioita. Maksullisten versioiden markkinointivaltteina ovat esimerkiksi parempi tietosuoja, suurempi työkaluvalikoima sekä suurempi määrä viestejä tai ominaisuuksia.

Osa alustoista on suoraan maksullisia. Maksullisuus perustuu useimmissa tapauksissa portaittasiin kuukausimaksuihin tai chatbottien tai viestien määrään. Hinnat liikkuvat muutamasta eurosta satoihin euroihin kuukaudessa. Osaan alustoista saa tarjouksen vain pyytämällä. Monia maksullisiakin alustoja voi päästä kokeilemaan koejaksolle esimerkiksi verkkoversiosta, jolloin alustasta voi saada käyttökokemuksen. Joillain alustoilla voi kehittää oman chatbotin ja vasta julkaisuvaiheessa alustasta tulee maksullinen.

Esimerkkejä ilmaisista avoimen lähdekoodin alustoista ovat Juji.io ja Tock. Esimerkkejä osittain maksullisista avoimen lähdekoodin alustoista ovat RASA, Botpress ja Microsoftin Bot Framework. Esimerkkejä maksullisista alustoista, joita voi testata ilmaiseksi, ovat Kompose chatbot, Landbot ja DialogFlow.

4 Johtopäätökset

Kartoitusraportin tarkoituksena oli agenttien kehitysalustojen yleisen tason kartoituksen lisäksi koostaa projektin arviointimatriisia varten lista potentiaalisista alustoista. Tässä luvussa esitellään alustojen kartoitusprosessia, perusteluita listalle valikoituneista ja listalta pois pudotetuista eli hylätyistä alustoista sekä käytännön esimerkkejä potentiaalisiksi todetuista vaihtoehtoista.

4.1 Kartoitettujen alustojen yhteenveto

Kartoitusvaiheen alkupuolella löydettyjä agentin kehitysalustoja oli noin 60, joista kaikki on listattu alustojen pitkään listaan (ks. liite 1). Alustoja kerättiin monipuolisesti eri lähteistä ja erityisesti sellaisista, joissa listattiin nykyhetkellä pinnalla olevia vaihtoehtoja. Lähteitä läpikäydessä tiedostettiin, että Top 10 -tyyliset listat on mahdollisesti koottu kaupallinen näkökulma priorisoiden, jolloin alusta ja sen käyttökelpoisuus arvioitiin omien vaatimusten mukaisesti. Alustojen listaukseen pyrittiin löytämään pääosin sellaisia vaihtoehtoja, jotka ovat edelleen aktiivisessa käytössä olevia avoimen lähdekoodin platform- tai framework-alustoja. Lisäksi listaan nostettiin näistä kriteereistä poikkeavia alustoja muun muassa siitä syystä, että niissä hyödynnettiin jonkin merkittävän toimijan, kuten Googlen tai Microsoftin, teknologioita. Mikäli alustasta paljastui nopeasti esimerkiksi teknologiaan tai käyttöönnottoon liittyvä ominaisuus, joka ei palvellut kartoituksen tavoitteita, kyseinen alusta hylättiin ilman perusteellisempaa tutustumista. Alustoja pyrittiin keräämään mahdollisimman monipuolisesti eri toimialoilta, mutta liian rajatut ja projektille epäolennaiset mahdollisuudet johtivat alustan hylkäykseen. Kartoitusta tehdessä kuitenkin tiedostettiin, että pedagogisia agentteja voidaan luoda alun perin eri käyttötarkoitukseen suunnitellulla alustalla. Alustakohtaiset tiedot valikoitiin ja hylkäykseen johtavista perusteluista löytyvät liitteenä 1 olevasta pitkästä listasta.

