LAB University of Applied Sciences

Azure OpenAI:n tietoturvariskit organisaatiossa ja lokaalin tekoälyn mahdollisuudet

LAB-ammattikorkeakoulu
Tradenomi (AMK)
2024
Eetu Paananen

Tiivistelmä

Tekijä(t)	Julkaisun laji	Valmistumisaika
Paananen, Eetu	Opinnäytetyö, AMK	2024
	Sivumäärä	
	37	

Työn nimi

Azure OpenAI:n tietoturvariskit organisaatiossa ja lokaalin tekoälyn mahdollisuudet

Tutkinto

Tietojenkäsittely Tradenomi (AMK)

Toimeksiantajan nimi, titteli ja organisaatio

Kristian Hentula, Digital Workplace Lead, Aalto IT

Tiivistelmä

Opinnäytetyön tavoitteena oli arvioida Azure OpenAI:n tietoturvallisuutta pilvipohjaisena tekoälynä ja vertailla sitä lokaali tekoälyn tarjoamiin ratkaisuihin. Tutkimuksessa keskityttiin tunnistamaan Azure OpenAI:n tietoturvariskit ja selvittämään, tarjoaako lokaali tekoäly paremman tietosuojan luokitellun datan käsittelyssä.

Laadullisen tutkimuksen menetelmillä toteutettu opinnäytetyö sisälsi teoreettisen viitekehyksen ja asiantuntijahaastattelut. Molemmat tietolähteet olivat samassa linjassa ja tukivat opinnäytetyön tulosta. Tutkimustuloksen mukaan Azure OpenAI:ssa ei havaittu merkittäviä tietoturvauhkia. Siitä huolimatta tutkimus osoitti, että lokaalin tekoälyn käyttö voi vähentää tietoturvariskejä huomattavasti, sillä se mahdollistaa datan säilyttämisen kokonaan organisaation omassa hallinnassa toisin kuin pilvipohjaiset tekoälypalvelut.

Opinnäytetyön tulokset ovat hyödyllisiä organisaatioille, jotka ovat suunnittelemassa tekoälyn käyttöönottoa ja pohtivat samalla eri vaihtoehtojen tietoturvakysymyksiä.

Asiasanat

Azure OpenAI, lokaali tekoäly, tietoturva, pilvipalvelut, GDPR, kielimallit

Abstract

Author(s)	Type of Publication	Published
Paananen, Eetu	Thesis, UAS	2024
	Number of Pages	
	37	

Title of Publication

Azure OpenAI's cybersecurity risks in organizations and the potential of local artificial intelligence.

Name of Degree

Business Information Technology (UAS)

Name, title and organization of the client

Kristian Hentula, Digital Workplace Lead, Aalto IT

Abstract

The aim of the thesis was to evaluate the cybersecurity of Azure OpenAI as a cloud-based artificial intelligence system and compare it to the solutions provided by local artificial intelligence systems. The research focused on identifying the cybersecurity risks associated with Azure OpenAI and determining whether local artificial intelligence provides better protection for processing classified data.

The thesis was conducted using qualitative research methods, incorporating a theoretical framework and expert interviews. Both sources were consistent and supported the findings. Research results indicated no significant cybersecurity threats within Azure OpenAI. Nevertheless, the study showed that local-AI could significantly reduce cybersecurity risks because it allows data to be entirely controlled within the organization's own infrastructure, unlike cloud-based artificial intelligence solutions.

These findings are valuable for organizations that are planning to implement artificial intelligence and are evaluating the cybersecurity aspects of different technological options.

Keywords

Azure OpenAI, local artificial intelligence, data security, cloud services, GDPR, language models

Sisällys

Liitteet

1 Johdanto

Taustaa tutkimukselle

Tässä opinnäytetyössä tutkitaan yhtä digitaalisen aikakauden isointa megatrendiä: tekoälyä ja sen ongelmia. Kyseessä on maailmaa mullistava ajankohtainen innovaatio, jonka hyödyntäminen yrityksissä ja organisaatioissa tarjoaa ennennäkemättömiä mahdollisuuksia, mutta samanaikaisesti se on herättänyt huolta tietoturva- ja yksityisyysriskeistä.

Edellä mainittuja riskejä on havaittu myös Aalto-yliopiston IT-osastolla, tämän opinnäytetyön toimeksiantajalla, jossa pohditaan parhaita keinoja turvalliseen tekoälyn hyödyntämiseen. Tällä hetkellä tekoälyn maksimaalinen hyödyntäminen ei ole kuitenkaan mahdollista, sillä organisaatiot pelkäävät datan vuotavan ulkopuolisille tahoille ilman heidän suostumustaan. Tästä varoittavana esimerkkinä voidaan pitää laajaa mediahuomiota saanutta ChatGPT:tä, jonka käyttö on kasvanut nopeasti. ChatGPT kerää ja analysoi käyttäjien antamia tietoja, kuten keskusteluja ja palautetta, parantaakseen kielimalliaan ja tarjotakseen entistä relevantimpia vastauksia. Lisäksi se kerää käyttödataa, kuten käyttöaikaa ja keskustelun aiheita, henkilökohtaisen käyttökokemuksen parantamiseksi. (Boina & Achanta 2022, 2).

Tekoälyyn liittyvien yksityisyys- ja luottamusongelmien kanssa kamppaillaan myös globaalilla tasolla, mikä rajoittaa merkittävästi kyseisen teknologian hyödyntämistä. Microsoft on hiljattain lanseerannut oman pilvitekoälynsä, Azure OpenAI:n, joka on jo käytössä useissa Microsoftin ympäristössä toimivista yrityksissä. Tässä tutkimuksessa tarkastellaankin Azure OpenAI -palvelun käyttöä organisaatioiden tietoturvan näkökulmasta. Opinnäytetyön hypoteesi rakentuu ennakko-olettamalle, joka on lähtöisin tutkimuksen yhteistyökumppanin Aalto IT:n ohjeistuksesta liittyen tekoälyjen käyttöön:

> *Julkista ChatGPT:tä tulisi käyttää ainoastaan julkisen tiedon kanssa* (*Ei Aalto-yliopiston sisäistä tietoa, ei Aalto-yliopiston luottamuksellista tietoa, ei Aalto-yliopiston salassa pidettävää tietoa, ei henkilökohtaisia tietoja. Esimerkiksi ei opiskelijamateriaaleja, ei sisäistä tutkimusmateriaalia. Sisältö, jonka lisäät julkiseen ChatGPT:hen, tulee OpenAI:n käyttöön heidän datansa ja palveluidensa parantamiseksi.)* (Aalto-yliopisto 2023, 2.)

Opinnäytetyön hypoteesin mukaan pilvitekoälyjä eli Azure OpenAI:ta sekä ChatGPT:tä tulee käyttää ainoastaan julkisen datan kanssa, jotta organisaatiot välttyvät heidän omistaman tiedon vuotamiselta. Vain julkisen datan käyttöä suosittelee Aallon lisäksi lukuisat muut tahot, ja OpenAI on kamppaillut tietosuojaongelmien kanssa jo pitkään. Useat yritykset ovat tehneet valituksia OpenAI:ta vastaan ja jotkin maat, kuten Italia, ovat jopa estäneet ChatGPT:n käytön kokonaan tietojen keräyksen vuoksi (Kallum 2023). Yritykset kuten Apple, Amazon ja JP Morgan ovat myös rajoittaneet ChatGPT:n käyttöä tietosuojaongelmien vuoksi (De Rose 2023), mikä korostaa tietoturvariskien merkittävyyttä organisaatiotasolla.

Toisaalta IT-journalisti Porterin (2023) mukaan OpenAI:n ChatGPT on maailman nopeimmin kasvava digitaalinen alusta, joka keräsi 100 miljoonaa aktiivista käyttäjää vain kahdessa kuukaudessa. Vertailun vuoksi Facebook tarvitsi tähän aikaa neljä vuotta. Tämä herättää kysymyksen tekoälyn ympärillä olevien tietosuojahuolten aiheellisuudesta ja siitä, tulisiko käyttäjien suhtautua varauksella näiden palveluiden käyttöön.

Tässä opinnäytetyössä käsitellään Azure OpenAI:n tietosuojaan liittyviä haasteita, joita monet yritykset, mukaan lukien Aalto-yliopiston IT-osasto, parhaillaan kohtaavat. Aihe on merkittävä ja haastava, sillä vaikka tietosuojavaatimukset laajenevat ja huoli digitaalisten järjestelmien haavoittuvuuksista kasvaa, kattavaa tietoa Azure OpenAI:n tietoturvasta ja lokaalin tekoälyn hyödyntämisestä on saatavilla vain niukasti.

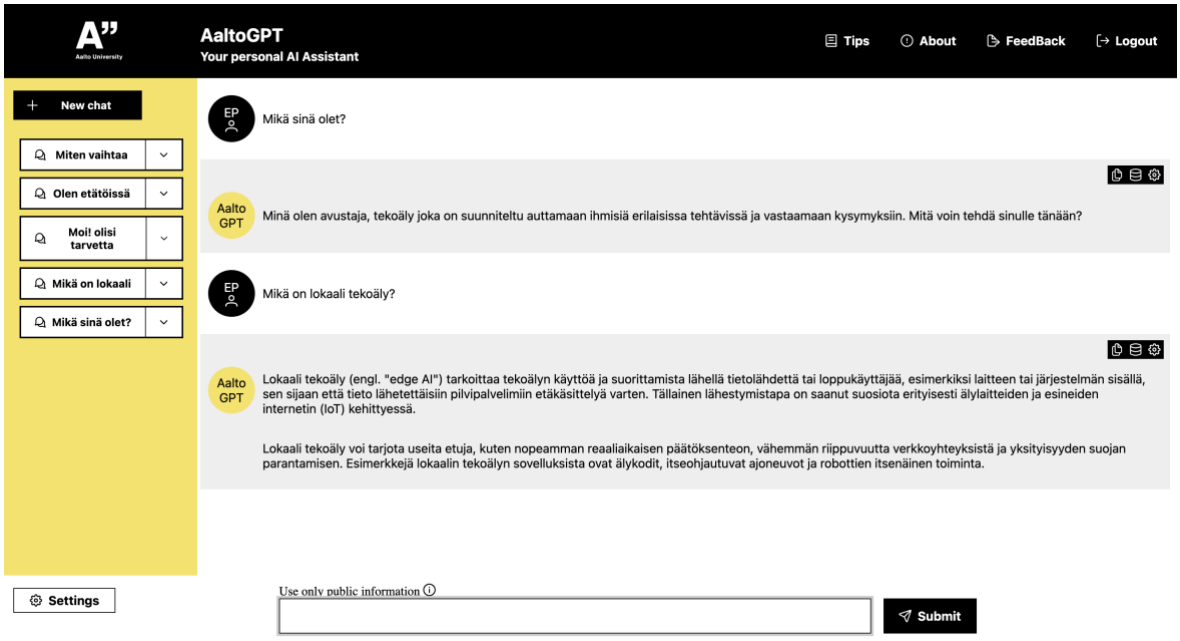
Pilvipohjaisien tekoälyjen mukaan lukien Azure OpenAI:n ja ChatGPT:n mahdollisen korvaajana on pidetty lokaalia tekoälyä, joka ei vielä ole saavuttanut valmista kaupallista muotoa. Tämän opinnäytetyön tavoitteena on arvioida, voidaanko Azure OpenAI:hin luottaa tekoälyn palveluntarjoajana sekä selvittää, millaisia ratkaisuja tietosuojaongelmiin lokaali tekoäly voisi tarjota.

Aalto-IT opinnäytetyön toimeksiantajana

Tämän opinnäytetyön toimeksiantajana toimii Aalto-yliopiston IT-osasto (Aalto-IT), jossa olen myös henkilökohtaisesti työskennellyt viimeisen vuoden ajan. Espoon Otaniemessä sijaitseva Aalto-yliopisto on yksi Suomen suurimmista oppilaitoksista, jossa opiskelijoiden ja henkilökunnan yhteenlaskettu lukumäärä on noin 17 000 henkeä. Laaja IT-osasto, joka työllistää noin 150 henkilöä, korostaa IT-osaamisen merkitystä modernissa opetuksessa ja tutkimuksessa.

Aalto-IT on kehittänyt oman tekoälyn Aalto GPT:n (Kuva 1), joka on Azure OpenAI:n päälle rakennettu, API-rajapintoja hyödyntävä tekoäly. Tämä GPT 3.5 ja GPT 4.0 kielimalleja

käyttävä chattibotti tuottaa vastauksia julkisen datan avulla, tukien Aallon turvallisuus- ja tietosuojalinjauksia.



Kuva 1. Aalto GPT:n käyttöliittymä (Aalto-yliopisto 2024)

Opinnäytetyön tavoite

Tämän opinnäytetyön keskeisenä tavoitteena on selvittää, voidaanko pilvitekoäly Azure OpenAI:hin luottaa tekoälyn palveluntarjoajana sekä selvittää, millaisia ratkaisuja tietosuojaongelmiin lokaali tekoäly tarjoaa. Tutkimus pyrkii vastaamaan keskeiseen kysymykseen: miten organisaatiot voivat hyödyntää tekoälyä turvallisesti, säilyttäen samalla datan yksityisyyden ja tietosuojan. Ensimmäisenä pyritään arvioimaan, onko Azure OpenAI tietoturvallinen tekoälyratkaisu yrityksille. Tämän ymmärtäminen vaatii siihen liittyvien käsitteiden hahmottamista.

Toisena keskeisenä tavoitteena on tutkia, voisiko lokaali tekoäly olla ratkaisu pilvitekoälyn tietoturvaongelmiin, arvioida sen käyttöä luottamuksellisen datan kanssa ja pohtia siihen liittyviä riskejä sekä mahdollisuuksia organisaatioille. Lokaali tekoäly on herättänyt kiinnostusta monissa organisaatioissa mahdollisena vaihtoehtona pilvitekoälylle.

Rajaukset

Opinnäytetyön tarkastelun kohteeksi valitaan pilvitekoälyistä vain yksi sovellus, Azure OpenAI, sillä se on yleisesti käytetty tekoälypalvelu, jota yritykset hyödyntävät

toiminnassaan. Pilvitekoälyn tai ChatGPT:n toiminta teoriassa ei ole opinnäytetyön varsinainen aihe, mutta tekoälyn tietoturvasta puhuttaessa on tärkeää ymmärtää, miten järjestelmät on kehitetty ja siksi myös kielimallin toimintaa sekä datan kiertoa järjestelmissä käsitellään. Varsinaisena tarkoituksena on selventää yrityksille mitä tietosuojariskejä pilvitekoälyn käytössä voi olla ja miten lokaali tekoäly voisi ratkaista pilvitekoälyn ongelmat. Näin ollen opinnäytetyö ei vain tarkastele tekoälyn ongelmakohtia, vaan pyrkii myös tarjoamaan mahdollisia ratkaisuja näihin haasteisiin. Tavoitteena on edistää ymmärrystä tekoälyn tietosuoja riskeistä sekä avata tietoisuutta uusista mahdollisuuksista lokaalin tekoälyn käytössä.

Tutkimusprosessi

Tämä opinnäytetyö toteutetaan käyttäen laadullista eli kvalitatiivista tutkimusmenetelmää. Laadullinen tutkimus pyrkii ymmärtämään ilmiöitä mielenkiintoisesta näkökulmasta ja tarjoamaan syvällisen kuvauksen tutkittavasta asiasta. Tämä tutkimus perustuu alustavaan hypoteesiin, jonka mukaan Azure OpenAI:ta saa hyödyntää vain julkisen datan kanssa. Tätä hypoteesia pohditaan ensin teoreettisen viitekehyksen perusteella, jonka jälkeen suoritetaan varsinainen tutkimus Aalto-IT organisaatiossa.

Tutkimuksen aineisto kerätään haastatteluita hyödyntäen, joiden tavoitteena on saada mielenkiintoisia näkökulmia ja ratkaisuja tekoälyn tietosuojaongelmiin. Haastateltavat valitaan harkittua otantaa käyttäen niin, että mukaan valikoidaan aiheesta mahdollisimman paljon tietävät henkilöt.

Tutkimuskysymykset

Opinnäytetyössä keskitytään kahteen laajaan tutkimuskysymykseen, joiden pohjalta koko tutkimusprosessi rakentuu. Tutkimuskysymykset auttavat selkeyttämään tutkimuksen päämäärää ja suuntaa tutkimuksen edetessä.

