



Sosiaalisen median botit

Oskar Berggren

Haaga-Helia ammattikorkeakoulu

Tradenomi

Opinnäytetyö

2023

Tiivistelmä

Tekijä(t) Oskar Berggren
Tutkinto Tradenomi
Raportin/Opinnäytetyön nimi Sosiaalisen median botit
Sivu- ja liitesivumäärä 27
<p>Sosiaalinen media on osana ihmisten arkielämää ja toimii työkaluna uutisten seuraamiseen, tuttavien väliseen yhteydenpitoon ja hauskojen videoiden katsomiseen. Lukiessamme uutisia tai toisien käyttäjien julkaisuita, olemme toisten alustoilla asioivien käyttäjien olevan ihmisiä, mitä jos seuraamasi henkilö tai kommentti, ajankohtaisessa uutisessa ei olekaan ihminen, vaan onkin botti?</p> <p>Botit ovat osittain tai täysin automaattisia ohjelmia, joiden yleinen käyttötarkoitus on automatisoida yksinkertaisia tehtäviä. Termi botti on lyhenne sanasta robotti. Ihmisten mielikuvissa robotit ovat kuva tulevaisuudesta, jossa robotit kävelevät keskuudessamme, ajavat autojamme toimivat tehdastyöntekijöinä, joten voidaanko botteja myös käyttää monimutkaisempiin tehtäviin tai voivatko ne toimia ihmisten vastikkeina sosiaalisessa mediassa? Teknologian kehittyminen koneoppimisessa ja kielimallien alueilla, on johtanut valtavaan kehitykseen tekstin ja kuvien luonnissa. Tekstien tai kuvien luontiohjelmia voidaan siis käyttää hyödyksi esimerkiksi boteissa julkaisuiden tekemisen automatisointiin.</p> <p>Tutkimuksessa nousi esiin kaksi ydin tutkimuskysymystä, kykeneekö tavallinen ihminen luomaan botteja sosiaaliseen mediaan ja voiko boteista tehdä uskottavia muiden ihmisten silmissä. Työhön on valittu kolme sosiaalisen median alustaa, jossa botit toteutetaan. Boteissa käytetään Python koodikieltä ja sen useita eri paketteja, joiden avulla bottien toiminnot rakennetaan. Ydin osana bottien toiminnoissa ovat OpenAI:n ChatGPT ja DALL-E kielimallit, joiden avulla boteille luodaan tekstiä ja kuvia. Tärkeänä osana työtä oli käytettyjen teknologioiden ja tekniikoiden saatavuus, joten kaikki työssä toteutetut osat ovat avoimesti saatavilla internetissä.</p> <p>Työn toteuttamisessa keskityttiin kolmeen osaan, bottien luontiin, toimintojen toteutukseen ja bottien automatisointiin. Tavoitteena työssä oli luoda prototyyppi, joka tuottaisi uskottavaa ihmisen luomalta vaikuttavaa sisältöä, mutta ajallisten rajoitteiden takia työssä ei toteutettu prototyyppiä, mutta toteutetut toiminnot toimivat pohjana uskottavan sisällön luonnille. Botit kykenivät luoda ihmismäistä sisältöä, mutta sisällön vaihtuvuus ja konteksti riippuvuus vaatii jatkokehitystä. Tutkimuksessa todistettiin myös, että sosiaalisen median bottien luominen on mahdollista amatööri tason osajalle.</p>
Asiasanat Botti, sosiaalinen media, koneoppiminen, kielimallit

Sisällys

1	Johdanto	1
2	Botit ja niiden taustaa	3
2.1	Bottien historiaa	3
2.2	Bottien hyödyt	3
2.3	Bottien haitat	3
2.4	Botit sosiaalisessa mediassa.....	3
2.5	Bottien tunnistus keinoja	4
3	Käytetyt teknologiat	5
3.1	Python.....	5
3.1.1	Pip	5
3.1.2	PRAW	5
3.1.3	Tweepy	5
3.2	OpenAI.....	5
3.2.1	DALL-E	5
3.2.2	ChatGPT.....	5
3.3	Zapier.....	6
4	Botti verkoston luonti	7
4.1	Tutkimuksen lähtökohdat.....	7
4.2	Henkilöprofiili, Käyttäjätilit ja Alustat	7
4.3	Bottien luonti	7
4.3.1	Reddit bot	7
4.3.2	X Bot.....	10
4.4	Bottien Toiminnot	13
4.4.1	OpenAI	13
4.4.2	Reddit	14
4.4.3	X.....	17
4.5	Automaatio.....	19
4.5.1	Reddit, X ja Schedule	19
4.5.2	Instagram automaatio	21
4.6	Jatko kehitys	23
5	Johtopäätökset.....	25
	Lähteet.....	26

1 Johdanto

Uutisissa on pitkään puhuttu boteista, jotka piileskelevät verkossa ja esittävät ihmisiä sosiaalisessa mediassa. Verkko on täynnä artikkeleita, joissa käsitellään aiheita liittyen botteihin ja tekoälyjen avustamiin käyttäjätileihin. Kuultuani paljon aiheesta päädyin lukemaan ja tutkimaan aihetta syvemmin. Minulle yllätykseksi ilmeni, että työkalut, joita automaatiolla toimivat botit käyttäjät, ovat julkisesti tarjolla, erimäärin heidän omilla alustoillaan. Mikäli työkalut ovat enemmän tai vähemmän julkisesti saatavilla, onko se ihme, että bottien määrä on valtava eri sosiaalisen median alustoilla?

Halusin työssä selvittää, voiko kuka ikinä pientä tekniikka ja koodaus taitoja omaava ihminen luomaan julkisen tiedon ja teknologian avulla botteja. Mitä mahdollisia ominaisuuksia ja toimintoja voisivat botit sisältää? Viimeisinä vuosina koneoppimisen avulla kehittyneet tekstin ja kuvien luontipalvelut ovat tehneet mullistuksia, jokaisen ihmisen elämään ja koska iso osa tästä teknologiasta on julkisesti käytettävissä, halusin selvittää, kuinka teknologiaa voitaisiin hyödyntää mahdollisesti boteissa, luomaan jatkuvaa sisältöä, kuvien ja tekstien muodossa.

Ydin tutkimuskysymyksessä työssä on ”uskottavan” botin luomien. Työssä uskottavalla botilla tarkoitetaan botia, joka toimii itsenäisesti julkaisemalla sisältöä ja joka on interaktiivinen alustojen sisällä sekä niiden kesken. Botin luoman sisällön on tärkeää olla satunnaista, mutta loogisesti johdannollista botin profiiliin kontekstissa.

Työssä käytettävät teknologiat ja alustat on valittu niin, että ne ovat ilmaisia ja julkisesti käytettäviä kaikille. Tutkimuksessa oleellisena osana oli siis tutkia botia, automaatio ja koneoppimisen teknologian saatavuutta, joten työssä ei käytetty julkisesti rajoitettuja/maksullisia ominaisuuksia tai teknologioita.

Tutkimuksen aihe on laaja ja haastava, joten sen skaala on rajoittunut työn edetessä teknisistä ja ajallisista syistä, mutta työn arvo on pysynyt korkeana, koska tutkimuksen tuloksia ja parempaa ymmärrystä boteista voivat hyödyntää niin yksityishenkilöt tai organisaatiot. Tutkimus tarjoaa tärkeän pohjan aiheen ymmärrykselle.