Alustojen pitkässä listassa on listattu vihreällä ne alustat, jotka on todettu potentiaalisiksi alustoiksi osaksi arviointimatriisia. Nämä alustat ovat Rasa, Botpress, Dialogflow, Juji.io ja Kompose. Näiden alustojen joukossa olevista Rasasta ja Botpressistä on myöhemmin tässä luvussa tarkempaa tietoa, minkä avulla on pyritty konkreettisemmin kuvaamaan alustojen toimintaa. Punaisella pohjalla olevat alustat on hylätty erinäisistä syistä. Näitä syitä ovat maksullisuus tai muu käyttöönoton vaikeus jo kokeiluvaiheessa, rajattu käyttötarkoitus esimerkiksi tietylle toimialalle, heikot mahdollisuudet kehittää omaa agenttia, alustan jatkon epävarmuus sekä liian yksinkertainen teknologia.

Alustat ovat valikoituneet tarkempaan tutustumiseen sen perusteella, kuinka helposti alustaa pääsee kokeilemaan ja onko tämä mahdollista tehdä ilmaiseksi. Alustoja on myös pyritty valitsemaan siten, että ne edustaisivat eri teknologioita ja niissä olisi mahdollisimman monipuolisesti hyödynnetty tekoälyllä varustetun agentin luomiseen tarvittavia ominaisuuksia. Kriteerinä oli myös se, että alusta soveltuu monipuolisemman agentin, kuten pedagogisen agentin, kehittämiseen, eikä palvele ainoastaan rajattua toimialaa.

4.2 Esimerkkejä alustoista

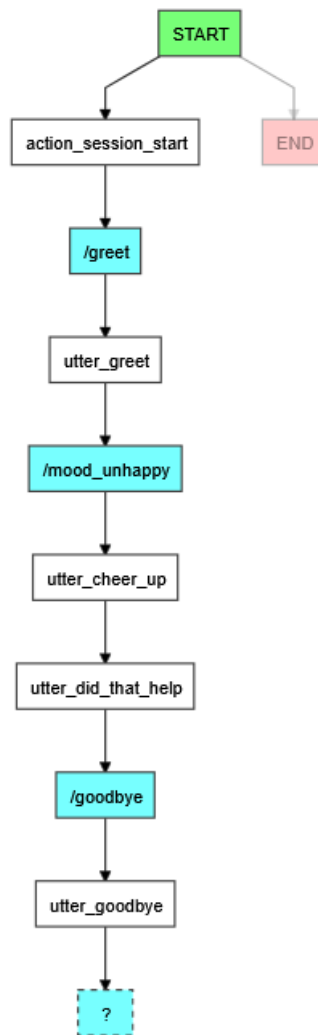
Tähän lukuun on koostettu tietoa kahdesta tarkempaan tutustumiseen valitusta alustasta, Rasasta ja Botpresistä. Nämä kaksi alustaa valikoituivat esiteltäviksi alustoiksi siksi, että ne löydettiin jo kartoitusvaiheen alkupuolella ja ovat siitä syystä hyvin tuttuja alustoja, sekä ne edustavat kehitysalustojen kahta eri tyyppiä (framework- ja platform-alusta).

4.2.1 Rasa

Rasa valikoitui tarkemman tutustumisen kohteeksi sen perusteella, että alustaa voi kokeilla helposti ja ilmaiseksi. Rasa Open Source on avoimen lähdekoodin framework-alusta, jota käytetään komentotulkilla python-ympäristössä, ja tiedostot kirjoitetaan YAML-merkintäkielellä.