Opinnäytetyön tutkimuskysymykset:

- Voidaanko Azure OpenAI:ta käyttää luottamuksellisen datan esimerkiksi henkilötietojen kanssa?
- Miten lokaalitekoäly voi ratkaista pilvitekoälyn ongelmia?

Ensimmäinen tutkimuskysymys keskittyy tekoälyn tietosuojaongelmiin organisaatiotasolla. Tarkoituksena on syventyä konkreettisiin haasteisiin, joita yritykset maailmanlaajuisesti kohtaavat tämän teknologian hyödyntämisessä. Ongelmat saattavat liittyä regulaatioihin, standardeihin sekä yleisesti tietosuojaan liittyviin seikkoihin.

Toisena tutkimuskysymyksenä tarkastellaan sitä, millaisia ratkaisuja lokaali tekoäly voi tarjota pilvitekoälyn kohtaamiin haasteisiin. Tutkimuksen tavoitteena on ymmärtää, miten lokaali tekoäly eroaa pilvitekoälystä suorituskyvyn ja tietoturvan näkökulmasta sekä millaisia hyötyjä ja haasteita sen käyttöön liittyy organisaatioissa. On myös tärkeää selvittää, mitä teknologisia vaatimuksia lokaali tekoäly edellyttää resurssien osalta itse tietokoneelta, jos kielimalli ladataan käyttäjien tietokoneille. Lisäksi tutkimuksessa käsitellään tekoälyn toimintaperiaatteita, tietoturvastandardeja sekä tekoälyn käytön rajoituksia yrityksissä.

Lisätutkimuskysymyksiä, jotka auttavat päämäärän selventämistä:

- Miten tekoäly toimii?
- Millaisia tietoturvastandardeja ja -tekniikoita niihin on kehitetty?
- Miksi tekoälyjen käyttö on useissa yrityksissä rajoitettua?
- Mitä eroa on pilvi- ja lokaalilla tekoälyllä?
- Miten data liikkuu lokaalissa- ja pilvitekoälyssä?
- Pystytäänkö Azure OpenAI:ta käyttämään luottamuksellisen datan kanssa?

Edellä mainitut kysymykset ovat tärkeitä, sillä tekoälyn edistyksellisyys on kiistatonta, ja sen käyttö voi merkittävästi tehostaa työntekoa ympäri maailmaa. Opinnäytetyön avulla pyritään löytämään vastauksia näihin kysymyksiin, jotka ovat olennaisia tekoälyn käytön kehittämisessä ja organisaatioiden turvallisuuden varmistamisessa. Työn tavoitteena on antaa kokonaisvaltainen kuva tekoälyn käytön mahdollisuuksista ja haasteista organisaatioille, edistäen turvallisen ja vastuullisen tekoälyn käyttöä.

2 Tekoälyn tietoturva organisaatioissa: Teoreettinen viitekehys

2.1 Tekoälyn määritelmä

Tekoäly (AI, engl. artificial intelligence) viittaa koneisiin tai ohjelmiin, jotka kykenevät suunnittelemaan, tekemään päätöksiä, ja luomaan uutta. Tällaiset järjestelmät kykenevät suorittamaan niille annettuja tehtäviä ja hyödyntämään syötteeksi saatuja tietoja ongelmien ratkaisemiseksi. Yleiskielessä termi "tekoäly" viittaa teknologioihin, jotka matkivat ihmisen ajatteluprosesseja. (Euroopan parlamentti 2020.)

Tekoälyn tekninen perusta rakentuu koneoppimisen, algoritmien ja kielimallien varaan, jotka yhdessä antavat järjestelmille kyvyn tuottaa ihmiselle luonnollista vaikuttavaa tekstiä (Salo 2023, 43). Kielimallit, kuten OpenAI:n GPT-sarja ja Googlen BERT, ovat tekoälyjärjestelmien kannalta keskeisiä tekniikoita, mahdollistaen chattibottien reagoinnin ihmisen esittämiin kysymyksiin.

Tekoälyteknologiat ovat olleet kehityksen alla yli viisi vuosikymmentä, mutta erityisesti viime vuosina laskentatehon, suurten datamäärien ja uusien algoritmien kehittymisen myötä tekoäly on kokenut huomattavia läpimurtoja (Euroopan parlamentti 2020). Vuonna 2024 tekoäly on profiloinut itsensä teknologia-alan huomattavimmaksi trendiksi, mikä oli ennakoitavissa edellisen vuoden suurimpia megatrendejä seuraamalla, kun tekoäly oli IT-alan suosituin teknologia. (Bolwell 2023.)

Nykypäivänä tekoälyä pidetään olennaisena osana yhteiskunnan digitaalista muutosta, ja se on noussut yhdeksi Euroopan unionin keskeisistä painopistealueista, mikä korostaa sen merkitystä yhteiskunnallemme. Tulevaisuuden tekoälyn kehityksen odotetaan tuovan mukanaan merkittäviä muutoksia, mutta on syytä muistaa, että se on jo nykypäivänä olennainen osa arkielämäämme esimerkiksi hakukoneiden, robottien ja chattibottien muodossa. (Euroopan parlamentti 2020.)

2.2 GPT-kielimalli yleisesti

Tekoäly tarvitsee toimiakseen kielimallin eli large language modelin (LLM), joka tarkoittaa mallia, mikä pyrkii arvioimaan ja generoimaan luonnollista tekstiä. GPT (Generative Pre-trained Transformer) on yksi suosituimmista kielimalleista, jonka Yhdysvaltalaisyhtiö OpenAI on kehittänyt. GPT-kielimalli perustuu Transformer-arkkitehtuuriin, jota on koulutettu laajoilla tekstiaineistoilla. Tämä tekee GPT:stä erään alan edistyneimmistä ja laajinten käytössä olevista kielimalleista. (Latterner 2023.)

Open AI julkaisi ensimmäisen GPT mallin vuonna 2018 ja useita versioita myöhemmin uusin versio on nimeltään GPT-4, joka julkaistiin maaliskuussa 2023. ChatGPT tarjoaa GPT-3.5-mallin käyttäjille ilmaiseksi, kun taas GPT-4 on saatavilla lisämaksusta (20 $/kk). (OpenAI b.) Nämä molemmat kielimallit ovat osa Azure OpenAI:n tarjontaa Microsoftin alustalla.

2.3 GPT-kielimallin toiminta teoriassa

GPT:n taustalta löytyy kielimallin lisäksi muutakin teknologiaa. Tekoäly-asiantuntija Salo (2023) kuvailee ChatGPT:n toimintamallia kirjassaan seuraavalla tavalla:

> *Sovelluksen toiminta perustuu syväoppimiseen ja neuroverkkoihin, jotka on mallinnettu ihmisaivojen toiminnan mukaan. Se on koulutettu valtavalla määrällä ”tokeneita”, jotka ovat peräisin ihmisten kirjoittamasta datasta, kuten kirjoista, artikkeleista ja muista dokumenteista eri aiheista, tyyleistä ja genreistä. Tämä suuri datasetti mahdollistaa sen, että ChatGPT oppii tunnistamaan kuvioita ja suhteita tekstidatasta ja ennustamaan, mikä teksti tulisi seuraavaksi missäkin lauseessa*. (Salo 2023, 42.)

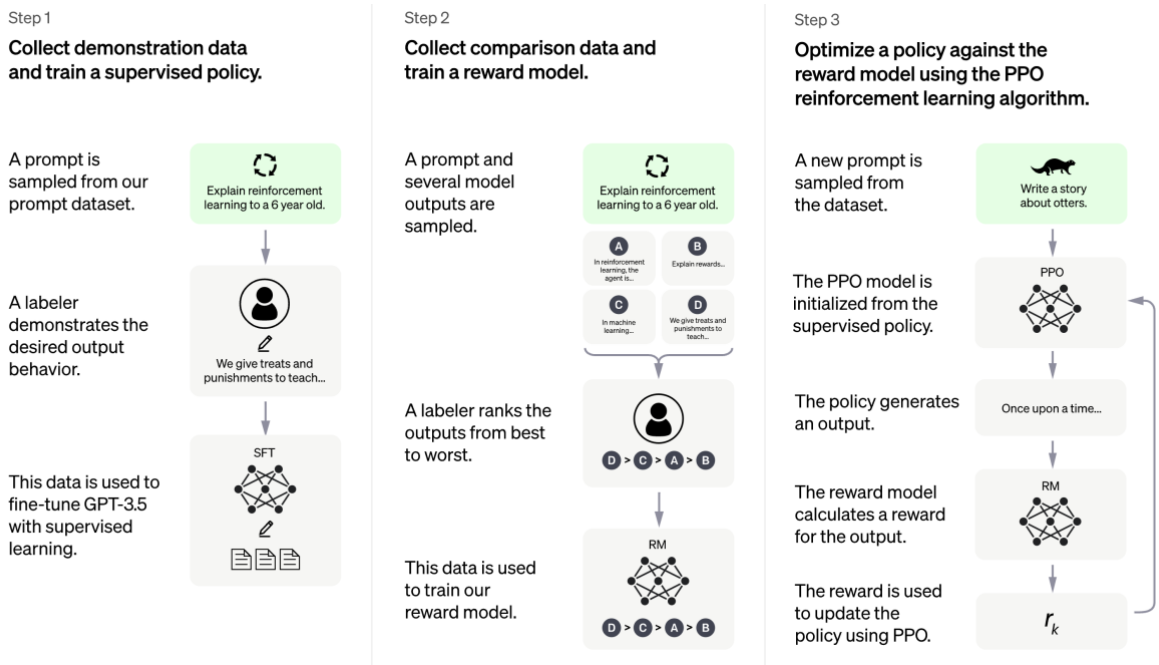
ChatGPT:n koulutuksessa on käytetty ihmispalautetta (Reinforcement Learning Human -Feedback, RLHF), joka on mahdollistanut kielimallin hienosäätämisen tarkemmaksi ja ihmismäisemmäksi. RLHF koulutuksessa AI-kouluttajat luovat esimerkkikeskusteluja chattibotin ja ihmisen välillä, ja nämä keskustelut sekoitetaan InstructGPT aineiston kanssa, jonka lopputuotoksena kielimalli osaa vastata paremmin käyttäjien kysymyksiin. (Salo 2023, 41–42.)

InstructGPT ChatGPT:n perustana

Harvard yliopiston tohtori Johri (2023) painottaa GPT kielimallin kouluttamisessa ohjeiden noudattamista ihmisten esimerkkien kautta. Hän viittaa tällä Salonkin mainitsemaan InstructGPT kielimalliin, joka sisältää 1,3 biljoonaa parametriä esimerkkikeskusteluita. OpenAI:n tieteellinen dokumentaatio (Ouyang ym. 2024, 1–2) selventää, miten GPT-kielimalli on rakennettu. Sen mukaan ChatGPT on saanut alkunsa InsturctGPT:n avulla, jonne on syötetty kysymysvastaus pareja, joita on käytetty hienosäätämään (Finetuning) kielimallia tässä valvotussa oppimisympäristössä. Nämä esimerkkimallit kouluttaja järjestää parhaimmasta huonoimpaan, jolloin tekoäly oppii tarjoamaan käyttäjilleen entistä parempia vastauksia.

Kuvassa 2 havainnollistetaan tekoälyn kehittämisen kolmevaiheista menetelmää, joka alkaa keräämällä opetusdataa ihmisten esimerkeistä, joiden perusteella koulutetaan malli valvotussa ympäristössä. Seuraavaksi prosessissa tekoälylle syötetään kysymyksiä, ja malli antaa useita eri vastauksia, jotka ihmisarvioijat luokittelevat parhaimmista huonoimpaan. Näitä tietoja käytetään kehittämään palkkiomalli, joka opastaa tekoälyä suosimaan parhaiksi arvioituja toimintoja.

Kolmannessa vaiheessa tekoälyn kehitysprosessissa keskitytään sen toimintalinjojen parantamiseen PPO eli proximal policy optimization -algoritmin avulla, joka on vahvistusoppimisen menetelmä (reinforcement learning). Tässä vaiheessa tekoälyä haastetaan uusilla prompteilla eli kysymyksillä, jotka eivät ole peräisin alkuperäisestä koulutusdatasta. Tämä pakottaa tekoälyn tuottamaan vastauksia tilanteisiin, joihin se ei ole suoraan valmistautunut. Saatuaan tekoälyn tuottaman vastauksen, edellisen vaiheen palkkiomalli arvioi vastauksen laatua. PPO-algoritmi käyttää näitä palkkioita ohjaamaan tekoälyn oppimista, jolloin se suosii niitä vastauksia, jotka saivat korkeamman palkkion. Näin kielimallin päivityksien myötä tekoälyn pitäisi oppia tuottamaan yhä laadukkaampia ja kontekstiin sopivampia vastauksia.



Kuva 2. Havainnollistava kuva GPT-kielimallin kouluttamisesta (OpenAI c.)

Tekoälyn vaatimuksina koneoppimista, neuroverkkoja ja algoritmeja

Tekoälysovellukset hyödyntävät koneoppimista keskeisenä osana toimintaansa, ja sen ymmärtäminen on välttämätöntä, jos halutaan saada kattava käsitys siitä, miten esimerkiksi Azure OpenAI toimii. Matematiikan tohtori Heikkilä (2023) käyttää koneoppimisen ymmärtämiseen vertauskuvana evoluutiota, jossa eliölajit sopeutuvat ympäristön muutoksiin. Hänen mukaansa koneoppimisessa on samankaltainen rakenne kuin evoluutiossa. Sovelluksen kohdatessa epäonnistumisia, ohjeita parannetaan ja yritetään uudelleen. Suuri määrä näitä toistoja kehittää algoritmeja ja parantaa sovelluksen toimintaa.

Heikkilä (2023) määrittelee koneoppimisen yläteorian käsitteenä, jota käytännössä toteuttavat neuroverkot. Neuroverkot hyödyntävät matemaattista tilastotiedettä valitsemalla aina optimaalisimman parametrin tai hylkäämällä vähemmän sopivan vaihtoehdon. Neuroverkot koostuvat yksinkertaisista matemaattisista lauseista, parametreista ja transformer-arkkitehtuurista. Parametrit, eli painokertoimet, säilövät opitun tiedon neuroverkoissa. Yleisesti ottaen, esimerkiksi GPT kielimallien vertailuissa, voidaan olettaa, että mitä enemmän parametrejä, sitä parempi kielimalli on kyseessä. Kielimallin transformer-arkkitehtuuri on puolestaan matemaattinen rakenne, joka määrittää, millaisia matemaattisia laskutoimituksia kielimallissa suoritetaan.

Heikkilän (2023) mukaan yritysten tekoälyprojekteissa kohtaama haaste on algoritmien optimointi, jossa oikeiden ongelmien konvertoiminen numeeriseen muotoon voi olla monimutkaista. Kurimo (2023) kertoo, että esimerkiksi tekoälyrobottien soveltaminen kotisiivoustehtävissä on erityisen haastavaa, sillä useat ei-numeeriset muuttujat vaikuttavat puhtaan kodin saavuttamiseen. Shakkirobotti sen sijaan toimii tehokkaasti, sillä shakkilauta ja pelinappulat voidaan määritellä numeerisesti, mikä mahdollistaa algoritmin paineistamisen niin, että robotti voittaa usein ihmispelaajan. Tämä esimerkki havainnollistaa, miten tekoälyhankkeita kehitetään ja kuinka AI-chattibotit voivat tuoda ratkaisuja yritysten haasteisiin, kunhan projektit suunnitellaan huolellisesti.

2.4 Pilvipalvelun määritelmä

Seuraavaksi tarkennetaan termejä pilvitekoälyihin liittyen, koska AI-alan haasteet kasaantuvat usein pilvitekoälyille, kuten esimerkiksi ChatGPT:lle. Pilvitekoälyn ymmärtäminen edellyttää ensin perustietämystä pilvipalveluiden toiminnasta. Pilvipalvelu mahdollistaa datan tallentamisen ja palveluiden hyödyntämisen verkkoyhteyden kautta, ilman fyysisten tallennusmedioiden, kuten kovalevyjen tarvetta. Pilvipalvelut haastavat perinteisen käsityksen siitä, että ihmiset säilövät dataa omiin fyysisiin tallennuspaikkoihinsa, joka tuo useita lisäominaisuuksia tiedon tallentajille. Pilvipalveluun tallentaminen tarjoaa

käyttäjille joustavuutta, skaalautuvuutta ja mahdollisuuden käyttää palveluita eri laitteilla mistä tahansa, kunhan verkkoyhteys on saatavilla. Lisäksi pilvipalvelut voivat tarjota käyttäjille tietoturvaa ja varmuuskopioinnin mahdollisuuksia tietojen säilyttämisessä. (Stagnitto 2024.)