Keskeisiä käsitteitä:**Ohjelmarajapinta (API)**

On ohjelmien käyttämä sarja sääntöjä, jotka määräävät miten ohjelma keskustelee muiden ohjelmien kanssa. API mahdollistaa ohjelman datan jakamisen turvallisesti ulkopuolisille ohjelmille rajaamalla datan jakamisessa käytettävät säännöt. API:n avulla on mahdollista integroida ohjelmia ja komponentteja toimimaan keskenään. (Hoffman 27.12.2021)

Koneoppiminen

On tekoäly teknologian alamuoto, jossa ohjelmia koulutetaan tunnistamaan kuvioita ja tekemään ennustuksia kuvioiden avulla. Ohjelmat muodostavat kuvioita ohjelmalle syötettävän tiedon avulla, ohjelman käyttämä tieto voi olla esimerkiksi tekstiä, kuvia tai numeroita. Kerätyn tiedon avulla koulutettua ohjelmaa kutsutaan kielimalliksi. Valmiita kielimalleja voidaan käyttää esimerkiksi tekstin tai kuvien luontiin. (Brown 21.4.2021)

Sosiaalisen median alustat

Ovat internetin kautta toimivia verkkosivuja tai verkkoalustoja, jotka mahdollistavat käyttäjien keskeisen viestinnän ja sisällön luomisen/jakamisen. Verkkoalustojen päätoimintona on luoda ja jakaa tietoa julkisesti alustoiden sisällä käyttäjien kesken, mutta alustat mahdollistavat myös käyttäjien välisen yksityisen viestinnän. Käyttäjien luoma sisältö voi esiintyä teksti, kuva, ääni ja video muodoissa, riippuen käytettävästä alustasta. (Britannica 2023)

2 Botit ja niiden taustaa

Botteja voidaan ajatella automaattisesti toimivina ohjelmistoina, jotka suorittavat eri tehtäviä, riippuen botin halutusta käyttötarkoituksesta. Arkielämässä on yleistä nähdä useita eri botteja päivittäin, esimerkiksi verkkokauppojen asiakaspalvelu viesteissä tai sosiaalisen median moderoinnissa. (Amazon Web Services)

2.1 Bottien historiaa

Ensimmäisiä botteja aloitettiin kehittämään jo ennen internetiä 1950-luvulla, esimerkkinä tästä on ELIZA niminen chatbot. Internetin saapuessa 1980-luvulla, ensimmäisiä verkossa toimivia botteja käytettiin Internet Relay Chat (IRC) kanavilla eri tehtävien automatisointiin. 1990-luvulla ilmestyivät ensimmäiset bottiverkot, jotka käyttivät kaapattuja koneita esimerkiksi tiedon varastamiseen ja levittämiseen. Toisena suurena 90-luvun botti mullistuksena, olivat web-kaappaajat, kuten Googlebot, joita hyödynnetään edelleen verkkoselainten haku tehokkuuden kehittämässä. (Potter 15.2.2021)

2.2 Bottien hyödyt

Botteja voidaan hyödyntää monissa osissa ihmisten elämää. Ydin hyöty bottien käytöstä löytyy yksinkertaisten tehtävien automatisoinnissa, jolloin ihmiset voivat keskittyä oleellisempiin tai monimutkaisempiin tehtäviin. Botteja voidaan hyödyntää esimerkiksi 24/7-palveluiden ylläpidossa, asiakaspalveluviestinnässä, tiedon keräämisessä tai ohjelmisto tai verkkotestauksen automatisoinnissa. (Stefanowicz 19.10.2023)

2.3 Bottien haitat

Botteilla ja niiden käyttötapauksilla on heikkoudet ja huonot puolensa. Botteja voidaan usein käyttää haittaohjelmien levittämiseen tai salaisen tiedon keräämiseen. Bottien käytössä ongelmallista on myös se, että niitä täytyy jatkuvasti kehittää ja ylläpitää, jotta ne pysyvät kunnollisesti toiminnassa. Valheellisen tiedon ja uutisten levittämistä on tehty pitkää hyödyntämällä bottien automaatio tekniikkaa, tekemällä kommentteja tai julkaisuita. (NICCS 2018)

2.4 Botit sosiaalisessa mediassa

Sosiaalisen median botit ovat sosiaalisen median sisällä toimivia botteja, jotka voivat toimia osittain tai täysin itsenäisesti. Botteja nähdään kaikissa sosiaalisen median alustoissa, mutta alustoilla nähdään eri lähestymistapoja bottien käyttöön. Useat alustat tarjoavat esimerkiksi tukevat botteja ja tarjoaa käyttäjille työkaluja niiden luomiseen ja käyttämiseen. Sosiaalisen median botteilla voidaan simuloida ihmisten käytöstä, jolloin niitä voidaan hyödyntää valetiedon, poliittisten

ideologioiden tai teollisen suosion levittämisessä. Botit voivat usein käyttää profiileja, jotka ovat tarkoituksen mukaisesti tehty näyttämään esimerkiksi kuuluisalta julkisuuden henkilöltä, kuten poliitikon käyttäjältä, jotta botti saisi uskottavuutta. (NICCS 2018)

2.5 Bottien tunnistus keinoja

Botti käyttäjä profiileja voidaan yrittää tunnistaa useilla eri tavoilla, tarkistelemalla epäilyttävän käyttäjän profiilia, seuraamalla mitä profiili julkaisee, ketkä ja ketä profiilia seuraavat. Useasti botti profiilit sisältävät geneerisiä profiilikuvia ja niiden nimet eivät aina ole ihmisten nimiä, vaan sekoitus satunnaisista sanoista ja numeroista. Toisena keinona voidaan seurata profiilin aikalinja toimintaa ja selvittää, kuinka usein tai harvoin profiili tekee päivityksiä, jos esimerkiksi käyttäjä tekee julkaisuja tauotta tai todella tiheään tahtiin, voi se olla merkki botti profiilista. (Combs 1.9.2020)

3 Käytetyt teknologiat

3.1 Python

Python on suosittu ohjelmointikieli, jota kutsutaan yksinkertaiseksi sen helpon luettavuuden takia. Python on tulkattu ohjelmointikieli, joka tarkoittaa sitä, että koodia ei tarvitse kääntää konekieleen, koska Python lukee ja suorittaa koodin suoraa lähdekoodista. Python kieltä käytetään usein skripteinä eri komponenttien yhdistämiseen. Kieli tukee myös moduuleja sekä paketteja, joita voidaan hyödyntää luomalla ne itse tai käyttämällä valmiiksi luotuja julkisia paketteja koodissa. (Python 2023)

3.1.1 Pip

Pip on pakettien hallintatyökalu Pythonille, sen avulla voidaan asentaa kirjastoja ja paketteja, jotka eivät sisälly Pythonin standardi kirjastoihin. (PyPI 2023)

3.1.2 PRAW

PRAW "Python Reddit API Wrapper" on Python paketti, joka mahdollistaa pääsyn Reddit palvelun ohjelmistorajapintaan. (PyPI 2023)

3.1.3 Tweepy

Tweepy on Python paketti, joka mahdollistaa pääsyn X palvelun ohjelmointirajapintaan. (Mahajan 2.11.2020)

3.2 OpenAI

OpenAI on tutkimuslaboratorio, jonka tavoitteena on tutkia ja kehittää tekoälyteknologiaa. Yritys on tuottanut useita eri koneoppimis kielimalleja, kuten GPT-3 ja GPT-4. (Hashemi-Pour 8.11.2023)

3.2.1 DALL-E

DALL-E on OpenAI:n luoma tekoälyllä toimiva kuvageneraatiomalli, jonka avulla on mahdollista luoda kuvia antamalla sille tekstimuotoisia kuvauksia, jonka pohjalta DALL-E luo kuvaukseen liittyvän kuvan. (Sheen 23.3.2023)

3.2.2 ChatGPT

ChatGPT on OpenAI:n kehittämä suuri kielimalli, joka perustuu GPT 3.5 arkkitehtuuriin. Toiminnallisesti ChatGPT pystyy vastaamaan kysymyksiin ja tuottamaan kysymykseen vastauksen käyttäen koneoppimismalleja. (Drapkin 28.10.2023)