Käytön voi aloittaa luomalla agentin valmiilla pohjalla, jota voi muokata tarpeen mukaan. Agenttia voi kouluttaa muokkaamalla tiedostoja ja ajamalla koulutuskäskyn komentotulkissa. Koulutusta voi myös tehdä interaktiivisesti, jolloin agentti tutkii reaaliaikaisesti olemassa olevia aikomuksia, entiteettejä, vastausvaihtoehtoja ja muistipaikkoja, jonka jälkeen se ehdottaa seuraavaa toimintoa. Ehdotus voidaan hyväksyä tai hylätä, jolloin agentti pyytää käyttäjää antamaan tietoja uuteen aikomukseen. Agentti luo koulutuksen aikana tulleesta materiaalista itselleen "tarinoita", jotka toimivat ennakkotapauksina ja ohjeina tulevaisuudessa. Käyttäjä voi myös luoda tarinoita koulutusmielessä sekä testikäyttöön. Testikäyttöön luodut tarinat ajetaan testikäskyllä ja toimivat samaan tapaan kuin testaaminen ohjelmoidessa, jolloin käyttäjä saa näkyviin testauksen tulokset sekä virheilmoitukset.

Agentin luomisessa tarvitaan useita tiedostoja, kuten nlu.yaml, stories.yaml, domain.yaml sekä config.yaml. NLU-tiedostossa käsitellään aikomuksia antamalla jokaiselle nimi ja esimerkkejä käyttäjän syöttämistä lauseista, ilmaisuista, sanoista ja kirjaimista, jotka liittyvät kyseiseen aikomukseen. Kun aikomus on tunnistettu, siirtyy agentti stories-tiedostossa sijaitsevassa tarinassa aikomusta vastaavaan toimintaan. Toimintoja käsitellään domain-tiedostossa, jossa agentin sanomiset ja tekemiset on kuvattu kunkin toiminnan ilmetessä. Lisäksi domain-tiedostossa esitellään työmuistipaikat ja kerrotaan, mistä ne saavat arvonsa, onko muistipaikalla oletusarvoa ja onko muistipaikan arvolle asetettu rajoituksia.



Kuva 2 Agentin kouluttamisen kaavio, Rasa

Kaavio kuvassa 2 visualisoi Rasalla tehdyn agentin kouluttamiseen liittyvää virtaa. Kaavio on luotu reaaliaikaisesti samalla, kun agenttia on koulutettu interaktiivisesti. Valkoisella pohjalla olevat toiminnot ovat agentin sanomisia, kun se reagoi turkoosipohjaisiin käyttäjän kommentteihin.

4.2.2 Botpress

Botpress valikoitui tarkemman tutustumisen kohteeksi, sillä sitä voi testata ilmaiseksi, ja alusta oli helppo ottaa käyttöön. Botpress on avoimen lähdekoodin platform-alusta, eli se toimii graafisen käyttöliittymän avulla. Alustan jotkut ominaisuudet vaativat JavaScriptin kirjoittamista. Botpressin voi ladata omalle tietokoneelle tai sitä voi käyttää selainversiossa. Alustalle luodaan virtoja visuaalisista elementeistä, jotka hyödyntävät tekoälyä, kuten luonnollisen kielen prosessointia ja tulkintaa. Kokeilun helpottamiseksi käyttäjä voi aloittaa agentin kehityksen valmiilla pohjalla, jota muokataan tarpeen mukaan. Valmis agentti voidaan integroida

osaksi viestintäkanavia, kuten Whatsappia ja Teamsiä, tai se voidaan upottaa vapaasti valittavalle verkkosivulle.