Pilvipalvelut luokitellaan yleensä kolmeen osaan: SaaS (Software as a Service), PaaS (Platform as a Service), sekä IaaS (Infrastucture as a Service). Nämä luokittelut määrittelevät, kuinka suuri osa palvelusta on palveluntarjoajan vastuulla suhteessa käyttäjään. Esimerkiksi ChatGPT on SaaS-sovellus, joka toimii suoraan selaimessa ilman, että käyttäjän tarvitsee asentaa tai ylläpitää palvelimia. Sen sijaan Azure OpenAI on PaaS-järjestelmä, tarjoten kehitysalustan käyttäjien omien tekoälyratkaisujen rakentamiseen (Lanfear 2022.)

Datan siirtyminen pilvipalveluun (Data in Transit)

Datan siirtyessä käyttäjältä pilvipalveluun, esimerkiksi chattibotille, tietoturva on avainasemassa. Pilvipalvelut käyttävät etäpalvelimia datan, kuten tiedostojen tai kuvien, tallentamiseen. Tieto siirretään tavallisesti internetin kautta, käyttäen HTTPS-protokollaa, mikä on salattu versio perinteisestä HTTP-protokollasta. Tämä salaus lisää sekä tietoturvaa että yksityisyydensuojaa. Pilvipalveluiden tarjoajat soveltavat monia tekniikoita ja protokollia, kuten SSL/TLS-salausta, mikä mahdollistaa luotettavan todentamisen sekä suojaa viestintää yksityisyyden loukkauksilta ja datan väärentämiseltä. Näin data liikkuu turvattuna asiakkaalta pilvipalveluun. (Google).

Kun data on saapunut pilveen, se tallennetaan yleensä virtuaalikoneisiin, jotka ovat osa fyysisiä palvelimia. Pilvipalveluiden tarjoajat sijoittavat dataa lukuisiin virtuaalikoneisiin eri puolilla maailmaa. Näin varmistetaan datan saatavuus jatkuvasti ja vastataan kasvaviin tallennustilatarpeisiin lisäämällä virtuaalikoneiden määrää tarpeen mukaan (Google).

Datan suojaus pilvipalveluissa (Data at Rest)

Tietoturvan näkökulmasta on olennaista ymmärtää, missä muodossa data säilytetään pilvipalveluissa. Nykyaikaiset pilvipalvelut tarjoavat useita suojausstandardeja ja tekniikoita vastauksena tietoturvariskeihin. Myös GDPR-asetus velvoittaa kaikkia EU:n säännösten alaisia yrityksiä, pilvipalveluntarjoajia mukaan lukien, takaamaan datan laajan suojauksen.

Palveluntarjoajista Microsoft kertoo verkkosivuillaan, että heidän Microsoft 365 for Business -palvelunsa on GDPR-yhteensopiva tallennusratkaisu, mutta opinnäytetyön kannalta on syytä huomioida, että Azure ei kuulu Microsoft 365 for Business -palvelun alaisuuteen (Davis ym. 2023). Pilviasiantuntija Stagnitto (2024) huomauttaa, että on epätavallista löytää palveluntarjoaja, joka ei tarjoaisi vankkoja turvaprotokollia ja suojattuja datakeskuksia.

Pilvipalvelut tyypillisesti antavat käyttäjille mahdollisuuden personoida turva-asetuksiaan, esimerkiksi ottamalla käyttöön kaksivaiheisen tunnistautumisen. Jotkut yritykset tarjoavat lisäsuojan muodossa nollatiedon salausta, mikä takaa, että ainoastaan tiedon omistaja voi purkaa tiedon salauksen.

Vaikka useat pilvipalvelut ovatkin huolehtineet tietoturva-asioista asianmukaisesti, on tärkeää muistaa, että ne ovat edelleen alttiita esimerkiksi kyberhyökkäyksille, mikä nostaa asiakastiedon vuotamisen riskiä. Yritysten datan yksityisyyden suoja on keskeinen huolenaihe, ja usein yritykset haluavat rajoittaa pilvipalveluntarjoajan pääsyn heidän tietoihinsa. On raportoitu, että esimerkiksi Google kerää enemmän käyttäjätietoja verrattuna Syncin kaltaisiin pilvipalveluihin, jotka eivät kerää eivätkä analysoi asiakastietoja. Tämän vuoksi on kriittistä, että yritykset valitsevat pilvipalveluntarjoajansa huolella, ottaen huomioon eri tarjoajien käytännöt yksityisyyden suojauksessa (Stagnitto 2024).

2.5 Pilvitekoälyn määritelmä

Pilvipalveluiden määrittämisen jälkeen on mahdollisuus ymmärtää mitä pilvitekoäly tarkoittaa. IT-asiantuntija Brennerin (2023) mukaan pilvitekoäly viittaa tekoälysovelluksiin esimerkiksi chattibotteihin, jotka toimivat selaimessa. Kun henkilö lähettää kysymyksen esimerkiksi ChatGPT:lle tai Azure OpenAI:lle, tieto siirtyy pilvipalveluun, jossa kielimalli käsittelee sen ja lähettää vastauksen. Tämä mahdollistaa sen, että palvelu ovat saatavilla ja käytettävissä verkossa eri laitteilla ilman, että kielimallia tarvitsee asentaa paikalliseen muistiin. Pilviteknologia tekoälyssä mahdollistaa skaalautuvuuden, koska pilviympäristössä laskentaresurssit voidaan helposti sovittaa vastaamaan kysynnän vaihteluita. Tämä tarkoittaa, että tekoälysovelluksia voidaan suorittaa tehokkaasti eri mittakaavoissa.

Pilvitekoäly tarjoaa sovelluskehittäjille pääsyn valmiisiin infrastruktuureihin, välttäen tarpeen rakentaa ja ylläpitää omia palvelinkeskusjärjestelmiä. API-rajapinnat mahdollistavat tekoälysovellusten, kuten ChatGPT:n, integroimisen kehittäjien omiin järjestelmiin ilman suuria infrastruktuuri-investointeja. Pilvitekoälyn taloudelliset edut tulevat selvästi esiin, erityisesti kun otetaan huomioon kielimallien tarvitsema merkittävä muistimäärä ja se, että oma datakeskus voi olla erittäin kallis pienille yrityksille. Pilvipohjaiset ratkaisut tarjoavat kustannustehokkaan vaihtoehdon, hyödyntäen pilviteknologian tarjoamaa joustavuutta ja resurssien hallintaa. (Brenner 2023). Pilvitekoälyjen kustannusrakenne voi jakautua joko kiinteään kokonaiskustannukseen tai per request -perusteiseen hinnoitteluun, jolloin hinta määräytyy palvelun käyttövolyymin perusteella.

2.6 Lokaalin tekoälyn määritelmä

Lokaali tekoäly, joka tunnetaan myös termeillä Edge AI tai On Device AI, on aihealue, josta suomenkielistä tietoa on saatavilla niukasti. Lokaali tekoäly (Local AI) viittaa järjestelmiin tai sovelluksiin, jotka toimivat itsenäisesti laitteissa. Erotuksena pilvipohjaisiin palveluihin lokaali tekoäly toimii paikallisesti laitteessa, mahdollistaen toiminnan ilman jatkuvaa verkkoyhteyttä. On tärkeää huomioida, että lokaalissa tekoälyssä itse kielimalli sijaitsee lokaalisti käyttäjän työasemassa esimerkiksi omalla kovalevyllä, mikä mahdollistaa yksilöllisen palvelun. Lokaali kielimalli voi sijaita myös organisaation palvelimella, jolloin se palvelee useita henkilöitä samanaikaisesti. Tällainen lokaali prosessointi voi tarjota merkittäviä etuja yksityisyyden, turvallisuuden ja vastausnopeuksien kannalta, koska tiedonkäsittely tapahtuu suoraan käyttäjän laitteella tai vaihtoehtoisesti yrityksen palvelimella. (Singh & Gill 2023, 72.)

Lawlor (2024) nostaa esille, että lokaali tekoäly kykenee toimimaan täysin autonomisesti, hyödyntäen laitteen sisäistä laskentakykyä ilman ulkoisia yhteyksiä. Tämä arkkitehtuuri ei ainoastaan alenna viiveitä, vaan tarjoaa myös henkilökohtaista apua käyttäjälle, samalla kun se minimoi tietoturvariskit, jotka liittyvät pilviyhteyksien kautta tapahtuvaan tiedonkäsittelyyn. Kun kielimalli asennetaan ja säilytetään henkilön omalla laitteella, kysymykset ja vastaukset eivät vuoda laitteen ulkopuolelle.

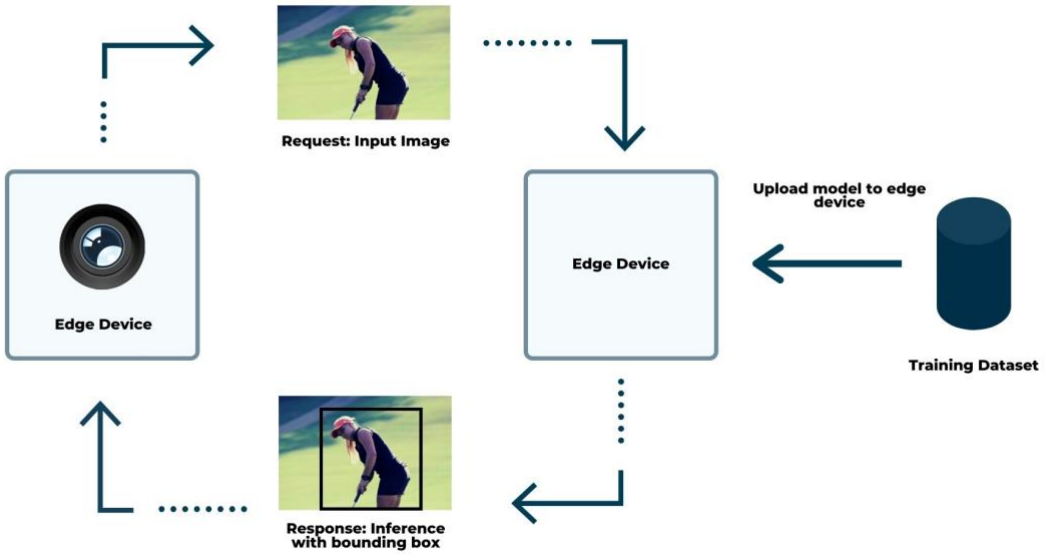
Lokaalin tekoälyn kielimallin ominaisuudet

Koska tekoäly vaatii toimiakseen kielimallin, on lokaalin kielimallin tarkastelu turvallisuus näkökulmasta välttämätöntä, kun halutaan selventää pilvitekoälyn ja lokaalin tekoälyn tietoturvaeroja. Advianin mukaan generatiivinen lokaali tekoäly voidaan asentaa tietokoneelle ilman merkittäviä tietoturvariskejä. Silloin tekoäly osaa vastata yleisiin kysymyksiin. Jos puolestaan halutaan luoda lokaali AI, joka osaa vastata yritystä koskeviin kysymyksiin, täytyy silloin kielimalli harjoittaa omaa dataa käyttäen. Tämä personoidun kielimallin luominen tuottaa osittain samoja tietoturva ongelmia kuin pilvitekoälyt, joten se ei suoraa ehkäise kaikkia pilvitekoälyn ongelmia. Sen avulla pystytään kuitenkin pelaamaan kolmannen osapuolen palveluntarjoajat ulos mahdollisista tietoturva riskeistä.

Lokaalin tekoälyn käyttöönottoon liittyvät tietosuojahaasteet ovat huomattavat, kun huomioidaan lainsäädännölliset velvoitteet, kuten GDPR, jotka asettavat tiukkoja vaatimuksia tiedon käsittelylle ja suojaamiselle. Omien keskitettyjen tietokantojen luominen lokaalissa kielimallin koulutusvaiheessa luo riskin kyberhyökkäyksille, mikä osoittaa, että

pelkkä lokaalin tekoälyn käyttö ei automaattisesti takaa parempaa tietoturvaa. Lokaali tekoäly vaatii organisaatioilta huolellista suunnittelua ja turvatoimien toteutusta. (Advian.)

Yhteenvetona voidaan todeta, että lokaali tekoäly tarjoaa edistyksellisiä ratkaisuja generatiivisen tekoälyn luomiseen, mutta personoidun kielimallin toteutuksessa on otettava huomioon tietoturva. Yhtenä ratkaisuna personoidun kielimallin kouluttamiseen käytetään yhdistettyä oppimista (Federated learning), jossa jokainen laite kouluttaa geneeristä kielimallia omilla tietokannoillaan paikallisesti ilman, että dataa lähetetään keskitettyyn palvelimeen. Kun aikaa kuluu ja yrityksellä on uutta tietoa, joka koetaan tarpeelliseksi työntekijöiden kielimallille, suoritetaan keskitetty ohjelmistopäivitys, jonka mukana uudet tiedot lähetetään lokaaleille malleille. Näin personoitu lokaali kielimalli kehittyy ilman että henkilöiden dataa lähetetään ulospäin omilta laitteilta, joka vähentää tietoturva riskejä merkittävästi. Alla olevassa esimerkkikuvassa 3 havainnollistetaan, miten sovellus pystyy tuottamaan pyydetyn rajauksen laitteen kamera sovellukseen, hyödyntäen lokaalia kielimallia. (Advian.)



Kuva 3. Lokaalin tekoälyn toimintamalli (Advian)

2.7 Lokaalin- ja pilvitekoälyn väliset erot

Keskeisin ero lokaalin ja pilvitekoälyn välillä liittyy datan sijaintiin. Kun käyttäjä avaa selaimessaan ChatGPT:n ja lähettää sille kysymyksen, tiedot siirtyvät pilvipalveluun käsittelyä varten. Lokaalia mallia käyttävä henkilö voi olla verkosta erillään, kuten puhelimen

ollessa lentotilassa, ja silti hyödyntää chattibottia, jolloin tiedot eivät kulje pilvipalveluun. Pilvitekoäly suorittaa päättelynsä etäpalvelimilla, kun taas lokaali tekoäly toimii suoraan laitteen omassa ympäristössä, mikä tarjoaa tietosuojan parannuksia ja toisaalta pienemmän kielimallin.

Qualcommin teknisen johtajan Lawlorin (2023) mukaan lokaali tekoäly suojaa käyttäjän yksityisyyttä tehokkaammin kuin pilvitekoäly, sillä vaikka pilvipalveluissa olevat tiedot ovat periaatteessa käyttäjän hallinnassa lainsäädännön ansiosta, lokaalissa tekoälyssä käyttäjä hallitsee dataa käytännössä. Lokaalin tekoälyn etuna on datan pysyminen laitteessa, jolloin minimoidaan riski tietomurrosta datan siirron aikana eikä ylimääräisiä datakopioita löydy pilvipalveluista, joihin saattaa myös kohdistua kyberhyökkäyksiä.

Yhteenvetona molemmilla toteutuksilla on omat ominaisuudet ja hyvät sekä huonot puolet. Lokaali tekoäly tarjoaa itsessään paremman tietosuojan verrattuna pilvitekoälyyn, kun data ei liiku laitteen ulkopuolelle. Toisaalta Azure OpenAI:n kaltaiset pilvitoteutukset tarjoavat skaalautuvuutta ja monipuolisempia vastauksia, kun itse kielimalli on laajempi ja infrastruktuuri on rajaton. Yksi merkittävä asiaa lokaaliin tekoälyyn liittyen on laitteen suorituskykyvaatimukset. Tehokkaiden kielimallien ajamiseen tarvitaan nykyaikaiset tietokoneet, joissa on vähintään 4 ydintä prosessoria kohden ja yli 16 GB RAM-muistia. Vanhempien ja vähemmän tehokkaiden laitteiden kanssa lokaalin tekoälyn suorituskyky voi olla alhainen tai käyttö kokonaan mahdotonta (Puget Systems).