3.3 Zapier

Zapier on verkkopohjainen automaatiopalvelu, jonka avulla yhdistetään eri verkkosovelluksia ja automatisoidaan niiden toimintoja. Palvelussa käytetään työnkulkuja, jotka sisältävät erillisiä tehtäviä. Tehtäville voidaan määrittää järjestys ja jokainen tehtävä sisältää triggerin, joka laukaisee seuraavan tehtävän. (Logreira 10.4.2023)

4 Botti verkoston luonti

4.1 Tutkimuksen lähtökohdat

Tutkimuksen tarkoituksena on saavuttaa tietämys ja ymmärrys bottien luomisesta ja käyttämisestä sosiaalisessa mediassa. Tutkimuksen ensimmäisenä vaiheena rajataan alustat, joissa automaatiota tai botteja tullaan käyttämään. Rajauksen jälkeen seuraavana vaiheena on aloittaa alusta kohteisten bottien luonti. Botit eivät ensimmäisessä versiossaan sisällä muita ominaisuuksia, kuin ensimmäisen julkaisun luonnin. Seuraavassa vaiheessa boteille tuotetaan niille haluttavat toiminnot. Viimeisessä vaiheessa tutkimusta tavoitteena on saada botit luomaan jatkuvaa sisältöä automaattisesti ja itsenäisesti. Kokonaisuuden valmistuttua tehdään johtopäätökset bottiverkoston toteuttamiskelpoisuudesta.

4.2 Henkilöprofiili, Käyttäjätilit ja Alustat

Bottien ja automaation toteuttamiseen eri sosiaalisen median alustoilla vaaditaan useita eri käyttäjätiliä ja ohjelmistorajapinnoissa vaadittuja API avaimia. Työssä on käytetty yhtä keskeistä sähköpostia, jota eri alustat käyttävät. Työn kulussa tarvitaan Reddit, X, Instagram Business ja Facebook tilit. Tilit ovat keskitetty google tiliin, jota käytetään valtuutusten antamiseen ja tilien palalutusta varten. Eri alustojen käyttäjätilit käyttävät google tiliin pohjautuvia käyttäjätietoja, kuten nimet ja syntymäajat.

4.3 Bottien luonti

Valittaessa bottien käytettäviä alustoja, tuli tehdä alustava selvitys bottien käytön sallittavuudesta eri vaihtoehtoisilla alustoilla. Mikään kohteena ollut alusta ei kiellä bottien käyttöä alustoillaan, mutta määrittää rajat niiden käyttötarkoituksille, esimerkiksi kuinka usein julkaisusta tykätään tai kuinka useaa käyttäjää botti voi seurata tietyllä aikavälillä. Tutkimuksessa olevat alustat päämääräisesti tukevat bottien käyttöä ja tarjoavat kattavia ohjeita bottien käytöstä heidän alustoillaan. Bottien luonnin aikana tukimateriaalina pystytään käyttämään alustakohtaisia kehittäjä ohjeita

4.3.1 Reddit bot

Botin luonti Reddit alustalla vaatii kolme ydin elementtiä. Ensimmäiseksi vaaditaan käyttäjätili, jota botti käyttää sen eri toiminnoissa, kuten julkaisuissa, kommentoinnissa tai julkaisujen tykkäyksissä. Toisena ydin elementtinä toimii API ohjelmointirajapinta, jonka avulla botti pääsee käsiksi palvelun dataan ja pystyy todentamaan Redditiille ohjelman toiminnan sallittavuuden. Kolmantena osana toimii Python koodi, joka mahdollistaa eri haluttujen toimintojen toteuttamisen.

Käyttäjätilin luonnin jälkeen tili pitää rekisteröidä Redditin kautta ohjelmaksi, saadakseen pääsyn Redditin ohjelmointirajapintaan. Rekisteröinti toteutetaan palvelun tarjoaman lomakkeena kautta. Lomake on esitetty kuvassa (Kuva 1). Lomakkeessa tulee täyttää ohjelman nimi, tyyppi, kuvaus ja uudelleenohjaus url. Tutkimuksen kohdalla ohjelma tyyppinä toimii skripti ja ohjelma uudelleenohjataan työaseman omaan verkko-osoitteeseen.

GET NEW REDDIT | OMAT ALIREDDITIT | KOTI - POPULAR - KAIKKI - SATUNNAINEN - USERS | TECHNOLOGY - ANNO

reddit OMAT ASETUKSET

valinnat | **sovellukset** | RSS-syötteen | kaverit | estetty | salasana/sähköpostiosoi Sebast

luo sovellus

By creating an app, you agree to Reddit's Developer Terms and Data Api Terms. You must also register


nimi

verkkosovellus Verkkopohjainen sovellus
 asennettu sovellus Asennettavaksi tarkoitettu sovellus, esimerkiksi kännykässä
 skripti Skripti henkilökohtaiseen käyttöön. Sisältää pääsyn vain kehittäjien tileille

kuvaus

tietoja url:stä

uudelleenohjaus url

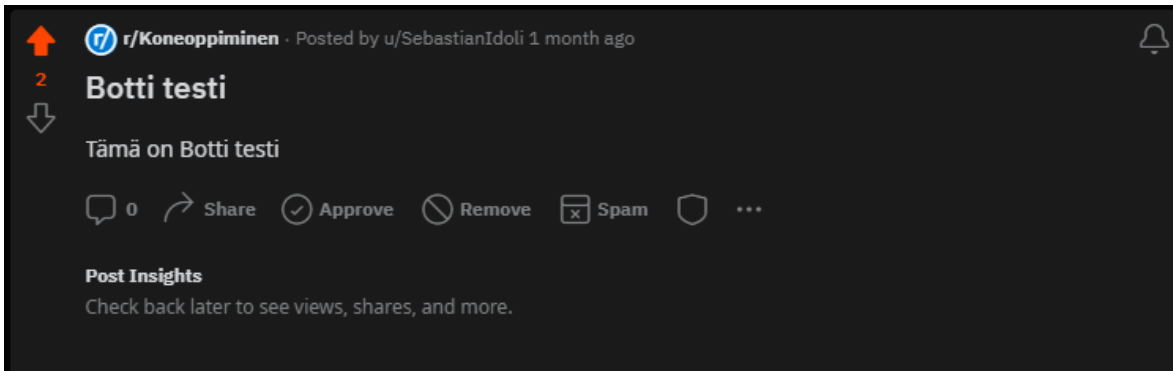
En ole robotti
 

reCAPTCHA
Tietosuojia - Ehdot

Kuva 1. Reddit sovellus lomake

Botin rekisteröinnin jälkeen saadaan tarvittavat API-tunnistetiedot, joilla botti saa tarvittavat valtuudet keskustella Reddit ohjelmointirajapinnan kanssa. Tunniste tietoina toimii `client_id` ja `client_secret`. `Client_id` on julkinen merkkijono, jonka avulla ohjelma tunnistetaan. `Client_secret` toimii luodun ohjelman salasana.

Ennen koodiin kirjoittamista tulee järjestelmään asentaa Python kirjasto, joka yksinkertaistaa botin ja Redditin väliset interaktiot, käytettävänä kirjastona toimii PRAW. Kirjasto asennetaan pip paketti asentajan avulla komentokehotteen kautta. Asennus tapahtuu ajamalla komento "pip install praw" komentokehotteessa.



Kuva 4. Reddit botti julkaisu

Yksinkertaisen Reddit julkaisu botin luonti on suoraviivaista ja ei vaadi suurta tietämystä ohjelmoinnista tai syvää käsitystä esimerkiksi ohjelmointirajapinnoista tai verkkoliikenteestä.