Lähteet

- Abbaschian, B. J., Sierra-Sosa, D., & Elmaghraby, A. (2021). Deep Learning Techniques for Speech Emotion Recognition, from Databases to Models. *Sensors*, 21(4), 1249. MDPI AG. <http://dx.doi.org/10.3390/s21041249>
- Alburger, J. (2021). How to Train a Chatbot: 8 Effective Tips for Training AI. Haettu 17.2.2023 osoitteesta: <https://www.hubtype.com/blog/how-to-train-a-chatbot-2>
- Alburger, J. (2022). Conversational AI: The Complete Guide to Conversational Artificial Intelligence. Haettu 16.2.2023 osoitteesta: <https://www.hubtype.com/blog/conversational-ai>
- Avinash. (2021). How to Build and Integrate a Chatbot in Moodle. Haettu 22.3.2023 osoitteesta: <https://www.kommunicate.io/blog/moodle-chatbot/>
- Bandita, S. & Bandeani, N. (n.d). Moodle Integration with DialogFlow (chatbots). Haettu 22.3.2023 osoitteesta: <https://www.springml.com/blog/moodle-integration-with-dialogflowchatbots/>
- Boost.ai. (2023). What is Conversational AI, anyway? Haettu 16.2.2023 osoitteesta: <https://www.boost.ai/knowledge/what-is-conversational-ai>
- Carpentier A. (n.d). Open source conversational agent or not: what to choose? Botnation AI. Haettu 19.2.2023 osoitteesta: <https://botnation.ai/site/en/conversational-agent-open-source/>
- Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology*, 37(1), 51-89.
- Dai, L., Jung, M., Postma, M., & Louwerse, M. (2022). A systematic review of pedagogical agent research: Similarities, differences and unexplored aspects. *Computers & Education*, vol 190. <https://doi.org/10.1016/j.compedu.2022.104607>
- Dhyani, M. & Kumar, R. (2021). An intelligent chatbot using deep learning with bidirectional RNN and attention model. *Mater Today Proc.* Vol 34, 817-824. <https://doi.org/10.1016/j.matpr.2020.05.450>
- Dilmegani, C. (2023a). Top 5 Expectations Concerning the Future of Conversational AI. Haettu 15.2.2023 osoitteesta: <https://research.aimultiple.com/future-of-conversational-ai/>
- Dilmegani, C. (2023b). Top 5 Expectations Regarding the Future of NLP in 2023. Haettu 15.2.2023 osoitteesta: <https://research.aimultiple.com/future-of-nlp/>
- Drenik, G. (2023). Large Language Models Will Define Artificial Intelligence. Haettu 25.2.2023 osoitteesta: <https://www.forbes.com/sites/garydrenik/2023/01/11/large-language-models-will-define-artificial-intelligence/?sh=13f4d7abb60f>
- Engati (2023). 6 types of chatbots - Which is best for your business? Haettu 28.2.2023 osoitteesta: <https://www.engati.com/blog/types-of-chatbots-and-their-applications>

- Henderson, M. (2019) A Technical Guide to Building an AI Chatbot. Haettu 27.2.2023 osoitteesta: [How to Build AI Chatbot: A Complete Development Guide | AltexSoft](#)
- Hmida, F. (2022). How to build your own Speech Emotion Recognition? Haettu 24.2.2023 osoitteesta: <https://vivoka.com/how-to-speech-emotion-recognition/>
- IBM (2023). What is machine learning? Haettu 19.2.2023 osoitteesta: <https://www.ibm.com/topics/machine-learning>
- IBM (2023). Natural Language Processing (NLP). Haettu 19.2.2023 osoitteesta: <https://www.ibm.com/en/topics/natural-language-processing>
- Integrately. (2022). 8 Million+ Ready Automations For 1000+ Apps. Haettu 22.3.2023 osoitteesta: <https://integrately.com/integrations>
- Iuchanka, A. (2022). How do chatbots work? Often with a little help from AI. Haettu 23.2.2023 osoitteesta: <https://www.itechart.com/blog/how-do-chatbots-really-work/>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business horizons*, 62(1), 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Karajgi, A. (2021). Building a Chatbot with Rasa. Haettu 16.2.2023 osoitteesta: <https://towardsdatascience.com/building-a-chatbot-with-rasa-3f03ecc5b324>
- Kavlakoglu, E. (2020a). AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference? Haettu 19.2.2023 osoitteesta: <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>
- Kavlakoglu, E. (2020b). NLP vs. NLU vs. NLG: the differences between three natural language processing concepts. Haettu 19.2.2023 osoitteesta: <https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/>
- Kärkkäinen, T. & Hänninen, J. (2022). Supplementary for 'An Additive Autoencoder for Dimension Estimation'*
- Less than 5 minutes. (2022). How to make a chatbot for Moodle with Tidio? [video]. Haettu 22.3.2023 osoitteesta: <https://www.youtube.com/watch?v=MALDkclnrFQ>
- Merritt, A. (2022). The Future of Conversational AI – Trends and predictions from the experts. Haettu 16.2.2023 osoitteesta: <https://medium.com/conversationalai/the-future-of-conversational-ai-c3ea4e31e8e1>
- Miller, S. (2022). Top 6 Programming Languages for Chatbot Development. Haettu 22.2.2023 osoitteesta: <https://www.codecademy.com/resources/blog/top-6-programming-languages-for-chatbot-development/>