2.8 Tietosuoja, regulaatiot sekä standardit

Regulaatiot ja standardit ovat välttämättömiä yhteiskunnan turvallisuuden, vakauden ja luottamuksen kannalta, sillä ne määrittävät alakohtaisia yleisesti hyväksyttyjä käytäntöjä ja kriteerejä, joita alalla toimivat yritykset noudattavat. Säädökset, joita kansainväliset tai kansalliset viranomaiset asettavat, muodostavat perustan koko alan toiminnalle ja takaavat reilun kilpailun.

Standardit antavat yksityiskohtaisia ohjeita näiden regulaatioiden käytännön soveltamiseen, mahdollistaen yritysten toiminnan ja tuotteiden laadun sekä turvallisuuden varmentamisen. Yritykset ottavat käyttöön näitä standardeja osoittaakseen, että heidän toimintatapansa vastaavat regulaatioita, joka vahvistaa asiakasluottamusta ja edistää yritysten toiminnan läpinäkyvyyttä ja vastuullisuutta. Turvallisuuden sekä luottamuksen lisääminen on erityisen tärkeää digitaalisessa ympäristössä, jossa tietoturva ja yksilöiden tietoisuus oikeuksistaan ovat jatkuvasti esillä. Alla on käsitelty opinnäytetyön kannalta merkittävimpiä regulaatioita ja standardeja.

Yleinen tietosuoja-asetus (GDPR)

IT-alan yksi merkittävimmistä regulaatioista on GDPR. Euroopan Unionin tietosuoja-asetus GDPR (General Data Protection Regulation) astui voimaan toukokuussa 2016 EU-alueella, jonka jälkeen yritysten on huolehdittava turvallisesta tietojenkäsittelystä ihmisten identiteetti turvattuna. Tietosuoja-asetus asettaa yrityksille ja organisaatioille tiukat vaatimukset siitä, miten henkilötietoja kerätään, säilytetään ja hallinnoidaan. (Euroopan parlamentin ja neuvoston asetus (EU) 2016/679, §4)

GDPR-asetus koskee sekä EU:n alueella toimivia yrityksiä sekä sen ulkopuolisia yrityksiä, jotka käsittelevät EU:n kansalaisten henkilötietoja. Asetus laajentaa yksilöiden oikeuksia, antaen heille mahdollisuuden tarkastella, muokata ja poistaa tietojaan organisaatioiden järjestelmistä. Lisäksi GDPR vaatii organisaatioilta proaktiivisia tietoturvatoimia ja velvoittaa ne ilmoittamaan tietoturvaloukkauksista niin viranomaisille kuin asianosaisillekin. (Euroopan parlamentti 2022.) Suomessa GDPR:n vaikutusta täydentää kansallinen tietosuojalainsäädäntö (Tietosuojalaki (1050/2018)), joka täsmentää yleisen tietosuoja-asetuksen määräyksiä Suomessa.

Tietosuojaa koskeva vaikutustenarviointi (DPIA)

DPIA, eli Data Protection Impact Assessment, on GDPR:n edellyttämä prosessi, jota sovelletaan, kun henkilötietojen käsittelyyn liittyy merkittäviä riskejä yksilön perusoikeuksille ja -vapauksille. Tämä menettely on erityisen tärkeä uutta teknologiaa käyttöönotettaessa tai suuria henkilötietomääriä käsiteltäessä, joista kumpikin voi lisätä tietoturvaloukkausten riskiä. DPIA:n avulla organisaatiot voivat demonstroida ja dokumentoida vastuullista henkilötietojen käsittelyä GDPR:n mukaisesti. (Wolford 2024.)

Suomen tietosuojaviranomainen korostaa, että tietosuojauhkien tunnistaminen ja arvioiminen on välttämätöntä. Tämä arviointi sisältää esimerkiksi riskien pisteytyksen, mikä on keskeinen osa organisaation riskienhallintaa ja uusien sovellusten tai palveluiden suunnittelun tueksi tehtävää päätöksentekoa. DPIA edesauttaa riskien ennakointia ja auttaa kehittämään strategioita niiden minimoimiseksi, suojellen näin sekä organisaatiota että sen asiakkaiden yksityisyyttä. (Tietosuojavaltuutetun toimisto).

ISO/IEC 27000-Tietoturvallisuuden standardisarja

ISO/IEC 27000-sarjan tietoturvastandardit ovat IT-alan keskeisiä työkaluja yritysten tietoturvan ja tietohallinnan parantamiseksi. Suomen Standardisoimisliitto SFS:n mukaan nämä standardit ovat elintärkeitä, ja niiden kehittämisestä vastaa ISO eli Kansainvälinen standardisoimisjärjestö sekä IEC eli Kansainvälinen sähköteknillinen komissio. Sarjaan

kuuluu useita standardeja, jotka määrittävät tietoturvan perusperiaatteet, riskien arvioinnin ja hallinnan sekä tietoturvan ylläpidon menetelmiä.

Microsoftin asiantuntija Vidich (2023) kertoo, että Microsoftin pilvipalvelut kuten Azure ja Dynamics 365 auditoidaan säännöllisesti riippumattomien kolmannen osapuolen organisaatioiden kanssa, jotta heidän palvelunsa noudattavat ISO/IEC 27001 -sertifikaattia. Microsoftin Azure Policy dokumentaatio tarjoaa lisäksi asiakkaille työkaluja tietoturvan ja tietosuojan hallintaan, perustuen samaisiin standardivaatimuksiin. Azure Policy:n avulla organisaatiot voivat varmistaa, että heidän Azure-pilvipalvelussa säilytettävät tiedot ovat suojattuja ja täyttävät asetetut tietoturva- ja tietosuojavaatimukset.

Tekoäly regulaatio EU AI ACT

Euroopan parlamentin raportin mukaan EU:n digitaalistrategian keskiössä on tekoälyn sääntelyn kehittäminen ja kehittyneen teknologian turvallisen sekä kestävän käytön edistäminen. Huhtikuussa 2021, Euroopan komissio otti historiallisen askeleen esittelemällä ensimmäisen tekoälyä koskevan sääntelykehyksen, joka asettaa globaalilla tasolla vaatimuksia tekoälyjärjestelmille. Tämän regulaation päämääränä on taata tekoälyjärjestelmien turvallisuus ja läpinäkyvyys, varmistaa niiden jäljitettävyys ja syrjimättömyys, sekä edistää ympäristön kannalta kestäviä ratkaisuja (Benifei & Tudorache 2023, 3.)

Euroopan parlamentin (Benifei & Tudorache 2023, 3.) mukaan tekoälyä koskevat uudet säännöt perustuvat riskiperusteiseen lähestymistapaan. Tämä tarkoittaa, että tekoälyjärjestelmät on luokiteltava Euroopassa niiden aiheuttaman riskin mukaan ja tietyt korkean riskin sovellukset, kuten ihmisiä manipuloivat tai reaaliaikaisen kasvontunnistuksen AI-sovellukset, kielletään kokonaan. Tämä regulaation tavoitteena on suojella ihmisten turvallisuutta ja yksityisyyttä. AI ACT edellyttää myös tekoälymallien toiminnan läpinäkyvyyttä, joka on välttämätöntä käyttäjien luottamuksen sekä teknologian kestävän kehityksen varmistamiseksi. Tietosuojan ja yksityisyyden näkökulmasta nämä uudistukset ovat tärkeitä, koska ne ehkäisevät henkilöiden ja yritysten datan vuotamista, joka on olennainen osa tämän opinnäytetyön aihetta.

Euroopan parlamentti (2023) kertoo, että se saavutti alustavan yhteisymmärryksen neuvoston kanssa tekoälysäädöksestä 9. joulukuuta 2023. Sekä parlamentin että neuvoston on vielä hyväksyttävä regulaatio, jotta siitä tulee osa virallista EU:n lainsäädäntöä. Ennen lopullista päätöstä on tärkeää, että kaikki parlamentin jäsenet voivat ilmaista näkemyksensä asiasta. Sopimuksesta äänestävät Euroopan parlamentin sisämarkkina- ja kansalaisvapausvaliokunnat, jotka ovat keskeisiä tahoja tässä päätöksentekoprosessissa.

2.9 Microsoftin ja OpenAI projekti Azure OpenAI

Vuonna 2015 perustetun yhdysvaltalaisen tekoälyyn keskittyvän yrityksen OpenAI:n tarkoituksena on omien verkkosivujen mukaan tuottaa geneerisen tekoälyn hyötyjä koko yhteiskunnalle (OpenAI a.). Yhtiö on kehittänyt laajalti tunnetun ChatGPT-tekoälyn, jota UBS:n tutkimuksen mukaan on kuvailtu maailmanhistorian nopeimmin kasvavaksi sovellukseksi (Hu 2023). Microsoft on osoittanut merkittävää kiinnostusta OpenAI:ta kohtaan ja on tullut sen suurimmaksi sijoittajaksi, omistaen 49 prosenttia yhtiön voittoa tuottavasta osuudesta, mikä on tuonut heidät OpenAI:n hallitukseen ja syventänyt yritysten välistä yhteistyötä (Tienari 2023).

Tämän yhteistyön tuloksena on syntynyt Azure OpenAI palvelu, jonka Microsoft kuvaa ratkaisuksi yrityksille, jotka haluavat integroida OpenAI:n kielimallin omiin järjestelmiinsä turvallisesti käyttäen REST API -rajapintaa. Microsoftin omistama Azure OpenAI hyödyntää Azure-pilvipalvelua, joka sisältää OpenAI:n kehittämiä kielimalleja turvallisen pilviympäristön sisällä. Turvallisuus, alueellinen rajoitus ja vastuullinen AI-sisällön suodatus ovat palvelun keskeisiä ominaisuuksia. (Browne ym. 2024).

Erityisen huomionarvoista on, että Microsoftin esitelmä verkkosivuillaan 'Vastuullisesta tekoälystä' (Responsible AI) vastaa suurelta osin aiemmin mainittua EU:n AI Act -regulaation vaatimuksia, joka edellyttää tekoälyn olevan vastuullinen eikä tuottavan harhaanjohtavaa tai vahingollista sisältöä. Microsoft on myös ilmoittanut, että johtuen suunnitelluista merkittävistä päivityksistä, heillä ei tällä hetkellä (14. helmikuuta 2023) ole resursseja tarjota palvelua uusille asiakkaille. Sen sijaan yhtiö pyrkii jatkamaan yhteistyötä vanhojen asiakkaiden kanssa ja valmistelee suurta tietoturvapäivitystä Azure OpenAI-palveluun. (Browne ym. 2024).

2.10 ChatGPT:n tietosuojahaasteet Azure OpenAI:lle

Toimittaja Drapkinin (2023) mukaan ChatGPT:n tietosuojaongelmat tarjoavat oppeja Microsoftin Azure OpenAI:lle. OpenAI:n kehittämän ChatGPT:n on raportoitu käyttävän ja mahdollisesti levittävän käyttäjien syöttämää dataa kielimallien parantamiseksi ilman käyttäjien nimenomaista suostumusta, mikä herättää huolta erityisesti yrityksissä. Organisaatioiden tulee olla varovaisia arkaluontoisen tiedon syöttämisessä ChatGPT:lle, ja useat yritykset ovat tämän vuoksi rajoittaneet sen käyttöä. On myös mainittu, että ChatGPT:n koulutuksessa käytetään julkisesti saatavilla olevia aineistoja, jotka voivat sisältää tekijänoikeuksien alaista materiaalia (Lomans 2023).

GDPR-standardit huomioiden ChatGPT:ssä on ominaisuus, jonka kautta käyttäjät voivat pyytää kaiken henkilökohtaisen datansa poistamista, ja tällöin tiedot tulisi hävittää OpenAI:n tietokannoista 30 päivän sisällä. Tästä huolimatta monilla yrityksillä on edelleen epävarmuutta heidän datansa potentiaalisesta hyödyntämisestä, mikä on johtanut ChatGPT:n kieltämiseen. Microsoft on vastannut näihin huoliin Azure OpenAI:n kautta, tarjoten sitä turvallisena vaihtoehtona tekoälysovelluksiin, missä tietoturvaa ja yksityisyyden suojaa korostetaan.

2.11 Azuren tietosuojaratkaisut

Microsoftin Azure OpenAI tarjoaa ympäristönsä kautta asiakkailleen tietosuojan takaavan palvelun. Toisin kuin monet muut pilvipalvelut, Azure ei jaa asiakkaan tietoja kolmansille osapuolille, kuten OpenAI:lle, eikä hyödynnä asiakastietoja kielimallien parantamiseen. Azure OpenAI on Microsoftin itsenäinen alusta, joka toimii eristyksissä muilta toimijoilta. Microsoftin merkittävä sijoitus OpenAI:hin on mahdollistanut erityisen turvallisen tekoäly-ympäristön luomisen Azureen, jossa voidaan käyttää OpenAI:n kielimalleja ilman, että se vaarantaa datan turvallisuutta. (Farley ym. 2023.)

Lisäksi Azure OpenAI sisältää väärinkäytön seurantasuodattimen (abuse monitoring), jonka avulla Microsoftin valtuutetut työntekijät voivat puuttua epäilyttävään toimintaan. Yritykset voivat halutessaan deaktivoida tämän seurannan, mikä antaa heille lisäkontrollia omien tietojensa yksityisyydestä. Azure OpenAI:n tietosuojaa edistävät myös erinäiset arkkitehtuuriominaisuudet, kuten Azure Active Directoryn autentikointijärjestelmä ja verkon eristys (Network Isolation). Valavala (2023) toteaa, että nämä ominaisuudet ovat jo laajasti käytössä eri organisaatioissa, mikä tekee Azure OpenAI:n integroimisen olemassa oleviin järjestelmiin sujuvaksi ja tehokkaaksi ratkaisuksi.

2.11.1 Microsoftin sopimukset

Azure on Microsoftin pilvipalvelualusta, joka on sitoutunut tarjoamaan selkeitä ja turvallisia käyttöehtoja ja sopimuksia asiakkailleen. Yritys korostaa tietosuojan merkitystä, antaen asiakkaille mahdollisuuden hallita ja poistaa henkilökohtaisia tietojaan palveluistaan. Asiakas säilyttää tiedon omistajuuden ja on vastuussa niiden hallinnasta, noudattaen samalla Microsoftin palvelujen käytölle asetettuja laillisia rajoituksia. Tietosuojalausekkeet vahvistavat asiakkaiden määräysvallan omien tietojensa suhteen. (Microsoft 2024a.)

Microsoftin tietojenkäsittelysopimuksessa (2024b.) on määritelty ehdot, jotka velvoittavat yhtiön varmistamaan tiedonkäsittelyn turvallisuuden. Sopimus korostaa Microsoftin käyttävän teknisiä ratkaisuja ja menetelmiä, jotka täyttävät tietosuojalakien vaatimukset:

> *Lisäksi näiden* (tietoturva*) toimenpiteiden tulee noudattaa standardien ISO 27001, ISO 27002, ja ISO 27018 vaatimuksia. Näiden vaatimusten edellyttämien tietosuojamenetelmien ja -käytäntöjen kuvaus on asiakkaiden saatavilla.* (Microsoft 2024b. 8)

Tietojenkäsittelysopimuksen liitteessä 1 käsitellään EU:n yleisen tietosuoja-asetuksen ehtoja. Sen perusteella Microsoftin suorittama henkilötietojen käsittely noudattaa EU:n lainsäädännön mukaisia tietosuojaa koskevia ehtoja. Erityisesti yhtiö sitoutuu toteuttamaan kaikki tietosuoja-asetuksen (GDPR) 32 artiklassa määritellyt toimenpiteet, kuten asianmukaiset tietoturvatoimet, riskien arvioinnin ja rekisterinpitäjän valvonnan (Microsoft 2024b. 20). Yhteenvetona voidaan todeta, että sopimusten perusteella Microsoft on sitoutunut takaamaan turvallisuuden myös Azure OpenAI -palvelussa käytettävälle datalle.

Azure OpenAI sertifikaatit sekä konesalien sijainti luomassa lisäturvaa

Azure OpenAI palvelu täyttää Compliance Offering-dokumentaation perusteella useita standardeja, joita nykypäivän yritysten on noudatettava (Yuen & Sanghavi 2024, 9.) Alla olevasta kuvasta nähdään, että Azure OpenAI täyttää esimerkiksi useita eri ISO/IEC standardeja, joista merkittävänä on ja aiemmin mainittu ISO/IEC 27001-sertifikaatti tietoturvan hallintajärjestelmä sekä sen jatke ISO/IEC 27701, joka tukee GDPR vaatimuksia.