4.3.2 X Bot

X alustalla botin luonti vaatii lähtökohtaisesti saman lähestymistavan, kuin Reddit alustalla, eli toimivan botin luontiin vaaditaan käyttäjätili, API valtuudet ja koodi, joka ajaa haluttavat toiminnot. Ensimmäisenä vaiheena luodaan X käyttäjä, jolle hankitaan kehittäjä oikeudet. Kehittäjä oikeuksien hankkiminen on askeleen monimutkaisempaa, kuin Reddit alustalla ja vaatii kirjallisen selityksen siitä, miksi ja mihin tarkoitukseen vaadit pääsyn X palvelun ohjelmistorajapintaan. Kirjallisessa selityksessä tulee kertoa luotavien ohjelmien mahdolliset ominaisuudet, haluatko käyttää X palvelun dataa, onko ohjelman käyttötarkoitus kaupallista, akateemista tai henkilökohtaista. Kirjallisen selityksen jälkeen saadaan heti pääsy kehittäjä työkaluihin, mutta mikäli huomataan, että ohjelmaasi käytetään selityksessä annettujen toiminnallisuuksien lisäksi muita ominaisuuksia, voidaan käyttäjäoikeutesi kieltää.

Saatuasi kehittäjäoikeudet pääset käsiksi X palvelun ympäristöön, jossa voit luoda projekteja, joiden alla toimii yksittäiset ohjelmat, jotka toimivat botteina. X palvelu tarjoaa kolme eri käyttäjäluokkaa, joista yksi on ilmainen. Käyttäjäloukkien erot ilmenevät sallittujen projektien määrästä, kuinka monta julkaisua voit tehdä kuukaudessa ja mitä toimintoja ohjelmasi voi tehdä. Ilmaisella käyttäjäluokalla voidaan tehdä 1500 julkaisua kuukaudessa ja sinulla voi olla vain yksi projekti kerrallaan käytössäsi. Ohjelmalle tulee myös asettaa tarvittavat oikeudet kuvan (Kuva 5) lomakkeessa, jotta se voi tehdä ohjelman tarvitsevia toimintoja. Ohjelmalle voidaan määrittää oikeudet lukea, kirjoittaa ja lähettää suoria viestejä. Asetuksissa määritellään myös ohjelman tyyppi.

sebastian_tweet_bot

User authentication settings

You can change these selections anytime.

App permissions (required)

These permissions enable OAuth 1.0a Authentication. ⓘ

- Read
Read Tweets and profile information
- Read and write
Read and Post Tweets and profile information
- Read and write and Direct message
Read Tweets and profile information, read and post Direct messages
- Request email from users
To request email from users, you are required to provide URLs to your App's privacy policy and terms of service.

Type of App (required)

The type of App enables OAuth 2.0 Authentication. ⓘ

- Native App ⓘ
Public client ⓘ
- Web App, Automated App or Bot ⓘ
Confidential client ⓘ

Kuva 5. X sovellus asetukset

Seuraavana vaiheena luodaan ohjelmistorajapintaan käsiksi pääsyyn vaadittavat valtuutusavaimet. Tässä vaiheessa luodaan viisi eri avainta, jotka ovat `consumer_key`, `consumer_secret`, `bearer_token`, `access_token` ja `access_token_secret`. `consumer_key` ja `consumer_secret` toimivat ohjelman tunnistus id:nä ja salasanana. `Bearer_token` on valtuutusavain, jota ohjelma käyttää tehdessään pyyntöjä suoraan X palvelun ohjelmistorajapintaan. `Access_token` ja `access_token_secret` avainten tarkoitus on saada lupa käyttää käyttäjän dataa, ohjelman keskustellessa X ohjelmistorajapinnan kanssa.

Valtuutusten hankkimisen jälkeen asennetaan komentokehotteen ja pip pakettien asentajan avulla Tweepy kirjasto, joka toimii välikätenä käyttäjän ja ohjelmistorajapinnan välillä.

Koodin kirjoittaminen alkaa importoimalla Tweepy ja määrittelemällä valtuutusavaimet, jonka jälkeen ohjelma valtuutetaan määritellyillä avaimilla. Seuraavaksi voidaan luoda haluttu toiminto, joka on tässä tapauksessa yksinkertaisen julkaisun tekeminen ohjelman avulla. Osion koodi näkyy kuvassa. (Kuva 6)

4.4 Bottien Toiminnot

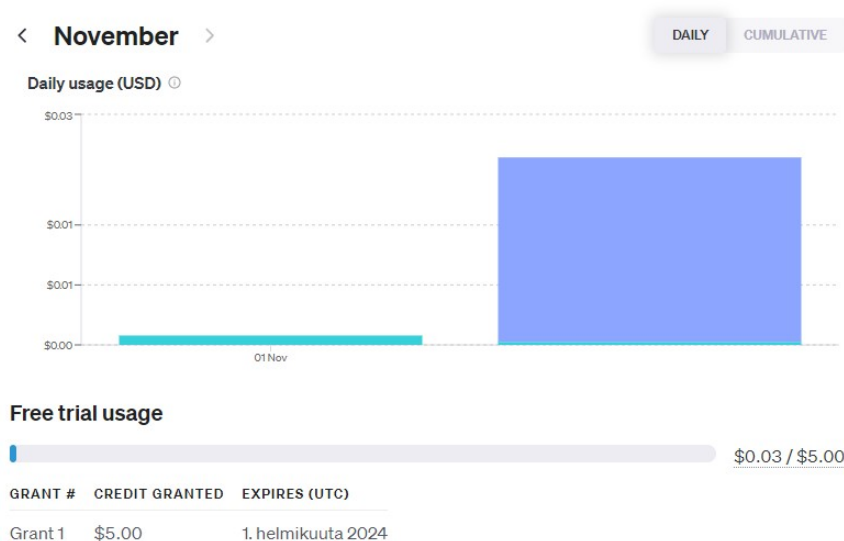
Tässä osiossa työtä lisää ominaisuuksia aikaisemmin luoduille boteille. Tavoitteena tässä vaiheessa on demonstroida bottien mahdollisia käyttötarkoituksia, itsenäisen sisällön luonnissa. Botteihin lisätyt ominaisuudet eivät luo kokonaiskuvaa itsenäisestä henkilöstä, mutta antavat näytön siitä, kuinka uskottavan henkilön luonti on mahdollista, käyttäen toteutettuja toimintoja.

4.4.1 OpenAI

Kaikkien alustojen julkaisuihin tullaan käyttämään OpenAI:n "gpt-3.5turbo" kielimallia, tuottamaan sisältöä bottien tekemiin julkaisuihin. Tässä osiossa luodaan siis koodi, joka antaa kielimallille kehoitteen ja saa takaisin siltä vastauksen. Kielimallilta saatuja vastauksia hyödynnetään julkaisuiden sisältönä. Päästäkseen hyödyntämään OpenAI:n kielimalleja koodissa, tarvitaan sen käyttöön vaadittava API avain. Avaimen saamiseen tarvitaan OpenAI käyttäjätunnukset. Käyttäjätunnusten luomisen jälkeen API avain voidaan luoda ilman lisä vaatimuksia OpenAI:n sivuilta. Open AI ohjelmistorajapinnan käyttö on ensirekisteröinnin jälkeen kolme kuukautta ilmaista, viiden dollarin hintakattoon asti. Sivuilta pääsee kätevästi näkemään API käyttömääräsi päiväkohtaisesti. Käyttöliittymä näkyy kuvassa. (Kuva 8)

Usage

Below you'll find a summary of API usage for your organization. All dates and times are UTC-based, and data may be delayed up to 5 minutes.



Kuva 8. OpenAI datan käytön, käyttöliittymä

importoitu random niminen paketti, jota hyödynnetään koodin funktiossa. Funktion tarkoituksena on palauttaa satunnainen esine listalta. Viimeiseksi koodissa lisätään funktion palauttama arvo muuttu-
tujaan, jota hyödynnetään botin päätiedostossa.