- Nichol, A. (2020). 5 Levels of Conversational AI: 2020 Update. Haettu 16.2.2023 osoitteesta: <https://rasa.com/blog/5-levels-of-conversational-ai-2020-update/>
- Onerva (n.d.). Miten kouluttaa tekoäly ymmärtämään luonnollista kieltä? Haettu 22.2.2023 osoitteesta: <https://onervahoiva.fi/miten-kouluttaa-tekoaly-ymmartamaan-luonnollista-kielta/>
- ProjectPro (2023). Speech Emotion Recognition Project using Machine Learning. Haettu 24.2.2023 osoitteesta <https://www.projectpro.io/article/speech-emotion-recognition-project-using-machine-learning/573>
- Python Software Foundation (2023) Botframework-connector. Haettu 27.2.2023 osoitteesta: [botframework-connector · PyPI](https://github.com/python-botframework-connector)
- Ravinutala, R. (2022). Next-gen Conversational AI: Technology that is redefining experiences. Haettu 14.2.2023 osoitteesta: <https://aijournal.com/next-gen-conversational-ai-technology-that-is-redefining-experiences/>
- Revang, M., Mullen, A., & Elliot, B. (2022). Critical Capabilities for Enterprise Conversational AI Platforms. Gartner.
- Stefanowicz, B. (2023). 9+ Best Open Source Chatbot Frameworks Compared. Haettu 17.2.2023 osoitteesta: <https://www.tidio.com/blog/chatbot-framework/#what-is>
- Toews, R. (2023). The Next Generation of Large Language Models. Haettu 25.2.2023 osoitteesta: <https://www.forbes.com/sites/robtoews/2023/02/07/the-next-generation-of-large-language-models/?sh=72f4e35418db>
- Weber, F., Wambsganß, T., Rüttimann, D. & Söllner, M. (2021). Pedagogical Agents for Interactive Learning: A Taxonomy of Conversational Agents in Education. International Conference on Information Systems (ICIS) 2021. Austin, USA.
- Yegulalp, S. (2022). What is Tensorflow? The machine learning library explained. Haettu 25.2.2023 osoitteesta: <https://www.infoworld.com/article/3278008/what-is-tensorflow-the-machine-learning-library-explained.html>
- Yellow.ai. (2022). The Future of Conversational AI: Progress and Advancement. Haettu 14.2.2023 osoitteesta: <https://yellow.ai/conversational-ai/future-of-conversational-ai/>