Microsoft Azure Compliance Offerings

Azure Service	CSA STAR Certification	CSA STAR Attestation	ISO 20000-1:2018	ISO 22301:2019	ISO 27001:2022	ISO 27017:2015	ISO 27018:2019	ISO 27701:2019	ISO 9001:2015	SOC 1, 2, 3	GSMA SAS-SM	HIPAA BAA	HITRUST	K-ISMS	PCI 3DS	PCI DSS	Australia IRAP	Germany C5	Singapore MTCS Level 3	Spain ENS High	Singapore OSPAR
Azure Object Anchors													✓								
Azure Open Datasets	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓		✓	✓		✓
Azure OpenAI Service	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓		✓	✓		✓

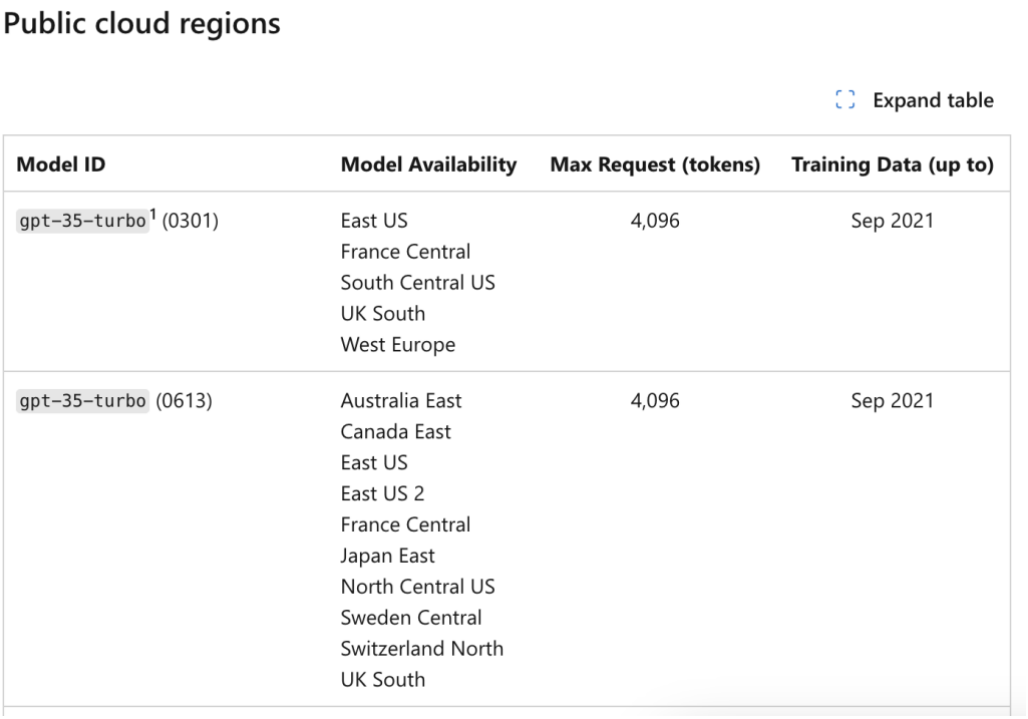
Kuva 4. Azure OpenAI:n täyttämät sertifikaatit (Yuen & Sanghavi 2024, 9)

Azuren dokumentaatiosta puuttuu maininta GDPR regulaatiosta, joka määrittele henkilötietojen käsittelyä, mutta dokumentaatio sisältää esimerkiksi ISO/IEC 27701-standardin, joka on suunniteltu auttamaan organisaatioita täyttämään GDPR:n vaatimukset. Microsoftin työntekijä on kommentoinut GDPR kysymystä heidän sivustollansa seuraavasti:

> *Microsoft ottaa tietosuoja- ja tietoturvakysymykset erittäin vakavasti ja on toteuttanut toimenpiteitä varmistaakseen, ettei henkilökohtaisesti tunnistettavaa tietoa jaeta OpenAI:lle Azure OpenAI -palvelun käytön aikana. Ymmärrän täysin huolesi GDPR:stä ja yksityisyyden noudattamisesta. Microsoft suhtautuu tietosuojaan vakavasti, ja Azure-palvelut, mukaan lukien OpenAI, on suunniteltu noudattamaan GDPR-säädöksiä.* (Norris 2023.)

Yhteenvetona voidaan todeta, että Azure OpenAI on suunniteltu toimimaan GDPR-säädösten mukaisesti. Vaikka Azure OpenAI:n suoraa GDPR-sertifiointia ei erikseen mainita, viittaa Microsoft useilla sivuillaan siihen, että Azure yleisesti tukee GDPR:ää ja on saavuttanut ISO/IEC-27701-sertifikaatin, mikä on osoitus tietoturvan ja tietosuojan hallinnasta (Mazzoli ym. 2024). On tärkeää huomata, että viittaus koskee koko Azurea, eikä erityisesti Azure OpenAI:ta. Microsoft korostaa käyttäjien roolia GDPR-vaatimusten täyttämisessä heidän käyttäessään palveluita.

ISO/IEC-27701-sertifikaatin lisäksi yritykset voivat valita, missä maassa niiden data sijaitsee, mikä on erityisen tärkeää EU-alueella toimiville vastuullisille yrityksille. Datakeskukset on sijoitettu useisiin maihin, ja esimerkiksi Suomen kannalta lähin datakeskus löytyy keski-Ruotsista. Kuvasta 5 ilmenee, että GPT-3.5 Turbo -version kielimallia ylläpidetään useissa eri maissa ympäri maailmaa (Microsoft 2024c).



Public cloud regions

Expand table

Model ID	Model Availability	Max Request (tokens)	Training Data (up to)
gpt-35-turbo[1] (0301)	East US France Central South Central US UK South West Europe	4,096	Sep 2021
gpt-35-turbo (0613)	Australia East Canada East East US East US 2 France Central Japan East North Central US Sweden Central Switzerland North UK South	4,096	Sep 2021

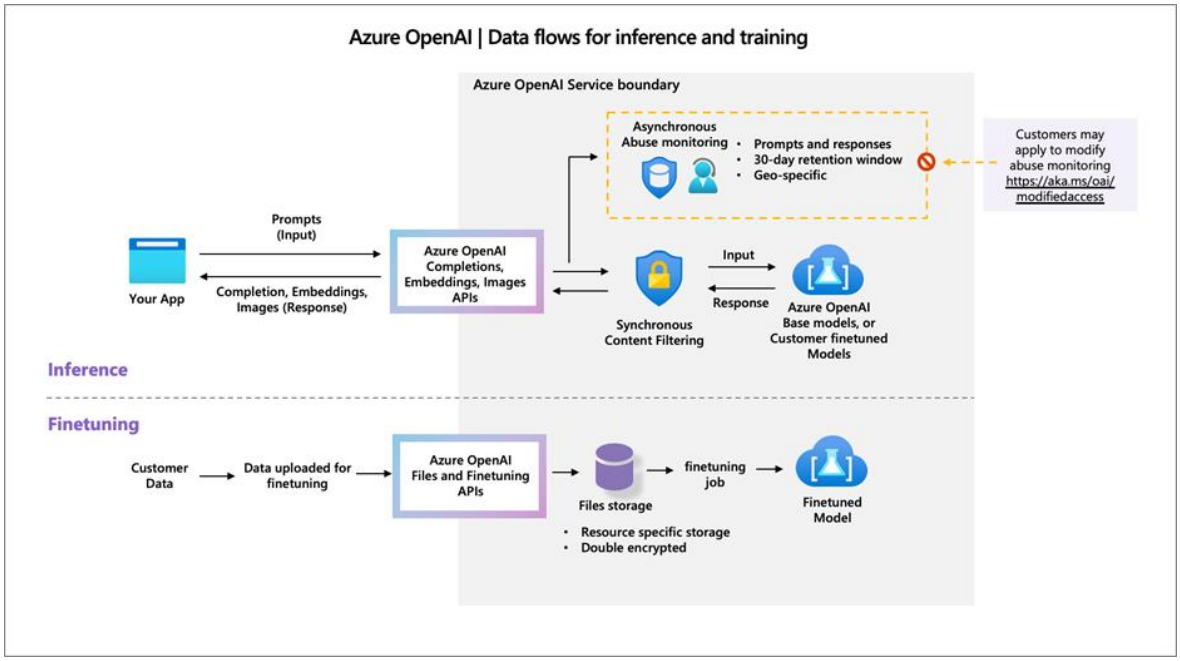
Kuva 5. Pilvipalvelun sijainti vaihtoehdot (Microsoft 2024c.)

2.11.2 Datan kulkeminen Azure OpenAI:n läpi

Microsoftin verkkosivut tarjoavat avointa tietoa siitä, miten dataa käsitellään Azure OpenAI -sovelluksessa. Yritys on julkistanut arkkitehtuurikuvauksen (kuva 6), joka selventää datan kulkua sovelluksen käytössä sekä datan kulkua Finetunting-mallin koulutusvaiheessa. Kuvasta ilmenee, että sovellukseen syötetty teksti eli prompti lähtee internetin välityksellä API rajapinnan välityksellä Azure OpenAI:hin, jossa suoritetaan ensisijainen suodatus. Tämän jälkeen käyttäjän prompti siirtyy kielimalliin, josta suodatetun vastauksen saa asiakkaan sovellus. Suodatuksen välissä on vielä Microsoftin oma väärinkäytön valvonta (Abuse Monitoring), jota käsitellään omassa kappaleessa tarkemmin. (Farley ym 2023.)

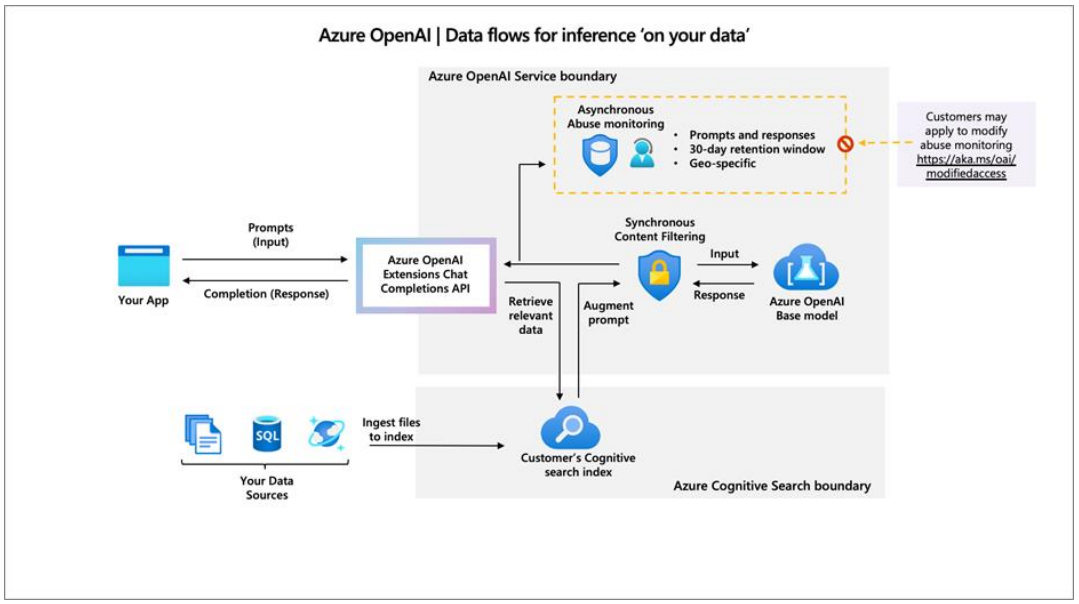
Kuvassa 6 on myös mallinnettu prosessia nimeltä “Data flow for training”, joka kuvaa Finetuning-mallin koulutusprosessia. Finetuning viittaa tekoälymallin räätälöintiin, joka vastaa paremmin yrityksen erityistarpeisiin käyttämällä yrityksen omaa dataa. Tämä vaatii mallille syötettyä koulutus dataa. Kuvan 6 perusteella koulutusdata säilötään Azure-pilvessä kaksoiskryptattuna koulutus materiaalille osoitetussa lokaatiossa. Organisaatioilla on oikeus ja mahdollisuus poistaa kyseinen koulutusaineisto Azuren palvelimilta milloin tahansa. On tärkeä huomioida, että Azure OpenAI ei hyödynnä käyttäjien syöttämiä

prompteja eikä yritysten dataa kielimallin kouluttamiseen, eikä mitään dataa jaeta kolmansien osapuolien esimerkiksi OpenAI:n kanssa. (Farley ym 2023.)



Kuva 6. Azure OpenAI-dataprosessi (Farley ym. 2023)

Alla oleva kuva 7 mallintaa, Azure OpenAI datankäsittelyn prosessia tilanteessa, jossa kielimalli integroidaan organisaation datan kanssa. Mallissa käyttäjän sovelluksesta lähetetty prompti siirtyy API-rajapinnan kautta Azure OpenAI-palveluun. Täällä tekoäly generoi vastauksen kielimallin sekä organisaation datan perusteella, minkä jälkeen asiakkaalle lähetetään vastaus. Prosessin aikana Microsoftin suodattimet tarkkailevat järjestelmän väärinkäyttöä, jottei järjestelmää käytetä laittomiin tarkoituksiin. Asiakkailla on mahdollisuus mukauttaa Microsoftin suodattimien asetuksia tarpeidensa mukaisiksi, mikä lisää joustavuutta ja kontrollia heidän omaan dataansa. (Farley ym. 2023.)



Kuva 7. Arkkitehtuuri personoidun kielimallin tilanteessa (Farley ym. 2023)

2.11.3 Azuren tietoturvatuotteet

Abuse monitoring

Abuse monitoring on Microsoftin kehittämä järjestelmä, jonka on tarkoitus lisätä palvelun kokonaisturvallisuutta. Microsoft kertoo artikkelissaan (2023), että järjestelmä havaitsee toistuvia käyttäytymissääntöjen (code of conduct) rikkomista. Yleiset käyttäytymissäännötkoostuvat eettisitä ohjeista, joissa kielletään seksuaalisuus, väkivalta, ilkeys ja manipulointi. Abuse monitoring -toiminto luokittelee sisään tulevaa dataa sen kielen perusteella ja pyrkii tunnistamaan väärinkäytön algoritmeja hyödyntäen.

Jos väärinkäytön indikaattorit täyttyvät, puuttuu asiaan todennäköisesti Microsoftin valtuutettu työntekijä. Henkilöt tarkastavat väärinkäytetyn sisällön ja ovat yhteydessä asiakasyritykseen. Nämä henkilöt työskentelevät eurooppalaisten organisaatioiden tilanteissa Euroopan sisäpuolella ja he tarvitsevat tiedon näkemiseen esimiehen erillisen luvan.

Abuse monitoring eli väärinkäytön valvonta on mahdollista ottaa pois käytöstä, mikä voi parantaa organisaation tietoturvaa. Kun väärinkäytön valvonta on poistettu käytöstä, Microsoft ei tallenna prompteja eikä vastauksia kyseisen palvelun käyttäjiltä. Tällöin Microsoftin työntekijöiden suorittama tarkastusprosessi ei ole käytännössä mahdollista, koska tietoja ei säilytetä lepotilassa Microsoftin Service Results Store -tietokannoissa. Tämä poikkeuslupa edellyttää, että tekoälyn väärinkäytön riski organisaatiossa on pieni.

Lisäksi toiminnon poistaminen voi olla hyödyllistä, jos valtiossa on voimassa olevia lakeja, jotka rajoittavat Microsoftin turvallisuusvalvontaa. (Farley ym 2023.)

Datan automaattinen kryptaus

Microsoft käyttää FIPS 140-2 -standardin mukaista 256-bittistä AES-salausta suojatakseen pilvessä säilytettyä koulutusdataa, Kun yritykset kouluttavat kielimallia omalla datallaan. Tämä salausmuoto muuttaa datan alkuperäisestä muodostaan tunnistamattomaksi, mikä tarkoittaa, että vaikka tiedot vuotaisivat, niitä ei voida ymmärtää ilman oikeaa salausavainta. Salausavaimet ovat oletusarvoisesti Microsoftin hallinnassa, mikä varmistaa datan turvallisuuden. Tiedon tallentaminen Azuren palveluun tulee tarpeelliseksi Finetunting-mallin käytössä ja sen kehittämisvaiheessa, jolloin kielimallia halutaan opettaa oman organisaation datalla (Urban ym. 2023.)