The image shows a code editor window with two tabs: 'list.py' and 'testi.py'. The active tab is 'list.py', which contains the following Python code:

```
Koodi > list.py > ...
1  import random
2
3  def get_random_language(language_list):
4      if language_list:
5          return random.choice(language_list)
6
7
8  coding_languages = [
9      "Python",
10     "Java",
11     "C++",
12     "JavaScript",
13     "C#",
14     "Ruby",
15     "Swift",
16     "Go",
17     "R",
18     "TypeScript",
19     "PHP",
20     "Kotlin",
21     "Objective-C",
22     "Perl",
23     "Scala",
24     "Rust",
25     "HTML",
26     "CSS",
27     "SQL",
28     "Assembly"
29 ]
30
31 random_language = get_random_language(coding_languages)
32
33
34
```

Kuva 10. Koodikieli lista

Tässä vaiheessa päätiedostoon importoidaan list.py (Kuva 10) tiedosto ja lisätään muuttuja arvot, joita käyttäen annetaan kielimallille kehoite, joka luo julkaisuun otsikon ja tekstin. Muuttujien merkki-
jonoissa annetaan kielimallille käsky luoda viikoittainen fakta satunnaisesti valitusta koodauskie-
lestä. Seuraavaksi määritellään muuttujat post_title ja post_body, jotka sisältävät tekstin luonti
funktion hyödyntäen aikaisemmin luotuja muuttujia, kuvan (Kuva 11) mukaisesti.

```

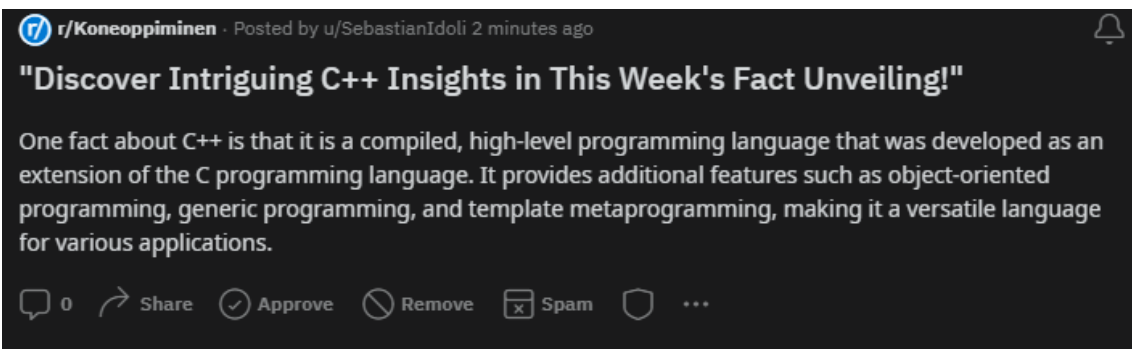
heading = f"give an non specific post title about an weekly {list.random_language} fact"
body = f"give a fact about {list.random_language}"

post_title = text_generation.prompt_text(heading)
post_body = text_generation.prompt_text(body)

```

Kuva 11. Reddit botti koodi, tekstin luonti ja julkaisu funktio

Koodi on nyt valmis ja se voidaan ajaa. Tulos näkyy julkaisuna Reddit keskustelupalastalla. Kuvassa (Kuva 12) näkyy botin tekemä julkaisu.



Kuva 12. Reddit koodikielijulkaisu

Toisena botin ominaisuutena botille luodaan kommentointi toiminto, joka valitsee aikaisemmin luodusta listasta (Kuva 10) kielen ja käy läpi keskustelupalstan viisi suosituinta julkaisua, jos missään viiden suosituimman julkaisun sisällä mainitaan satunnaisesti valittu koodauskieli, botti kommentoi julkaisua. Koodissa luodaan muuttujat subreddit ja top_post, jotka määrittävät kohde palstan ja mitä julkaisuja botti käy läpi. Seuraavaksi luodaan funktio, joka skannaa julkaisut ja mikäli julkaisun tekstistä löytyy yhteensopiva nimi, botti kommentoi julkaisua. Valmis koodi on esitetty kuvassa (Kuva 13).

```

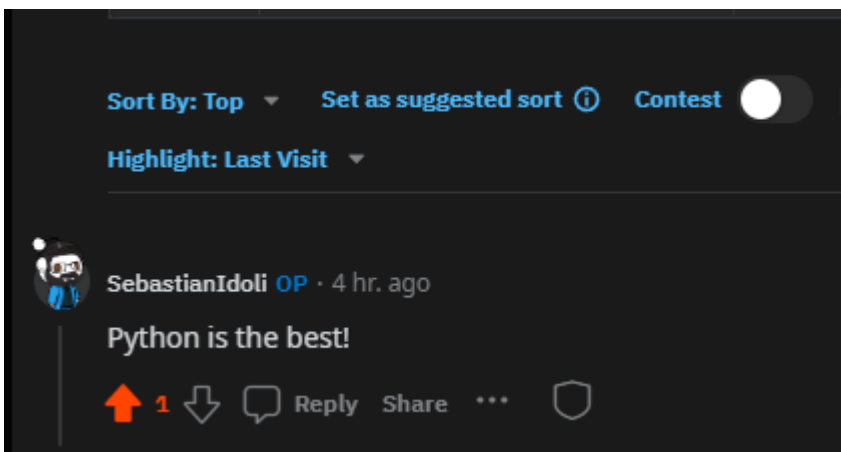
subreddit = reddit.subreddit("Koneoppiminen")
top_post = subreddit.top(limit=5)

def comment_post():
    for post in top_post:
        if list.random_language in post.selftext:
            post.reply(f"{list.random_language} is the best!")
            print("Kommentoitu")
            print(list.random_language)

```

Kuva 13. Reddit botin koodi, kommentointi

Seuraavaksi koodi ajetaan ja tuloksena voidaan nähdä kommentteja julkaisuiden alla. Valmis julkaisu näkyy kuvassa (Kuva 14).



Kuva 14. Reddit botin kommentti

4.4.3 X

X palvelun toimintojen luonti on tutkimuksessa rajallista, sillä isoin osa ohjelmistorajapinnan kautta käytettävistä ominaisuuksista ovat maksullisia. Palvelun ilmainen kehittäjä luokka tarjoaa vain julkaisuiden tekemisen ja käyttäjätiedon osittaisen keräämisen. X palvelun maksullisissa käyttäjäluokissa on mahdollista kerätä käyttäjien julkaisut ja niiden ID arvot, joka mahdollistaisi esimerkiksi valittujen käyttäjien julkaisuiden uudelleen julkaisun. Ohjelmistorajapinnan käyttörajoitteiden takia botin toiminnallisuutena on julkaisuiden tekeminen, käyttäen Reddit botissa olevaa rakennetta.

Botille tehtävän toiminnon tarkoituksena on saada botti siteeraamaan satunnaisesti valittua poliitikoita tai filosofiä. Ensimmäisenä luodaan Python kielen muodossa oleva lista kuuluisista filosofeista ja poliitikoista. Koodissa käytetään Reddit osiossa luotua list.py (Kuva 10) tiedostoa pohjana ja vaihdetaan koodissa käytettävä listan sisältö, osa uudesta sisällöstä näkyy kuvassa (Kuva 15). X botin päätiedostossa tehdään muuttuja, joka luo kehoitteen kielimallille. Viimeiseksi tehdään kuvan (Kuva 16) mukainen julkaisu funktio, jossa määritellään tekstin sisältö olemaan kielimallin vastaus muuttuja kehoitteeseen.