Liitteet

Liite 1: Alustojen pitkä lista

Värikoodit	Tutkitaan	Ei jatkoon
------------	-----------	------------

Nimi	Ilmainen/hinta	Avoin lähdekoodi (ohjelmointikieli)	Toimintaympäristö	Muuta	Hyv-/hyl-perustelut
Botpress https://botpress.com/	Ilmainen osittain (tiettyyn viestimäärään asti)	Avoin lähdekoodi (JavaScript /no-code)	Ladattava, graafinen ohjelma, myös selainversio	Ei välttämättä vaadi koodausta.	Voi kokeilla ilmaiseksi, hyvä dokumentaatio, monipuolinen teknologia
Dialogflow https://cloud.google.com/dialog-flow	Ilmainen kokeilu pelkistetystä versiosta, hinta tapauskohtainen	Ei avoin lähdekoodi	Selaimessa	GoogleCloudin. Ei vaadi koodausta.	Usein käytetty alusta, integroitavuus, mahdollisuus kokeilla ilmaiseksi, vahva toimija taustalla
Juji.io https://juji.io/	Ilmainen	Avoin lähdekoodi	Selaimessa, graafinen	Ei vaadi koodausta.	Voi kokeilla ilmaiseksi, helppo ottaa käyttöön, selaimella toimiva eli ei vaadi latausta
Kompose chatbot https://www.komunicate.io/chatbot-templates	Ilmainen testaus	Ei avoin lähdekoodi	Selaimessa, graafinen	Mahdollinen Moodle-integraatio. Ei vaadi koodausta.	Voi kokeilla ilmaiseksi, helppo ottaa käyttöön
RASA https://rasa.com/	Ilmainen perusversio, maksullisia versioita saatavilla	Avoin lähdekoodi (Python, YAML)	Komentotulkilla käytettävä		Monipuolinen alusta, ladattavissa ja kokeiltavissa ilmaiseksi, hyvä dokumentaatio
Ada https://www.ada.cx/	Maksullinen			Kansainvälisille yrityksille. Ei edellytä koodausta.	Maksullinen
Alexa https://developer.amazon.com/en-US/alexa/alexa-	Maksullinen	JavaScript, Java, Python		Monenlaiseen käyttöön.	Maksullinen

skills-kit/get-deeper					
Amazon Lex Framework https://aws.amazon.com/lex/	Maksullinen			Teksti ihmispuheeksi	Maksullinen
Ana.chat https://www.ana.chat/		Github (TypeScript)	Ladattava	Ensimmäinen avoimen lähdekoodin kehitysalusta	Vanhentunut teknologia, ei käytä ilmeisesti tekoälyä lainkaan.
atSpoke				Ei ole enää olemassa, nykyään Okta?	Ei enää olemassa tai vaihtanut nimeä.
Bold360 https://developer.bold360.com/help/EN/Bold360API/Bold360API/Bold_API_Welcome_Header.html	Maksullinen			Tekoälypohjainen. Ei edellytä koodausta.	Maksullinen
Boost.ai https://www.boost.ai	Maksullinen	Ei avoin lähdekoodi	Graafinen	Jyväskylän yliopistolla jo käytössä Minerva-chatbotin taustalla, ei vaadi koodausta	Kallis kaupallinen alusta
Botkit https://github.com/howdyai/botkit	Ilmainen	TypeScript, JavaScript	Ladattava	Osa Microsoft Bot Frameworkia	Väärä käyttötarkoitus: suunniteltu keskustelualustoille.
BotMan https://botman.io/	Ilmainen	GitHub (PHP)	Ladattava, komentotulkillä käytettävä	Keskustelualustoille, avoin lähdekoodi.	Ohjelmointikieli on vieras, käyttöönotto vaatisi paljon resurssia.
Botonic https://botonic.io/	Ilmainen	Avoin lähdekoodi (JavaScript)	Selaimessa, komentoriviltä ohjattava		Liian haastava käyttöönotto.
Botsify https://botsify.com/	Maksullinen			Koulutusikäyttöön suunnattu.	Maksullinen
Botshot https://github.com/botshot/botshot	Ilmainen	GitHub (Python)		Keskustelukäyttöön.	Ei enää tuettu.