Asiakasyrityksellä on mahdollisuus ottaa käyttöön customer-managed keys (CMK), jolloin yritys pystyy itse hallinnoimaan kryptausavaimia. CMK:n avulla yritys voi itse luoda, kiertää, poistaa ja tarkastaa pääsynhallintaa salausavaimille. Microsoft painottaa, että datan säilyttäminen kryptatussa muodossa auttaa täyttämän tietoturvavelvoitteet, joita tietyt regulaatiot saattavat vaatia. (Urban ym. 2023.)

3 Haastattelututkimus Aalto IT:n organisaatiossa

3.1 Haastattelututkimuksen kuvaus

Opinnäytetyön teoreettista viitekehystä haastaa asiantuntijahaastattelut, joiden avulla pyrittiin syventämään ymmärrystä tutkimuksen aiheeseen liittyen sekä löytämään uusia kiinnostavia näkökulmia. Haastattelututkimuksen kysymykset johdettiin opinnäytetyön teoriapohjasta, joka lisää tutkimuksen läpinäkyvyyttä, mahdollistaen vastakkaiset näkemykset teorialähteiden ja asiantuntijoiden välillä. Laadullisen tutkimuksen aineisto kerättiin haastattelemalla Aalto IT:n asiantuntijoita, jotka pyrkivät selventämään miksi pilvitekoälyn tietosuojariskejä sekä lokaalin tekoälyn mahdollisia ominaisuuksia

Tutkimukseen osallistui viisi asiantuntijaa, jotka valitsin heidän kokemuksensa perusteella. Haastateltavien asiantuntemus kattoi laaja-alaisesti pilvipalveluiden, tekoälyn ja tietoturvan eri osa-alueita. Aineistonkeruu suoritettiin yksilöllisten haastattelujen muodossa, joiden aikana tehtiin yksityiskohtaisia muistiinpanoja asiantuntijoiden vastauksista.

3.2 Vastausten analysointi

Asiantuntijahaastatteluista saadut vastaukset analysoitiin teemoittain, mikä auttoi vahvistamaan tutkimuksen rakenteellista tiiviyttä ja tärkeimpien löydösten esiin tuomista. Yhtenäisen kysymysrungon käyttäminen kaikissa haastatteluissa varmisti vastausten vertailtavuuden ja johdonmukaisuuden.

Haastattelujen alussa käsiteltiin yleisiä käsityksiä tekoälystä ja sen integroinnista organisaatioiden toimintaan. Sen jälkeen keskustelu keskittyi syvällisemmin Azure OpenAI:n tietoturvaan ja lokaalin tekoälyn hyödyntämiseen, tavoitteena ymmärtää niiden vaikutuksia ja roolia organisaatioiden tietoturvapolitiikassa. Käytetyt haastattelukysymykset ja niiden yksityiskohtainen sisältö ovat esitetty opinnäytetyön liitteessä 1.

Mielipide tekoälystä

Tutkimukseen osallistuneet asiantuntijat osoittivat huomattavaa mielenkiintoa opinnäytetyötä ja sen aihetta kohtaan, mikä näkyi heidän halukkuutenaan osallistua haastatteluihin. Eritoten tekoälyn tietoturvaa koskevat kysymykset herättivät vilkasta keskustelua ja näkökulmien vaihtoa.

Haastateltavat yhtyivät näkemykseen siitä, että tekoäly edustaa myönteistä kehitystä, jolla on potentiaalia muokata maailmaamme perustavanlaatuisella tavalla tulevaisuudessa. He ovat kiinnittäneet huomiota erityisesti kielimallien nopeaan kehitykseen viime aikoina, pitäen sitä merkittävänä esimerkkinä tekoälyn kehittymisestä. Asiantuntijat ilmaisivat

näkemyksensä siitä, että tekoäly tulee vähitellen korvaamaan monia rutiininomaisia työtehtäviä ja tehostamaan prosessien suorituskykyä monilla eri toimialoilla.

Datan turvaaminen organisaatioissa

Haastattelussa Aalto-yliopiston asiantuntijat painottivat datan suojelun kriittistä merkitystä, korostaen esimerkiksi tutkimusdataa, joka on yliopiston kallisarvoista omaisuutta ja joka ei saa vuotaa yliopiston ulkopuolelle. Vahva datan luokittelujärjestelmä on heidän mukaansa välttämätön, jotta henkilöstö ymmärtää, miten tietoja saa käsitellä ja hyödyntää.

Regulaatiot ja standardit suojaavat asiantuntijoiden mukaan ihmisiä nyky-yhteiskunnassa, ja esimerkiksi GDPR sekä DPIA-regulaatiot tulisi olla oletusarvoisesti integroituja organisaatioiden päätöksentekoprosesseihin. Asiantuntijat ilmaisivat myös huolensa yksityishenkilöiden verkkokäyttäytymisestä, koska yksityishenkilöt lopulta tekevät päätöksen siitä mihin he omaa dataa syöttävät. Tästä esimerkkinä kiireinen professori tai johtaja, jolla ei ole aikaa lukea kaikkia sähköposteja, joten hän lähettää viikon sähköpostit ChatGPT:lle, jolloin hän rikkoo GDPR regulaatiota jakamalla sähköpostien lähettäjien henkilötiedot tekoälylle.

Azure OpenAI tietoturvan näkökulmasta

Asiantuntijat näkevät Azure OpenAI:n lähtökohtaisesti turvallisena järjestelmänä, sillä se sisältää useita erilaisia suojaustyökaluja, joiden avulla Azureen voidaan luoda yrityksen omassa hallinnassa oleva ympäristö, joka ei tallenna dataa minnekään, poikkeuksena koulutusdata, jonka yritys voi halutessaan poistaa. Vaikka Azure OpenAI sisältääkin suojausominaisuuksia, perustuu kaikki viime kädessä luottamukseen palveluntarjoajan ja käyttäjän välillä. Asiantuntijoiden mukaan on tärkeää luoda erittäin tarkat sopimukset kaikista palveluista, mikä lisää luottamusta osapuolten välillä.

GDPR-regulaation mukaisesti Azure-ympäristöä luodessa käyttäjä pystyy itse määrittämään, missä valtiossa data sijaitsee. Tämä on erityisen tärkeää EU-alueen yrityksille, jotka haluavat luoda mahdollisimman turvallisen ympäristön. Asiantuntijat ymmärtävät myös Azure OpenAI:n riskit, ja nämä on otettu huomioon Aalto GPT:n kehityksessä, varmistaen, että tietoturvariskit minimoidaan.

Pilvipalveluiden yleinen haaste on datan poistamisen vaikeus palveluntarjoajan tietokannoista, mikä ei aina ole mahdollista, vaikka GDPR sitä vaatiikin. Tämän varmistamiseen käyttäjillä ei ole monia vaihtoehtoja, mutta Azure vakuuttaa sopimuksissaan, että asiakkaalla on mahdollisuus poistaa omat datansa heidän palveluistaan.

Aalto GPT, joka on Aalto-yliopiston hallinnoima palvelu, ei tallenna käyttäjien dataa (promptit ja vastaukset), mikä minimoi tietovuodon riskin Microsoftin kautta. Tämä eliminoi käytännössä datan vuotamisen riskin ulkopuolisille tahoille, sillä kaikki keskusteluhistoria tallentuu vain käyttäjän selaimelle. Useat asiantuntijat myös huomauttavat tietoisuudestaan siitä, että Yhdysvalloilla saattaa olla mahdollisuus tiedustella heidän maaperällään olevaa dataa Patriot Actin, Cloud Actin ja FISA-lakien mukaan. Tämä on yksi riski, joka on otettava huomioon, jos yrityksen data sijaitsee Yhdysvalloissa.

Lisäksi suuret yritykset kykenevät luomaan omat, räätälöidyt pilviympäristönsä, mikä parantaa tietoturvaa verrattuna pienempiin yrityksiin, jotka eivät ehkä taloudellisista syistä pysty investoimaan vastaaviin ratkaisuihin. Pahimmassa tapauksessa pienet yritykset saattavat käyttää yksilöille tarkoitettuja edullisia tekoälypalveluita, mikä lisää riskiä tietovuodoille.

Kriittiset havainnot Azure OpenAI:n tietoturvaan liittyen

Kysyttäessä asiantuntijoilta, miksi yritykset rajoittavat Azure OpenAI:ssa käytettävän datan vain julkiseksi luokiteltuun, yksi keskeinen vastaus oli Azuren abuse monitoring. Tämä valvontajärjestelmä mahdollistaa Microsoftin työntekijöiden tarkastella yrityksen prompteja, mikä herättää epäluottamusta palvelun tietoturvaa kohtaan. Tämä havainto on merkityksellinen tarkasteltaessa Azure OpenAI:n käyttöä ja sen tietoturvakäytäntöjä yritysten näkökulmasta.

Asiantuntijoilta kysyttiin miksi Word-sovellusta voi käyttää salaisen datan kanssa mutta Azure OpenAI:ta ei, vaikka molemmat sovellukset ovat samalta palveluntarjoajalta. Asiantuntijat vastasivat että, Azuren API-yhteydessä tai sovelluksessa on lisäriskejä, jotka eivät tue salaisen datan käyttöä Azure OpenAI:ssa. Word puolestaan sisältää ominaisuuden, jossa tiedoston voi kryptata salasanan avulla. jolloin tiedosto sijaitsee tallennetussa paikassa kryptatussa muodossa. Samaan aikaan Azure OpenAI tarvitsee pääsyn käsiteltävään dataan tekstimuodossa reaaliaikaisesti, mikä heikentää sen soveltuvuutta salaisen tiedon käsittelyyn.

Haastatteluissa tuli ilmi huoli siitä, kuinka Azure erottaa eri yritysten dataa pilvipalvelussaan. Sekä asiantuntijoiden että julkisten lähteiden tarjoama tieto tästä aiheesta on olematonta. Mikäli yritysten datan eristäminen toisistaan on toteutettu Azure-palvelussa turvallisesti, olettaisi tiedon olevan helposti saatavilla ja dokumentoituna. Tämän tiedon puute herättää kysymyksen siitä, voiko datan mahdollinen sekoittuminen Azuren palvelussa luoda tietoturvariskejä. Organisaatioiden on siksi syytä omaksua varovainen lähestymistapa ja tehdä perusteellinen riskiarvio. On myös suositeltavaa kysyä lisätietoja suoraan Azuren edustajilta, mikäli epäselvyyksiä ilmenee.

Ylläpitäjien oikeuksien vaikutus tietoturvaan

Asiantuntijoiden havaintojen mukaan Azure-palvelun yksi potentiaalinen tietoturvariski liittyy ylläpitäjien laajoihin oikeuksiin. Jos näitä oikeuksia käytetään väärin tai ylläpitäjän käyttäjätunnukset päätyvät vääriin käsiin, se voi johtaa merkittäviin turvallisuusuhkiin. Ylläpitäjillä on laaja pääsy organisaation järjestelmiin ja tietoihin, mikä tekee heistä potentiaalisen riskin tietoturvalle. Arkaluontoisen tiedon vuoto tai tietomurrot ovat realistisia seurauksia, mikäli ylläpitäjien tunnukset vuotavat väärille tahoille

Turvallisuusriskien hallitsemiseksi on välttämätöntä toteuttaa tiukat ylläpitäjätilien hallinnointikäytännöt, jotka sisältävät käyttöoikeuksien minimoinnin välttämättömään toiminnallisuuteen sekä ylläpitäjien toiminnan jatkuvan seurannan. Tämä ei ainoastaan paranna tietoturvaa vaan myös auttaa ennaltaehkäisemään mahdollisia väärinkäytöksiä. Lisäksi on suositeltavaa ottaa käyttöön monivaiheinen todentaminen ja muita edistyksellisiä turvamekanismeja, jotka tehostavat ylläpitäjätilien suojelua ja minimoiden niiden väärinkäytön riskit.

Asiantuntijan näkökulma lokaalista tekoälystä

Asiantuntijat ilmaisivat positiivisen suhtautumisen lokaaliin tekoälyyn. Haastattelun aikana keskustelimme lokaalin tekoälyn toimintaperiaatteista asiantuntijoiden kanssa, ja he kokivat kielimallin sijainnin Aallon sisällä merkitykselliseksi. Lokaalin tekoälyn käyttöönotto Aallossa on mahdollista, mikäli se läpäisee tietoturvatarkastuksen ja sille on nähtävissä todellinen tarve. Asiantuntijat myös suhtautuivat positiivisesti ajatukseen salaiseksi luokitellun datan käytöstä paikallisen tekoälyn kanssa. He eivät pitäneet tätä ongelmana, mikäli paikallinen malli täyttää kaikki Aalto-yliopiston tietoturvastandardit.

Asiantuntijat kertoivat, että lokaalin kielimallin täytyy pysyä tuottamaan laadukkaita vastauksia, jotka riippuvat sille syötetystä datasta. Useiden asiantuntijoiden mukaan lokaali tekoäly saattaa olla tehoton ja vastaukset heikompia verrattuna Azure OpenAi:hin. Eräs asiantuntija mainitsi, että lokaali tekoäly on huomattavasti halvempi vaihtoehto raskaassa datan analysoinnissa verrattuna esimerkiksi GPT 4 mallin hinnoitteluun, jos järjestelmä on raskaalla käytöllä. Toisaalta osa asiantuntijoista painotti Azure OpenAI:n olevan halvempi tämän hetken käyttäjämäärällä, jolloin lokaalia mallia ei ehkä vielä kannattaisi rakentaa. Kustannukset voivat kuitenkin vaihdella pilvipalveluissa, kun taas lokaalin mallin kustannuksia on helpompi arvioida.

Erään asiantuntijan mukaan Aallossa on valmiudet tuotteistaa lokaali tekoäly organisaation käyttöön, jos sille nähdään tarve. Tekninen valmius on jo olemassa Aallon lisäksi useissa muissakin yrityksissä mutta, hankkeella pitää pystyä tuottamaan lisäarvoa yritykselle.

Lokaalin tekoälyn tuotteistamisen ja sen hyötyjen avaaminen verrattuna pilvitekoälyyn täytyy olla tarpeeksi merkittävä, jotta tällainen hanke aloitetaan. Azure OpenAI -ympäristö nähdään turvallisena ympäristönä, johon on panostettu paljon. Tulevaisuudessa oman paikallisen mallin toteuttaminen on mahdollista, ja asiantuntijat uskovat, että tällöin voitaisiin käsitellä myös salaiseksi luokiteltua dataa lokaalissa tekoälymallissa.

Organisaatioiden vastuu tietoturvasta

Asiantuntijat kertoivat, että organisaation käytössä olevista sovelluksista on aina pääsääntöisesti vastuu itse organisaatiolla. Jos tietoa vuotaa esimerkiksi pilvipohjaisesta tekoälystä, on Aalto-yliopisto vastuussa mahdollisista tietoturvapoikkeamista. Pilvipalveluiden käyttöönotossa on tärkeää luoda tarkat sopimukset, joissa määritellään tilanteet, joissa palveluntarjoaja on korvausvelvollinen organisaatioille. Tämä ei kuitenkaan pelasta jo vuotanutta tietoa, joten organisaatioiden on aina arvioitava mahdolliset riskit uusia järjestelmiä otettaessa käyttöön. Tähän tarkoitukseen sopii mainiosti DPIA, eli tietosuojavaikutusten arviointi.

Asiantuntijoiden mukaan organisaatioiden vastuulla on myös kouluttaa työntekijänsä toimimaan vastuullisesti verkossa laatimalla säännöt ja ohjeet, jotka tukevat organisaation tietoturvapolitiikkaa. Asiantuntijat korostivat yksilön käyttäytymisen merkittävyyttä ja toivoivat yksilöiden arvioivan aina tekoälystä saatavaa hyötyä suhteessa siihen liittyviin riskeihin. Esimerkiksi patenttidokumentaation käyttäminen ChatGPT:ssä ei missään nimessä ole järkevää, jos siitä saatava hyöty on vähäinen riskeihin verrattuna.

Asiantuntijat antoivat kaksi vinkkiä tekoälyn käyttöön kaikille henkilöille. Ihmisten pitäisi aina miettiä, saako kyseinen data vuotaa, ja muistaa, että tekoälyt voivat tehdä virheitä, joten vastaukseen ei kannata luottaa sokeasti.