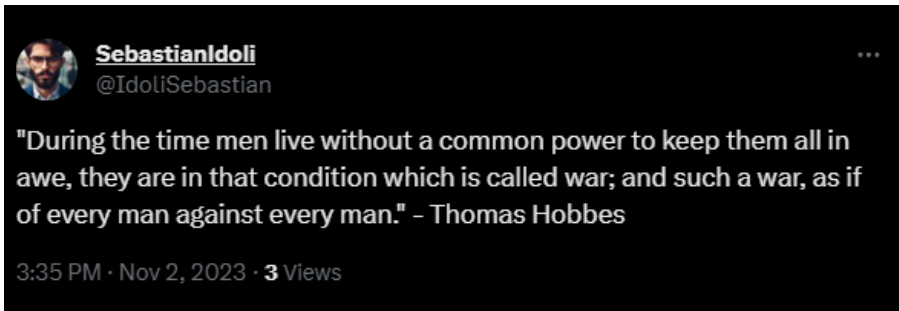
```
names = [  
    "Aristotle",  
    "Plato",  
    "Confucius",  
    "Socrates",  
    "Karl Marx",  
    "Friedrich Engels",  
    "John Locke",  
    "Jean-Jacques Rousseau",  
    "Voltaire",  
    "Thomas Hobbes",  
    "Immanuel Kant",
```

Kuva 15. lista filosofien nimiä

```
19 |  
20 | body =f"tell a quote from {list.random_name}"  
21 |  
22 | def tweet():  
23 |     client.create_tweet(  
24 |         text = tweet_content.tweet_text(body)  
25 |     )  
26 |
```

Kuva 16. X botti koodi, filosofi julkaisu

Botin toimivuudesta varmistutaan ajamalla ohjelma ja menemällä X käyttäjän aikalinjalle, jossa näkyy kuvan (Kuva 17) mukainen julkaisu.



Kuva 17. X botti filosofi julkaisu

4.5 Automaatio

Automaatio osion tarkoituksena on tehdä toimintojen ajamisesta automaattista ja käyttäjästä itsenäistä. Automaatiot toteutetaan Python koodilla ja Zapier automaatio palvelun avulla. Toimintojen automaatio määritellään joko päivittäiseksi tai viikottaiseksi.

4.5.1 Reddit, X ja Schedule

Tässä osiossa automatisoidaan Reddit ja X botit ajamaan toimintonsa itsenäisesti niille, määrättyinä aikoina. Tavoitteena on hyödyntää Python schedule ja random paketteja, joiden avulla määritellään halutut tehtävät ajamaan niille määrättyinä aikoina. Viimeiseksi aikataulutettujen tehtävien tulee ajaa satunnaisina kellon aikoina. Satunnaisen kellon ajan tarkoituksena on antaa kuva siitä, että ihminen ajaa toiminon, eikä se tapahdu samaan aikaan toistuvasti.

Reddit botin julkaisut ja kommentoinnit tapahtuvat satunnaisina kellon aikoina kerran viikossa ja X botti tulee tekemään julkaisun kerran päivässä. Reddit toimintojen automatisointi aloitetaan luomalla Python tiedosto, jota kautta aikataulutetut tehtävät ajetaan. Ensimmäisenä osana koodia importoidaan aikataulutukseen ja toimintojen ajamiseen tarvittavat kirjastot. Seuraavaksi määritellään kommentti ja julkaisu toiminnot tiedoston sisällä omiin funktioihinsa, joita käytetään aikataulutus funktiossa. Kolmantena vaiheena luodaan itse aikataulutus funktio, jossa määritellään muuttujiin random kirjaston avulla satunnaiset numerot tunneille ja minuuteille. Muuttujien määrittelyn jälkeen tehdään aikataulutus funktio, joka ajaa toiminnot maanantaisin muuttujilla määritettyyn kellon aikaan. Funktion sisällä tulostetaan myös tehtävä ja sen ajamis- kellonaika. Viimeisessä osassa koodia on While True silmukka, joka tarkastaa sekunnin välein ajettavat tehtävät ja ajaa toiminnot niille määrättyinä aikoina. Valmis koodi nähdään kuvassa (Kuva 18).

```
def post():
    sebastian_bot.reddit_post()
def comment():
    comment_bot.comment_post()

def weekly_post(task):
    hour = random.randint(12,20)
    minute = random.randint(10,59)
    schedule.every().monday.at(f"{hour}:{minute}").do(task)
    print(f"Task{task} at schedulet to {hour}:{minute}")
    print(schedule.get_jobs())

weekly_post(post)
weekly_post(comment)

while True:
    schedule.run_pending()
    time.sleep(1)
```

Kuva 18. Reddit botti koodi, automaatio

X palvelun automatisointi toimii sammalla metodilla, kuin Reddit botti, mutta ajaa toimintonsa ker-
ran päivässä. X botin koodi nähdään kuvassa (Kuva 19).

```
def tweet():
    sebastian_tweet_bot.tweet()

def daily_tweet(task):
    hour = random.randint(12,20)
    minute = random.randint(10,59)
    schedule.every().day.at(f"{hour}:{minute}").do(task)
    print(f"Tweeted at {hour}:{minute}")
    print(schedule.get_jobs())

daily_tweet(tweet)
```

Kuva 19. X botti koodi, automaatio

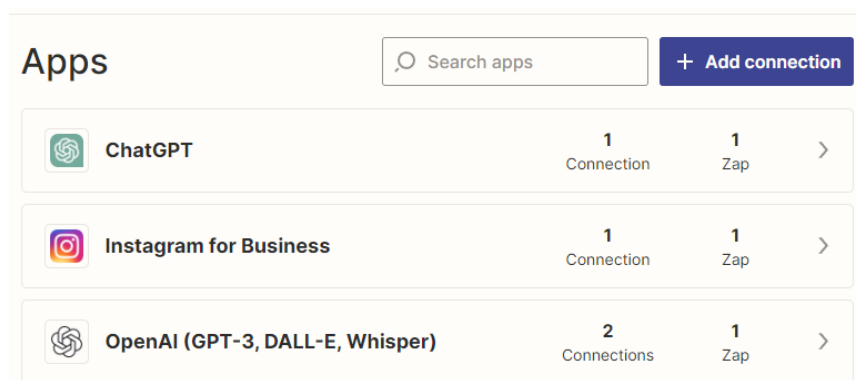
Reddit ja X botit voidaan nyt ajaa, jonka jälkeen ne tekevät toimintonsa automaattisesti, niin pit-
kään kunnes työasema suljetaan.

4.5.2 Instagram automaatio

Instagram/Meta ovat tarkoituksellisesti halunneet rajoittaa bottien toimintaa alustoillaan. Toisin kuin Reddit tai X, Meta tarjoaa rajoitetumpaa pääsyä heidän ohjelmistorajapintaansa ja valtuutusten hankkiminen on hankalaa. Julkisesti tarjolla olevia Python botti ohjelmia löytyy, mutta suurin osa niistä eivät ole tämän työn teon aikana ajantasolla. Instagram palveluun on mahdollista luoda botteja SeleniumWebDriver kehikon kautta, mutta sen käyttö ei ole suositeltavaa ja se voi johtaa käyttäjätilien sulkemiseen. Botin luomisen sijasta tässä työssä käytetään julkisia automaatio palveluita, joiden toimintaa Meta tukee. Tutkimuksessa käytettävän palvelun nimi on Zapier.