Bottender https://bottender.js.org/	Ilmainen	GitHub (TypeScript)		Vain keskustelukäyttöön.	Väärä käyttötarkoitus: suunnattu vain keskustelualueille. Lisäksi ei välttämättä enää tuettu.
ChatGPT https://openai.com/blog/chatgpt/	Ilmainen			Ei ole chatbotin kehitysalusta	Ei sovellu käyttötarkoitukseen
Chatterbot https://chatterbot.readthedocs.io/en/stable/	Ilmainen	Avoin lähdekoodi (Python)	Ladattava		Paljon koodausta vaativa
Claudia Bot Builder https://claudiaajs.com/claudia-bot-builder.html	Ilmainen	GitHub (JavaScript)		Vain keskustelukäyttöön, kirjasto claudia.js:lle.	Väärä käyttötarkoitus: suunnattu keskustelualustoille. Vanhentunut toteutus ja mahdollisesti maksulliseksi muuttuva.
csml.dev https://www.csml.dev/	Ilmainen tiettyyn viestimäärään asti ja rajatuilla ominaisuuksilla. Pro- ja yritysversiot 20\$/kk	Avoin lähdekoodi (Rust)	Selaimessa, koodattava	Oma rakennuskieli chatbotteille, ilmeisesti integroitavissa hyvin joustavasti	Ilmaisversio suppea
Dasha AI https://dasha.ai/en-us	Ilmainen testaus		Selain/ladattava	Ihmismäinen puhekeskustelu.	Liian rajattu: pääosassa keskustelu/ihmismäinen puhe.
DeepPavlov https://deeppavlov.ai/	Ilmainen	GitHub		Syvämpi oppiminen.	Taustatoimija.
Drift https://www.drift.com/	Ilmainen osittain			B2B-yritykset.	Kallis ja yrityskäyttöön suunnattu.
Elomia https://elomia.com/	Ilmainen testaus			Mielenterveys, tekoälyvirtuaaliterapeutti.	Ei kykene luomaan agentteja.
FreshChat https://www.freshworks.com	Ilmainen osittain (tietty määrä agenteja ilmainen)	(no-code)	Selaimessa, graafinen	Ei vaadi koodausta.	Liian yksinkertainen teknologia.
Golem	Ilmainen	GitHub	Python	Ei voimassa -> Botshot	Ei enää tuettu.

IBM Watson https://www.ibm.com/products/watson-assistant/docs-resources	Ilmainen osittain	Valmiita koodoja saavilla sovelluskäyttöön (maksun takana)			Ei palvele käyttötarkoitusta.
Imperson https://www.imperson.com/	Ilmainen rajallinen kokeilu		Selaimessa	AR/VR	Vaikea ottaa käyttöön, osittain maksullinen, vahvasti yrityskäyttöön.
Iris https://iris.ai/	Ilmainen pelkistetty versio, muuten alkaen 75 €/kk			Auttaa tutkimuksen teossa --> käsittelee suuria määriä artikkeleita, ns. "tekoäly-tutkimusavustaja"	Väärä ja liian rajoitettu käyttötarjoitus
Infeedo https://www.infeedo.ai/	Maksullinen			Henkilöstöpalvelu etätyötä tekeville, luo HR:lle raportit.	Maksullinen
Intercom https://www.intercom.com/	Maksullinen	JavaScript, Ruby, Python		Uusi AI-versio, integroitavissa useaan.	Maksullinen
Kasisto https://kasisto.com/	Maksullinen	C#, Java		Talouseläin AI-bot	Maksullinen
Landbot https://landbot.io/	Ilmainen testaus, käyttö 30-300 €/kk		Graafinen	Data-analytiikka sisältävä. Ei edellytä koodausta.	Liian lyhyt ilmainen kokeilu-aika.
LivePerson https://www.liveperson.com/	Maksullinen			AI-keskustelu, NLP	Maksullinen
ManyChat https://many-chat.com/	Maksullinen			Instagram-/Facebook-keskusteluun.	Maksullinen
MedWhat https://medwhat.com/				Terveystsekkaukset, AI.	Rajattu terveydenhoitoalalle, ei mahdollisuutta luoda omaa botia.
Microsoft Bot Framework https://dev.botframework.com/	Ilmainen osittain (Azure Bot Service maksullinen)	Tarjoaa avoimen lähdekoodin alustan	Ladattava	Avoin GitHubista. Vahva taustatoimija, pääsee	Käyttöönotto haastava, monimutkainen kustannusmalli