4 Yhteenveto ja pohdinta

4.1 Yhteenveto

Tämä opinnäytetyö laadittiin yhteistyössä Aalto-yliopiston kanssa, jossa tutkimuksen alkuvaiheessa oli mahdollista käyttää Aalto GPT:tä ja ChatGPT:tä ainoastaan julkisen datan kanssa. Tutkimuksen ensisijaisena tavoitteena oli selvittää, kuinka Azure OpenAI käsittelee dataa ja miksi vain julkisen datan käyttö on sallittua kyseisessä sovelluksessa. Toisena tavoitteena oli paneutua lokaalin tekoälyn mahdollisiin ratkaisuihin, joita se voisi tuoda tekoälytarjontaan tietoturvanäkökulmasta. Yrityksillä ja organisaatioilla on laajasti epäluuloja AI-palveluita kohtaan, mikä on ymmärrettävää ottaen huomioon ChatGPT:n tunnetut tietovuotokohut. Ongelman taustalla yksityishenkilöillä on mahdollisuus käyttää mitä tahansa palveluita henkilökohtaisessa elämässään, kun taas organisaatioita sitovat regulaatio ja lait, joiden myötä heidän datan käyttöön ja tietoturvaan liittyy tiukemmat ehdot. Näitä regulaatioita edustaa esimerkiksi GDPR, joka velvoittaa yritykset toimimaan tietoturvallisuuden huomioon ottaen.

Opinnäytetyön aikana määriteltiin teoriaosuudessa tekoälyä koskevat käsitteet, minkä jälkeen tutkimus keskittyi Azure OpenAI-palvelun toimintaan tietoturvan näkökulmasta. Tekoäly palveluntarjoajista Microsoftin kehittämä Azure OpenAI on alansa yksi suurimmista toimijoista ja suosittu valinta yrityksissä, joissa on jo valmiiksi käytössä Microsoftin ympäristö.

4.2 Oivalluksia ja näkökulmia

Opinnäytetyön tutkimusprosessi on valottanut Azure OpenAI:n ja lokaalin tekoälyn käyttöön liittyviä tietoturva- ja tietosuojakysymyksiä, tarjoten samalla syvällisempää ymmärrystä näiden teknologioiden mahdollisuuksista ja rajoitteista. Opinnäytetyön teoriaosassa on käynyt ilmi, että vaikka pilvipohjainen tekoäly tarjoaa skaalautuvuutta ja joustavuutta, se myös altistaa organisaatiot tietosuojahaavoittuvuuksille. Tämän vuoksi lokaalin tekoälyn merkitys tietoturvan ja -suojan parantamisessa korostuu, tarjoten ratkaisun, jossa data pysyy tiukasti organisaation hallinnassa. Pilvipalveluista on tarjolla lukuisia tietoturvaratkaisuja esimerkiksi datan liikkumisessa ja säilyttämisessä, mutta ne eivät silti yllä lokaalin mallin turvallisuustasolle, koska koko infrastruktuuri on silloin yrityksen omassa hallinnassa. Tämä vähentää ulkoisten uhkien riskiä ja luo luottamusta järjestelmään, jos organisaation oma tietoturva on korkealla tasolla.

Yksi merkittävä oivallus on, että tekoälyn käyttö ei ole mustavalkoinen valinta pilven ja lokaalin välillä, vaan organisaatioiden tulisi harkita hybridiratkaisuja, jotka yhdistävät molempien järjestelmien parhaat puolet. Esimerkiksi salaisen datan käsittely voitaisiin suorittaa lokaalisti, kun taas vähemmän arkaluontoiset tehtävät voisivat hyödyntää pilvipalveluiden tarjoamaa tehokkuutta.

Toinen tärkeä näkökulma on tietoisuuden lisääminen tekoälyn tietosuojakysymyksistä sekä käyttäjien, että kehittäjien keskuudessa. Organisaatioiden on panostettava ihmisten koulutukseen ja ohjeistukseen, jotta kaikki sidosryhmät ymmärtävät tekoälyn käytön mahdolliset riskit ja parhaat käytännöt. Esimerkiksi Aalto-yliopistoon saapuvien kansainvälisten opiskelijoiden tulee ymmärtää datan luokittelua ja tekoälyn käyttöön liittyvät säännöt.

Lisäksi on osoittautunut, että lainsäädännön ja standardien, kuten GDPR:n ja ISO/IEC 27001:n, rooli tekoälyn turvallisessa käytössä on ensiarvoisen tärkeää. Näiden ohjeistusten sekä standardien noudattaminen ei ainoastaan vähennä tietoturvariskejä, vaan myös rakentaa luottamusta teknologiaa kohtaan niin käyttäjien kuin organisaatioiden keskuudessa.

Tutkimus on mielestäni osoittanut, että tekoälyn eettisyys ja vastuullisuus ovat keskeisiä arvoja, jotka ovat ratkaisevia teknologian kestävän kehityksen kannalta ja jotka määrittelevät palveluiden regulaatioita. Myös Euroopan unioni tukee tätä kehityssuuntaa. Siksi organisaatioiden ja viranomaisten on oltava aktiivisia ei vain teknologisten innovaatioiden hyödyntämisessä, vaan myös niiden eettisten ulottuvuuksien pohdinnassa. Kokonaisuudessaan opinnäytetyö on osoittanut, että vaikka tekoäly tarjoaa huomattavia hyötyjä, sen turvallinen ja vastuullinen käyttö vaatii jatkuvaa valppautta, innovatiivisia ratkaisuja ja kaikkien osapuolten yhteistyötä. Tulevaisuudessa tekoälyn kehityksen ja käytön on kuljettava käsi kädessä eettisten periaatteiden ja tietosuojan parhaiden käytäntöjen kanssa, jotta voidaan maksimoida sen hyödyt minimoiden samalla potentiaaliset riskit.

4.3 Tutkimustulos

Teoreettisen viitekehyksen sekä IT-asiantuntijoiden haastattelujen perusteella Azure OpenAI:ta voidaan pitää tietoturvan näkökulmasta hyvänä vaihtoehtona yrityksille. Tieteelliset artikkelit, verkkohaut tai asiantuntijat eivät pysty osoittamaan Azure OpenAI:n merkittäviä tietoturvaongelmia.

Azure OpenAI säilyttää tiedot EU-alueella, ei luovuta niitä kolmansille osapuolille ja tarjoaa yrityksille mahdollisuuden luoda yksityisen ympäristön Azure-pilvipalveluun. Tiedonsiirto ja

säilytys tapahtuvat salatusti, ja yritykset voivat hallita salausavaimia, kun käytössä on fine-tuned-versio. Mikäli abuse monitoring -toiminto on kytketty pois päältä, Microsoft ei pysty näkemään asiakkaan tietoja edes väärinkäytöstilanteissa, koska silloin käyttäjän syötteitä ei tallenneta mihinkään. Tämän tutkimustuloksen perusteella voidaan todeta, että luokitellun datan käyttö Azuren-pilvitekoälyssä on tietoturvallisesti mahdollista.

Kuten todettu, tutkimuksen lähtötilanteessa joulukuussa 2023 Aalto-yliopiston ohjeistus Aalto GPT:n käytölle oli rajoitettu ainoastaan julkiseen dataan. Nyt tutkimusprosessin aikana Aalto-yliopisto teki omia selvityksiään, joiden perusteella on tehty uusia linjauksia luokitellun datan käyttöön kyseisessä sovelluksessa. Maaliskuussa 2024 tulleen uuden linjauksen myötä käyttäjille annettiin lupa käyttää järjestelmässä myös Aallon tiedonluokittelu määrittämää sisäistä ja luottamuksellista dataa, jotka löytyvät taulukosta 1.

Julkinen	Sisäinen	Luottamuksellinen	Salainen
Esimerkiksi: • julkaisut • tiedotteet • koko yliopistoa koskevat päätökset	Esimerkiksi: • työtiedostot • luonnokset • muistiot • tiedot, joita ei julkisteta	Esimerkiksi: • henkilötiedot • liikesalaisuudet • neuvottelutiedot • tutkimussuunnitelmat • yksityiskohtaiset järjestelmätiedot	Esimerkiksi: • yksityisyyden suojan piiriin kuuluvat tiedot (potilas- ja terveystiedot) • arkaluonteinen tutkimustieto • turvallisuustiedot

Taulukko 1. Aallon tiedon luokittelu (Aalto-yliopisto 2020)

Aalto-yliopistolle sisäisen ja luottamuksellisen datan, kuten henkilötietojen, käyttö Aalto GPT-palvelussa edustaa merkittävää muutosta. Tämän opinnäytetyön yhtenä tavoitteena oli tarjota Aalto-yliopistolle taustatietoa Azure OpenAI-palvelun tietoturvaominaisuuksista. Alkuperäisesti oletin, että kyseinen muutos tapahtuisi myöhemmin, mutta sen toteutuminen jo tutkimuksen aikana oli positiivinen yllätys. Tämä opinnäytetyö tukee Aalto-yliopiston päätöstä hyödyntää myös luottamuksellista dataa Azure OpenAI:ssa, perustuen edellä esitettyihin perusteluihin. Neljän kuukauden perehtymisen aikana Azure OpenAI -palvelussa ei havaittu merkittäviä tietosuojariskejä, jotka olisivat haitallisia Aalto-yliopistolle.

Lokaalia tekoälyä voidaan pitää puolestaan tietoturvan näkökulmasta järkevänä ratkaisuna organisaatioille, joilta löytyy tarvittavat resurssit kyseisen palvelun tuottamiseen. Kun kielimalli sijaitsee käyttäjän omalla laitteella, päästään tällöin minimoimaan tietovuodon riski kolmannen osapuolen kautta ja saavutetaan näin ollen hieman parempi tietoturva. Toisaalta

tutkimuksen aikana selveni että loppukäyttäjän oma toiminta on edelleen yksi merkittävä riski jos henkilökunta jakaa dataansa ei hyväksytyissä palveluissa.

Opinnäytetyö saavutti sille asetetut tavoitteet, eli selvensi Azure OpenAI:n riskien olevan suhteellisen vähäiset, jota tukee myös Aalto-yliopiston päätös sallia luottamuksellisen datan käyttö Azure OpenAI:ssa. Teoreettisen viitekehyksen ja asiantuntijoiden perusteella lokaali tekoäly voi vähentää organisaatioiden tietoturvariskejä erityisesti salaisen datan kanssa, koska data ei poistu silloin organisaatiosta. Organisaatioiden on itse tehtävä kartoitusta ja mietittävä, onko heillä tarvetta käyttää tekoäly palveluita salaisen datan kanssa.

Opinnäytetyön tulokset perustuvat teoreettiseen viitekehykseen ja asiantuntijahaastatteluihin, jotka ovat samassa linjassa keskenään ja tukevat tutkimustulosta. Vaikka teoreettiset lähteet ja asiantuntijoiden näkemykset ovat pääosin samansuuntaisia, jotkin asiantuntijat suhtautuivat kriittisemmin pilvi- ja lokaalin tekoälyn käyttöön verrattuna teoreettisiin näkemyksiin.

Asiantuntijoiden kommenttien perusteella tämän opinnäytetyön aihe vaikutti ajankohtaiselta ja mielenkiintoiselta. Eräs asiantuntija mainitsi, että aihe on haastava, sillä siitä on saatavilla vain vähän tietoa, vaikka suuri osa maailman yrityksistä kohtaa nämä ongelmat suunnitellessaan tekoälyn käyttöönottoa. Opinnäytetyö auttaa organisaatioita tekemään päätöksiä tietoturvaratkaisujen ja tietosuojakäytäntöjensä suhteen tekoälypalvelujen kontekstissa.

4.4 Tulosten käytettävyys muille organisaatioille ja eettisyyden pohdintaa

Opinnäytetyössä havaittiin, että pienemmillä organisaatioilla voi olla haasteita luoda yhtä turvallisia tekoälypalveluita kuin mitä suuret pilvipalvelut, kuten Azure, pystyvät tarjoamaan. Azure OpenAI tarjoaa monipuolisia tietoturvaominaisuuksia, jotka voivat olla hankalia sekä kalliita toteuttaa pienemmässä mittakaavassa omalla infrastruktuurilla. Tämän vuoksi on tärkeää, että pienemmät organisaatiot arvioivat tietoturvan tarpeitaan ja harkitsevat, miten ne voivat hyödyntää ulkoisia palveluntarjoajia, kuten pilvipalveluita tietoturvan tason nostamiseksi. Eräs asiantuntija myös painotti, että harvat yritykset käsittelevät tiedonluokittelultaan salaista dataa, jolloin voidaan ajatella turvallisen pilvitekoälyn olevan riittävä tällaisen yrityksen tarpeisiin.

Tutkimuksessa huomattiin, että organisaatioiden on suositeltavaa kehittää yhtenäiset käsitteet ja toimintatavat IT-hankkeiden parissa toimiessaan. Tämän opinnäytetyön haastatteluissa havaittiin, että kun asiantuntijoita on monia ja käsitteet vaihtelevat, yhtenäinen kielenkäyttö parantaa organisaation tehokkuutta ja tuottavuutta. Lisäksi tarkasti

määritellyt käsitteet esimerkiksi datan luokittelusta lisäävät organisaation tietoturvaa, kun kaikki työntekijät ymmärtävät mitä dataa saa käsitellä missäkin järjestelmässä.

Eettisestä näkökulmasta, kun yritykset ottavat käyttöön pilvitekoälyjärjestelmiä, on kriittistä suunnitella projekteja niin, että ne huomioivat mahdollisen sensuroinnin eettisesti arveluttavien käyttötarkoitusten, kuten rikollisten suunnitelmien, estämiseksi. Näissä tapauksissa palveluntarjoajat, kuten Azure, voivat keskeyttää palvelun, jos se havaitsee palvelun käytön sopimusehtojen vastaisesti. Asiantuntijoiden mukaan Azure OpenAI:n sisällönsuodatus (Content Filtering) voi poistaa osan laittomasta sisällöstä, mutta se ei välttämättä ole riittävä suoja kaikkea väärinkäyttöä vastaan. Tähän eettiseen haasteeseen pyrkii vastaamaan Euroopan komission AI ACT -regulaatio, jonka tavoitteena on luoda sääntelyä, joka estää tekoälyn sovellusten käytön epäeettisiin tarkoituksiin. Näin ollen sovellusten väärinkäyttöön on pyritty varautumaan jo palveluntarjoajan tasolla. Toisaalta on olennaista myös huomioida, että jos lokaalissa tekoälyssä ei ole filttereitä, saattaa tämä lokaali malli antaa herkemmin epäkohteliaita vastauksia.

Järjestelmän eettisyys on keskeinen tekijä pilvitekoälyn käyttöönottopäätöksissä, ja se tulisi ottaa huomioon hankintaprosesseissa sekä sisällyttää huolellisesti sopimuksiin. Tämä takaa, että tekoälyn käyttö on vastuullista ja eettiset periaatteet sekä sääntelyt ovat huomioitu, sitoen yrityksen noudattamaan asianmukaisia lakeja ja käytäntöjä.

4.4.1 Lokaalin tekoälyn käyttöönotto yrityksissä

Lokaalin tekoälyn käyttöönotto tarjoaa merkittäviä etuja, kun käsitellään luokiteltua tietoa. Yksi tärkeimmistä eduista on kyky prosessoida salaista tietoa, kuten terveysdataa, ilman että kolmannen osapuolen puuttumista tarvitaan, mikä parantaa tietoturvaa. Lisäksi lokaali tekoäly mahdollistaa palveluiden toimintavarmuuden sekä ennustettavammat kustannukset verrattuna pilvipohjaiseen tekoälyyn.

Lokaalin tekoälyn käyttöönoton haasteita ovat tiedon jatkokäsittelyyn liittyvät riskit, kuten tietoturvariskit, jotka voivat syntyä käyttäjän toimesta tallentaessaan dataa epäluotettaviin paikkoihin. Tämä ei kuitenkaan ole suoranaisesti lokaalin tekoälyn tuottama ongelma vaan yleinen riski esimerkiksi salaisen datan käsittelyssä. Eräs IT-asiantuntija mainitsi, että vastaava tilanne voi syntyä, vaikka tekoälyjärjestelmä olisi sinänsä turvallinen. Riskejä voi esiintyä, jos henkilöstö käsittelee salaista tietoa varomattomasti, esimerkiksi lähettäessään sitä sähköpostitse. Tällaiset riskit voidaan kuitenkin minimoida tehokkaan tiedonhallinnan ja tarkasti määriteltyjen tietoturvakäytäntöjen avulla, jolloin henkilöstölle on selkeät ohjeet salaisen datan säilyttämisestä ja käsittelystä

4.4.2 Turvallinen päätöksenteko tekoälyhankinnoissa

On tärkeää tunnustaa, että tekoälyn käyttöönotto vaatii aina perusteellista tietoturvan tarkastelua. Vaikka tämän opinnäytetyön tulokset ja Aalto-yliopiston käytäntö Azure OpenAI:n suhteen tarjoavat lisää luottamusta pilvipohjaiseen tekoälyyn, on jokaisen organisaation itse arvioitava riskejä huolellisesti ennen uusien teknologioiden hankintaa. Tähän arviointiin voi kuulua esimerkiksi tutkimuksessa käsitelty Data Protection Impact Assessment (DPIA), joka on yksi menetelmä vaikutusten arvioimiseen.