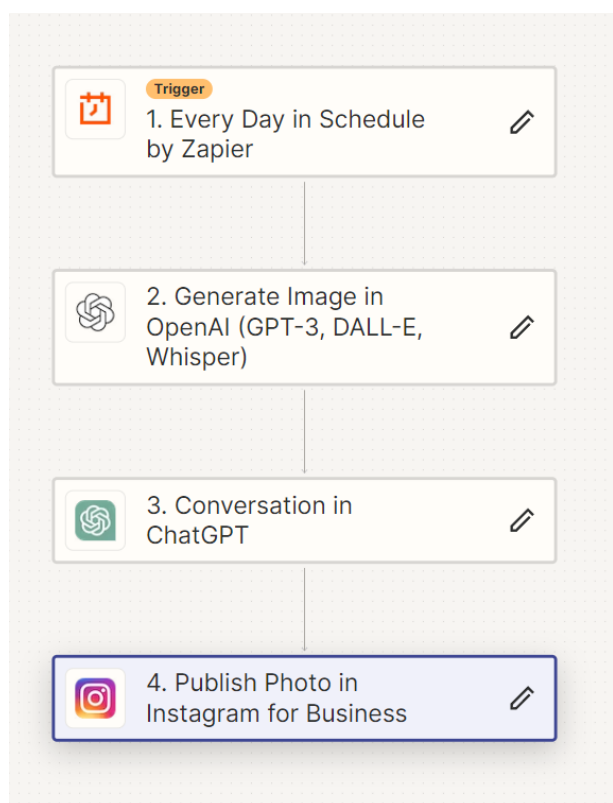
Zapier on automaatiopalvelu, jonka avulla voidaan yhdistää ja automatisoida eri palveluiden toimintoja. Palvelu hyödyntää laukaisu ja toiminta ominaisuuksia, joilla voidaan luoda työkulkuja tehtäville. Palvelussa on mahdollista luoda esimerkiksi työkulku, joka seuraa valittua henkilöä sosiaalisessa mediassa ja henkilön uusien julkaisujen jälkeen työkulussa lähetetään sähköposti haluttuun kohteeseen. Tässä työosassa luon työkulun, jossa tehdään päivittäinen kuva julkaisu Instagram palveluun. Julkaisussa oleva kuva luodaan OpenAI:n DALL E kuvangenerointimallilla ja julkaisun teksti luodaan ChatGPT:n avulla.

Zapier palvelun käyttämiseen tarvitaan käyttäjätili, joka sisältää sähköpostin, tässä tapauksessa käytetään projektiin luotua google tiliä. Palvelussa voidaan valita integrointiin useita eri palveluita, mutta tässä tapauksessa yhdistän siihen Instagram yritystilin ja OpenAI käyttäjän, joka sisältää API avaimen, jota Zapier käyttää keskustellessaan DALL E ja ChatGPT mallien kanssa. Käytetyt sovellukset nähdään kuvassa (Kuva 20).



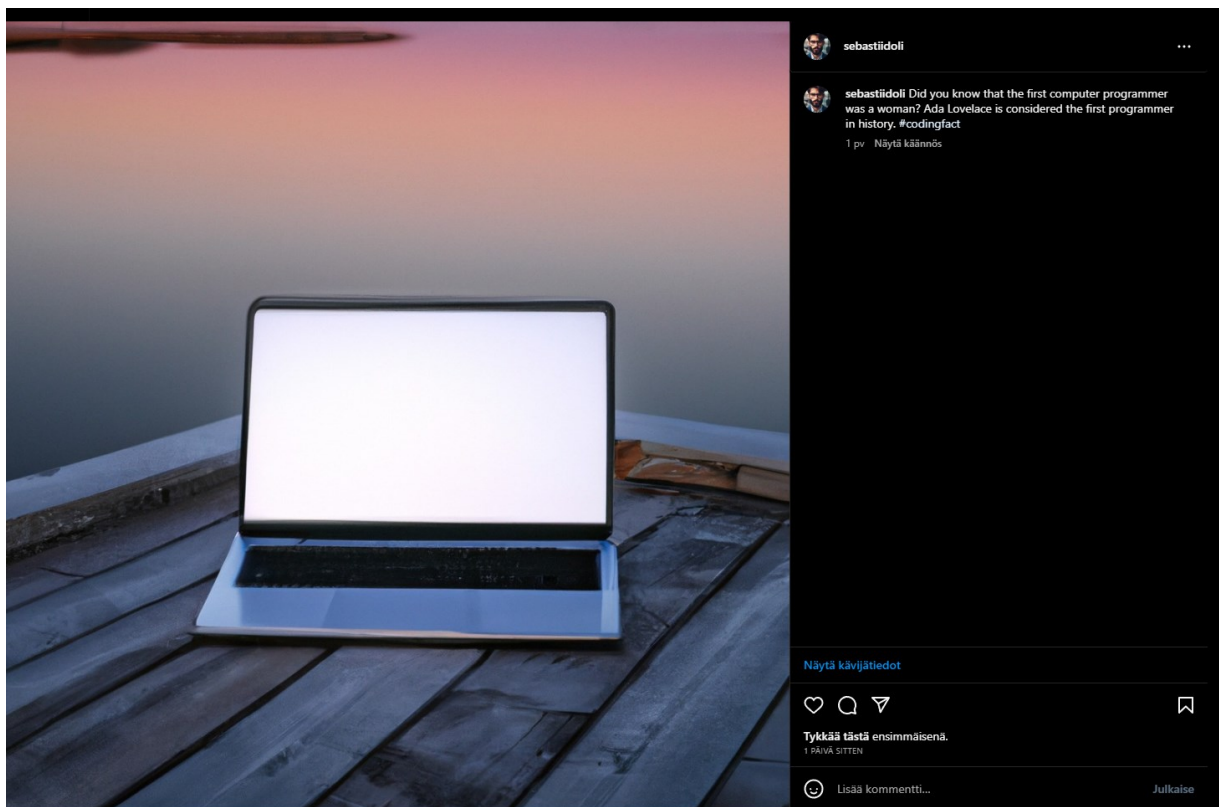
Kuva 20. Zapier käyttöliittymä, sovellukset

Ensimmäisenä osana työnkulkua käytetään Zapier palvelun schedule toimintoa, aktivoimaan työnkulku. Schedule ominaisuudella asetetaan työnkulku alkamaan päivittäin kello 19:00, lisäksi scedulen avulla voidaan määrittää, aktivoituuko toiminto viikonloppuisin, tässä tapauksessa kyllä. Seuraavana osana työnkulussa luodaan julkaistava kuva DALL-E mallin avulla. Toiminnossa tulee määrittää kehote, jolla kuva luodaan luotavien kuvien määrä ja kuvan resoluutio. Kolmantena osana työnkulkua luodaan kuvaus julkaisulle käyttäen ChatGPT:tä. Toiminnossa tulee määrittää tapahtuman tyyppi, joka määrittellään keskusteluksi. Tapahtuma tyyppin määrittämisen jälkeen kirjoitetaan kehote, johon kielimalli vastaa. Mallille lähetettävän viestin lisäksi asetetaan erityyppisiä sääntöjä, jotka auttavat mallia luomaan halutun vastauksen. Viimeisenä vaiheena työnkulussa asetetaan tehtäväksi Instagram julkaisu ja määrittellään julkaisun sisällöksi aikaisemmin luotu kuva ja teksti. Viimeisen toiminnon lisäyksen jälkeen on luotu haluttu työnkulku. Työnkulku nähdään kuvassa (Kuva 21).



Kuva 21. Zapier käyttöliittymä, työnkulku

Työnkulun toiminnallisuudesta voidaan varmistua tarkistamalla, työnkulussa kohdistettua Instagram sivua. Julkaisu nähdään kuvassa (Kuva 22).



Kuva 22. Instagram julkaisu

Kuvan julkaisu toiminnon automaatio Zapier palvelun kautta on tehty käyttäjäystävälliseksi, joten se on helppo toteuttaa.

4.6 Jatko kehitys

Työn kehitys mahdollisuudet ovat laajat ja botteja voi kehittää useaan eri suuntaan. Seuraavaksi määrittelen kolme mahdollista kehityskohdetta.

Botit toimivat työssä itsenäisesti, mutta seuraavana loogisena edistysenä boteille halutaan luoda interaktio mahdollisuuksia keskenään. Tämän voi toteuttaa mahdollisesti kahdella tavalla, joko hyödynnetään Zapier automaatio palvelussa olevia triggereitä tekemään julkaisuja, jos toisessa mediassa tehdään julkaisu. Ongelmana Zapier ratkaisussa on se, että se ei toimi X palvelun kanssa. Toisena mahdollisuutena X palveluun voidaan hankkia korkeampi kehittäjä luokka, jonka myötä saadaan kerättyä ID tietoja julkaisuista ja käyttäjistä. ID tietojen avulla voidaan uudelleen julkaista X julkaisuita, esimerkiksi Reddit alustalla.