		(C#, JavaScript, Java, Python)		kokeilemaan ilmaiseksi, nostetaan käsitteeseen, mikäli esiintyy tarve.	
MobileMonkey				Mobiilisovellus (MobileApp), someen suunniteltu.	Vain sosiaaliseen mediaan ja sovelluksiin
Netomi https://demo.netomi.com/learn	Maksullinen			Mainostettuja ominaisuuksia: NLP, ATA ja reinforcement learning. Ei edellytä koodaamista.	Maksullinen
OpenDialog https://github.com/opendialogai	Ilmainen testaus	Avoin lähdekoodi, GitHub	Graafinen	Ei edellytä koodaamista.	Maksullinen
Pandorabots https://home.pandorabots.com/home.html	Ilmainen osittain	Avoin lähdekoodi (AIML)			Liian lyhyt ilmainen kokeiluaika
Paradox https://www.paradox.ai/	Ilmainen			Rekrytointitaroituksiin.	Väärä ja liian rajoitettu käyttötarkoitus
Podium	Maksullinen (300 \$/kk)		Graafinen	Liian kallis	Kallis
ProProfsChat https://www.proprofschat.com/	Maksullinen (13-15 \$/kk)		Graafinen	Ei edellytä koodaamista.	Maksullinen
Replica https://replica.com/	Ilmaisiversion saatavilla	Scala, Python		Juttukaveri, joka oppii keskustelusta.	Väärä käyttötarkoitus: oman keskustelukumppanin luontiin tarkoitettu.
Salesforce/ Einstein https://www.salesforce.com/products/einstein/overview/	Maksullinen (25-300 \$/kk)		Graafinen	Tekoälypohjainen (AI-powered). Ei edellytä koodaamista.	Kallis ja suunnattu eri käyttötarkoitukseen.
SurveySparrow https://surveysparrow.com/	Ilmainen testaus, rajoitettu versio			Kyselybotti, joka ei edellytä koodaamista.	Väärä käyttötarkoitus: suunnattu enemmänkin kyselyiden tekoon.

Tars https://hello-tars.com/	Ilmainen testaus		Graafinen	Monipuolisesti valmiita pohjia eri käyttötarkoituksiin. NLP, API, analytics.	Liian lyhyt testijakso.
Tensorflow https://www.tensorflow.org/	Ilmainen	Python, C++, JavaScript		Kirjasto	Pelkkä kirjasto muiden alustojen hyödynnettäväksi
Theano https://github.com/Theano/Theano o		GitHub (Python)		Kirjasto	Pelkkä kirjasto muiden alustojen hyödynnettäväksi
Tock https://github.com/theopenconversationkit/tock	Ilmainen	GitHub (Kotlin, TypeScript)			Asentaminen koneelle ei onnistunut.
Tidio https://www.tidio.com/	Maksullinen (30-300 €/kk)		Graafinen käyttöliittymä	Ei koodausta	Maksullinen
Wati https://www.wati.io/	Maksullinen			WhatsAppan suunniteltu.	Maksullinen
wit.ai https://wit.ai/	Ilmainen			Kirjaututtava Facebook-tunnuksilla.	Ei voi kokeilla ilman Facebook-kirjautumista
WP-Chatbot	Ilmainen osittain		WordPress-alusta	WordPressiin suunniteltu. Ei edellytä koodaamista.	Suunniteltu vain WordPress-sivustoille. Integraatioiden myötä maksulliseksi muuttuva.
x.ai				Tapaamisten sopimiseen, kalentereihin integroitava.	Väärä ja liian rajattu käyttötarkoitus
Zendesk Answer Bot	Maksullinen			Netomin Zendeskiin syötettävän tiedon varassa toimiva.	Muiden sovellusten varassa oleva, eikä hyödyllinen Zendeskin ulkopuolella.