Tässä työn kehitysvaiheessa bottien toiminnot ovat ajettu paikalliselta työasemalta. Ratkaisu on toimiva, mutta käytännöllisempi ratkaisu olisi ajaa toiminnot ulkopuoliselta palvelimelta. Palvelin voi

olla itseluotu virtuaalikone tai ulkoisesti hankittu pilvipalvelu esimerkiksi Googlen tai Amazonin kautta.

Työn isoin kehitys mahdollisuus kohdistuu sisällön luomiseen. Sisällön yksilöllistämistä, monipuoleisuutta ja toiminnallisuutta voidaan kehittää määräämällä paremmat säännöt kielimalleille annetuille kehoitteille. Kielimallille annettuihin kehoitteisiin voidaan lisätä mahdollisesti tarkemmat kieliasu säännöt, parempi konteksti tehdylle kehoitteelle ja esimerkiksi vaihteleva aihe, joka perustuu suosittuun julkaisun nimeen tai aiheeseen.

5 Johtopäätökset

Työssä toteutettiin kolme ydin osaa, jotka olivat botit, niiden toiminnallisuus ja automaatio sosiaalisen median kontekstissa. Mitä tuloksia toteutuksesta ilmeni?

Ensimmäisenä tutkimuskysymyksenä työssä halusin selvittää bottien toteutuskelpoisuuden ammatillisuuden taitojen omaavalle ihmiselle. Työssä ilmennyt tulos oli selvä kyllä. Tutkimuksessa käytetyillä alustoilla X ja Reddit on yksinkertaista luoda ja käyttää botteja hyödyntäen palveluiden jakamisa ohjeita ja julkista materiaalia. Alustat tukevat bottien käyttämistä antamalla siihen käytettävät työkalut tai tukemalla työkaluja, jotka hyödyntävät palveluiden ohjelmistorajapintaa. Automaatiopalveluita, kuten Zapier on myös helppo hyödyntää julkaisuiden automatisoinnissa, koska palvelu on tehty käyttäjäystävälliseksi.

Toisena työn tutkimuskysymyksenä oli selvittää, onko mahdollista luoda uskottava ihmiseltä vaikuttava botti, joka toimii sosiaalisen median sisällä huomaamattomasti. Tämä kysymys ei saanut suoraa vastausta, sillä työssä ei päästy ajallista syistä luomaan prototyyppiä, jota päästäisiin testaamaan, mutta tutkimukseni perusteella riittävän prototyyppi botin luonti on mahdollista. Miksi uskon tutkimuskysymyksen vastauksen olevan kyllä.

Työssä toteutetut bottien ominaisuudet antavat selvän pohjan ”uskottavan” botin luonnille, sillä työssä luotujen ominaisuuksien kehittämismahdollisuudet ovat lähes rajattomat. Yksinkertaisimmat ominaisuudet uskottavuuden toteutukselle ovat automaatio ja aikataulutuksen satunnaisuus. Aikataulutuksen satunnaisuus toteutettiin boteilla ja automaatiopalvelussa, käyttämällä yksinkertaisia sääntöjä, koodia ja funktioita, joita voidaan hyödyntää julkaisuiden tai interaktioiden satunnaisessa toteutuksessa. Aikataulutuksessa ja automaatiossa käytettäviä sääntöjä on helppo muokata ja säätää jatkossa bottien käyttötarkoituksen mukaan. Toisena uskottavuuteen vaadittavana toiminnallisuutena botille olivat erityyppiset interaktiot alustojen sisällä. Tutkimuksessa toteutettu kommentointi interaktio Reddit alustalla on yksinkertainen, mutta sitä on mahdollista kehittää ja hyödyntää ”uskottavassa” botissa, antamalla koodin sisällä tarkempia sääntöjä kommentointiperusteille ja sen luomille kommentteille. Interaktioiden hyödyntäminen X alustalla on myös mahdollista, mutta vaatii rahallista sijoittamista lisäoikeuksien hankkimiseen. Viimeisenä osana uskottavuuden toteutuksessa oli sisällön luonti. Työssä toteutetut sisällön luonti ominaisuudet hyödynsivät kielimalleja, joille annettiin yksinkertaisia kehoitteita, joista saatuja vastauksia käytettiin julkaisuiden sisältönä. Tuloksena saatiin botit tuottamaan yksinkertaista ja itseään toistavaa sisältöä, mutta ominaisuuden kehittämismahdollisuudet ovat lähes rajattomat. Hyödyntäen paremmin määriteltyjä kehoitteita ja vaihtelevia aiheita, on mahdollista luoda uskottava käyttäjän aikalinja sosiaalisen median botille.

Lähteet

Amazon Web Services. What is a bot?. Luettavissa: <https://aws.amazon.com/what-is/bot/>. Luettu: 6.10.2023

Britannica 2023. social media. Luettavissa: <https://www.britannica.com/topic/social-media>. Luettu: 8.11.2023

Brown, S. 21.4.2021. Machine learning, explained. Luettavissa: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>. Luettu: 8.11.2023

Combs, V. 1.9.2020. Think before you share: Four tips and three tools for spotting a bot on social media. Luettavissa: <https://www.techrepublic.com/article/think-before-you-share-four-tips-and-three-tools-for-spotting-a-bot-on-social-media/>. Luettu 7.10.2023

Drapkin, A. 23.10.2023. What Is ChatGPT? A Beginner's Guide With Simple Explanations. Luettavissa: <https://tech.co/news/what-is-chatgpt-beginners-guide-simple>. Luettu:10.11.2023

Hashemi-Pour, C. 8.11.2023. OpenAI. Luettavissa: <https://www.techtarget.com/searchenterpriseai/definition/OpenAI>. Luettu:10.11.2023

Hoffman, C. 27.12.2021. What Is an API, and How Do Developers Use Them? Luettavissa: <https://www.howtogeek.com/343877/what-is-an-api/>. Luettu 9.11.2023

Logreira, S. 10.4.2023. What is Zapier and What are its Benefits?. Luettavissa: <https://www.linkedin.com/pulse/what-zapier-its-benefits-steven-q-logreira->. Luettu (10.11.2023)

Mahajan, S. 2.11.2020. Beginners Guide to Tweepy. Luettavissa: <https://www.loginradius.com/blog/engineering/beginners-guide-to-tweepy/>. Luettu: 10.11.2023

NICC 2018. SOCIAL MEDIA BOTS OVERVIEW. Luettavissa: https://niccs.cisa.gov/sites/default/files/documents/pdf/ncsam_socialmediabotsoverview_508.pdf?trackDocs=ncsam_socialmediabotsoverview_508.pdf. Luettu 20.10.2023

Potter, W. 15.2.2021. A Brief History of the "Bot": From IRC to ContentBot. Luettavissa: <https://www.contentbot.ai/blog/news/a-brief-history-of-the-bot/>. Luettu: 7.10.2023.

PyPI. Pip 23.3.1. Luettavissa: <https://pypi.org/project/pip/>. Luettu:10.11.2023

PyPI. praw 7.7.1. Luettavissa: <https://pypi.org/project/praw/>. Luettu:10.11.2023

Python. What is Python? Executive Summary. Luettavissa: <https://www.python.org/doc/es-says/blurb/>. Luettu: 8.11.2023

Sheen, T. 23.3.2023. DALL-E. Luettavissa: <https://www.techopedia.com/definition/dall-e>. Luettu: 10.11.2023

Stefanowicz, B. 19.10.2023. Top Real-Life Chatbot Use Cases That Work [2023]. Luettavissa: <https://www.tidio.com/blog/chatbot-use-cases/>. Luettu: 8.11.2023