Regulaatioihin ja lakeihin liittyen Aalto-yliopistossa on onnistuneesti integroitu säädöksiä koulun päätöksentekoprosesseihin, joilla halutaan lisätä kokonaisturvallisuutta päätöksenteossa. Tämä on esimerkillinen käytäntö, josta muutkin toimijat voivat ottaa mallia ja implementoida vastaavia toimintatapoja omassa toiminnassaan, edistääkseen turvallista päätöksentekoa hankintojen kontekstissa.

4.4.3 Salaiseksi luokitellun datan käyttäminen tekoälyssä

Salaiseksi luokitellun datan käyttö tekoälyssä edellyttää tarkkaa suunnittelua, joka kattaa teknologiset, tietoturva- ja eettiset kysymykset. Asiantuntijat tunnistivat pilvipohjaisissa tekoälyjärjestelmissä olevan riskejä, sillä datan hallinta siirtyy yrityksen ulkopuolelle, joka on selvä riski salaisen datan käsittelyssä. Vaikka pilvipalvelut tarjoavat kattavia tietoturvaominaisuuksia, saattaa silti olla turvallisempaa säilyttää arkaluonteinen data yrityksen omissa järjestelmissä.

Erityisesti terveysalalla toimivat julkiset organisaatiot ovat tiukkojen säädösten alaisuudessa terveystietojen käsittelyssä. Tällaisen datan käyttö pilvessä vaatii asiantuntijoiden mukaan raskaan suunnitteluprosessin, sisältäen tietojenkäsittelysopimuksen (DPA), tietosuojan vaikutusten arvioinnit (DPIA) sekä oikeudellisten asiantuntijoiden konsultaation varmistaakseen auditoinnin ja sopimusten asianmukaisuuden. Vaikka tämä prosessi olisikin tehtävissä, voivat nämä yritykset edelleen kohdata asiakkaidensa epäröintiä salaisen datan käytön suhteen pilvessä. Tämän vuoksi lokaalit tekoälyratkaisut saattavat olla houkuttelevampi vaihtoehto, sillä ne vähentävät monimutkaisten prosessien tarvetta ja mahdollistaa paremman datan hallinnan. Yhteenvetona voidaan todeta, että salaiseksi luokitellun datan käyttö tekoälyssä edellyttää organisaatioilta syvällistä pohdintaa, teknologista asiantuntemusta sekä kykyä navigoida monimutkaisten lainsäädännöllisten vaatimusten viidakossa.

4.5 Jatkotutkimusehdotus kustannuslaskennasta

Jatkotutkimus voisi keskittyä tutkimaan, millaisissa tilanteissa paikallinen tekoälyratkaisu on kustannustehokkaampi kuin pilvipohjainen palvelu. Tämä tutkimus voisi vertailla kattavasti kustannuksia eri käyttötapauksissa, ottaen huomioon muun muassa alkuperäiset investoinnit, jatkuvat ylläpitokulut ja järjestelmän skaalautuvuuden. Yksi haastatelluista asiantuntijoista toi esille, että suuren datamäärän käsittely on erityisen hyvä esimerkki tilanteesta, jossa paikallisesti isännöity tekoäly voi tarjota merkittäviä säästöjä. Hän mainitsi, että vaikka lokaali tekoäly vaatii alkuinvestointina kymmeniä tuhansia euroja, voisi intensiivisessä ja laajamittaisessa käytössä olevan suuryrityksen näkökulmasta tämä investointi osoittautua pitkässä juoksussa edullisemmaksi.

Kustannustutkimus voisi lisäksi syventyä analysoimaan erilaisia tekijöitä, kuten organisaation kokoa, erityistarpeita, tietoturvavaatimuksia ja lainsäädännön asettamia ehtoja, jotka vaikuttavat päätöksentekoon lokaalin ja pilvipohjaisen tekoälyratkaisun välillä. Tämän tutkimuksen avulla organisaatiot voivat tehdä tietoon perustuvia, taloudellisesti järkeviä päätöksiä tekoälyinvestoinneissaan, valitsemalla optimaalisimman tekoälypalvelun niiden yksilöllisiin tarpeisiin.

Lähteet

Aalto. 2020. Tiedon luokittelu. Viitattu 10.4.2024. Saatavissa rajoitetusti https://www.aalto.fi/fi/tiedon-kasittely/tiedon-luokittelu

Aalto. 2023. How to use Chat GPT. Viitattu 10.1.2024. Saatavissa rajoitetusti https://www.aalto.fi/system/files/2023-05/How%20to%20use%20ChatGPT.pdf

Aalto GPT. 2024. Aalto GPT your personal AI Assistant. Viitattu 1.2.2024. Saatavissa rajoitetusti https://gpt.aalto.fi/

Advian. What is Edge AI. Viitattu 18.1.2024. Saatavissa https://www.advian.fi/en/what-is-edge-ai

Benifei, B & Tudorache D. 2024. Euroopan parlamentti. Artificial intelligence act. Viitattu 26.1.2024. Saatavissa https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

Boina, R., & Achanta, A. 2022. Balancing Language Brilliance with User Privacy: A Call for Ethical Data Handling in ChatGPT. Viitattu 6.4.2024. Saatavissa https://www.ijsr.net/archive/v12i9/SR23903065711.pdf

Bolwell, A. 2023. HP Megatrends. Top 10 tech trends to watch in 2023. Viitattu 31.3.2024. Saatavissa https://hpmegatrends.com/top-10-tech-trends-to-watch-in-2023-53578d128b25

Brenner, M. 2023. Nutanix. The role of AI in cloud computing. Viitattu 14.1.2024. Saatavissa https://www.nutanix.com/theforecastbynutanix/technology/ai-in-the-cloud

Browne, K., Farley, P., Urban, E., Jenks, A. 2024. Microsoft. What is Azure OpenAI Service. Viitattu 1.2.2024. Saatavissa https://learn.microsoft.com/en-us/azure/ai-services/openai/overview

Davis, C., Edwards, E. Fox, C., Jupudi, A., Baumgarher, P, Mazz R., Johnsson M., Pablo, A., Cole L. 2024. Microsoft. Yksinkertaistettu GDPR: Pienyrityksesi opas. Viitattu 14.1.2024. Saatavissa https://learn.microsoft.com/fi-fi/microsoft-365/admin/security-and-compliance/gdpr-compliance?view=o365-worldwide

De Rose, A. 2023. HR Brew. These companies has banned ChatGPT. Viitattu 27.2.2024. Saatavissa https://www.hr-brew.com/stories/2023/05/11/these-companies-have-banned-chatgpt-in-the-office

Drapkin, A. 2023. Does ChatGPT Save My Data? OpenAI's Privacy Policy Explained. Viitattu 29.3.2024. Saatavilla https://tech.co/news/does-chatgpt-save-my-data

Euroopan parlamentti. 2020. Mitä tekoäly on ja mihin sitä käytetään. Viitattu 27.2.2024. Saatavissa https://www.europarl.europa.eu/news/fi/headlines/society/20200827STO85804/mita-tekoaly-on-ja-mihin-sita-kaytetaan

Euroopan parlamentti. 2022. Yleinen tietosuoja-asetus. Viitattu 25.1.2024. Saatavissa https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_fi.htm

Euroopan parlamentti. 2023. EU AI Act: first regulation on artificial intelligence. Viitattu 25.3.2024. Saatavissa https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Euroopan Parlamentin ja neuvoston asetus (EU) 2016/679, §4

Farley, P., Urban, E., Mehrota, N., Hill, A. 2023. Data, privacy, and security for Azure OpenAI Service. Viitattu 20.2.2024. Saatavissa https://learn.microsoft.com/en-us/legal/cognitive-services/openai/data-privacy

Google. What is Cloud Storage. Viitattu 13.1.2024. Saatavissa https://cloud.google.com/learn/what-is-cloud-storage

Heikkilä, M. 2023. Tekoäly: Koneoppiminen, neuroverkot, ChatGPT ja ChatGPT-4. Spotify. Viitattu 13.2.2024. Saatavissa https://open.spotify.com/episode/4f9Sz9VLUQVv1hkVtKNMQ9?si=V6OaQZ7uSRCQlQllGHPnbQ

Hu, K. 2023. Reuters. ChatGPT sets record for fastest-growing user base - analyst note. Viitattu 31.3.2024. Saatavilla https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

ISO/IEC 27001:2022 Standardi. Saatavissa https://www.iso.org/standard/27001

Johri, S. 2023. Harvard University. The Making of ChatGPT: From Data to Dialogue. Viitattu 12.2.2024. Saatavissa https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/

Kallum, S. 2023. BBC. ChatGPT banned in Italy over privacy concerns. Viitattu 26.1.2024. Saatavilla https://www.bbc.com/news/technology-65139406

Kurimo, R. 2023. Tekoäly: Koneoppiminen, neuroverkot, ChatGPT ja ChatGPT-4. Spotify Viitattu 13.2.2024. Saatavissa https://open.spotify.com/episode/4f9Sz9VLUQVv1hkVtKNMQ9?si=V6OaQZ7uSRCQlQllGHPnbQ

Lanfear, T. 2022. Microsoft. Azure Data Encryption at rest. Viitattu 13.3.2024. Saatavissa https://learn.microsoft.com/en-us/azure/security/fundamentals/encryption-atrest

Latterner, T. Medium. LLM & GPT: what are they, and how do they work. Viitattu 14.2.2024. Saatavissa https://medium.com/@thomas.latterner/llm-gpt-what-are-they-and-how-do-they-work-2df1b5925f6

Lawlor, P. 2023. Qualcomm. Getting personal with on-device-ai. Viitattu 10.1.2024. Saatavissa https://www.qualcomm.com/news/onq/2023/10/getting-personal-with-on-device-ai

Mazzoli., R. O´Sullivan, S., Berhnard, J. 2023. Microsoft. Azure and Dynamics 365 accountability readiness checklist for the GDPR. Viitattu 21.2.2024. Saatavissa https://learn.microsoft.com/fi-fi/compliance/regulatory/gdpr-arc-azure-dynamics

Microsoft. 2024a. Microsoftin tietosuojalauseke. Viitattu 24.2.2024. Saatavissa https://privacy.microsoft.com/fi-fi/privacystatement

Microsoft. 2024b. Tietojenkäsittelysopimus. Viitattu 14.2.2024. Saatavissa https://www.microsoft.com/licensing/docs/view/Microsoft-Products-and-Services-Data-Protection-Addendum-DPA

Microsoft. 2024c. Azure OpenAI Service Models. Viitattu 22.2.2024. Saatavissa https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/models#model-summary-table-and-region-availability

Norris, J. 2023. Microsoft. Is the Azure Open API in compliance with GDPR regulations. Viitattu 21.2.2024. Saatavissa https://learn.microsoft.com/en-us/answers/questions/1339718/is-the-azure-open-api-in-compliance-with-gdpr-regu

OpenAI. a. About. Viitattu 20.2.2024. Saatavissa https://openai.com/about

OpenAI. b. Pricing. Viitattu 14.2.2024. Saatavissa https://openai.com/chatgpt/pricing

OpenAI. c. Viitattu 12.2.2024. Saatavissa https://images.openai.com/blob/cf717bdb-0c8c-428a-b82b-3c3add87a600/ChatGPT_Diagram.svg?width=10&height=10&quality=50

Ouyang, L. Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin P., Zhang, C. 2024. OpenAI. Training language models to follow instructions with human feedback. Teoksessa:

Viitattu 12.2.2024. Saatavissa https://cdn.openai.com/papers/Training_language_models_to_follow_instructions_with_human_feedback.pdf

Porter, J. 2023. TheVerge. ChatGPT continues to be one of the fastest-growing services ever. Viitattu 12.2.2024. Saatavissa: https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference

Puget Systems. Hardware Recommendations for Machine Learning / AI. Viitattu 11.2.2024. Saatavissa https://www.pugetsystems.com/solutions/scientific-computing-workstations/machine-learning-ai/hardware-recommendations/

Salo, I. 2023.Luova tekoäly mullistaa kaiken – ChatGPT näyttää tietä. Viro: Printon. Saatavissa https://kauppakamaritieto-fi.ezproxy.saimia.fi/ammattikirjasto/teos/luova-tekoaly-muuttaa-kaiken-2023#kohta:3.20Miten20ChatGPT20toimii

SFS-EN ISO/IEC 27000:2020. Saatavissa https://sfs.fi/standardeista/tutustu-standardeihin/suositut-standardit/iso-iec-27000-tietoturvallisuuden-standardisarja/ Helsinki: SFS Suomen Standardit

Singh R, & Gill, S. Keai. 2023. Internet of Things and Cyber-Physical Systems. (72) Viitattu 6.4.2024. Saatavissa https://doi.org/10.1016/j.iotcps.2023.02.004

Stagnitto, J. 2024. Cloudwards. What Is Cloud Storage and How Does It Work? 2024 Guide to Online Data Storage. Viitattu 20.1.2024. Saatavissa https://www.cloudwards.net/how-cloud-storage-works/

Tienari, M. 2023. Tivi. Microsoft pääsee OpenAI:n hallitukseen. Viitattu 10.2.2024. Saatavissa https://www.tivi.fi/uutiset/microsoft-paasee-openain-hallitukseen/9a795d0d-830d-438e-96cf-f5f5e48bcca9

Tietosuojavaltuutetun toimisto. Vaikutustenarviointi. Viitattu 12.1.2024. Saatavissa https://tietosuoja.fi/vaikutustenarviointi

Urban, E., Downs, J., Huff, A. 2023. Microsoft. Azure OpenAI Service encryption of data at rest. Viitattu 21.2.2024. Saatavissa https://learn.microsoft.com/en-us/azure/ai-services/openai/encrypt-data-at-rest

Valava, P, 2023. Microsoft. Is there a difference between OpenAI ChatGPT Enterprise and ChatGPT version provided with Azure OpenAI. Viitattu 21.2.2024. Saatavissa https://learn.microsoft.com/en-us/answers/questions/1396203/is-there-a-difference-between-openai-chatgpt-enter

Vidich, S. 2023. Microsoft. ISO/IEC 27001:2022. Viitattu 30.1.2024. Saatavissa https://learn.microsoft.com/en-us/azure/compliance/offerings/offering-iso-27001

Wolford, B. 2024. Euroopan Unioni. Data Protection Impact Assessment (DPIA). Viitattu 14.1.2024. Saatavissa https://gdpr.eu/data-protection-impact-assessment-template/

Yuen, C. & Sanghavi, N. 2024. Azure. Microsoft Azure Compliance Offering. (9) Viitattu 22.2.2024. Saatavissa https://servicetrust.microsoft.com/DocumentPage/7adf2d9e-d7b5-4e71-bad8-713e6a183cf3

Liitteet

Liite 1. Kyselylomake

Yleisesti tekoäly

- Mielipide tekoälystä? Mitä hyvää mitä huonoa? Muuttaako se maailman?
- Miten data siirtyy Azure OpenAI:hin?

Azure OpenAI pilvitekoäly

- Luotetaanko Aallossa Microsoftiin?
- Onko Azure OpenAI tietoturvallinen?
- Azure OpenAI kierrättää dataa filtterin läpi, abuse monitoring on pois päältä, Data sijaitsee Azuren pilvessä, jossa myös kielimalli sijaitsee esim. ruotsissa. Mikä tässä datan kierrossa Azuren kautta on isoin huolenaihe?
- Miten Azure OpenAI:ssa on jaoteltu eri käyttäjä yritysten data?
- Mitä riskejä Azuressa on? Miten Azuren riskit voidaan minimoida?
- Jos nyt jo esim. Word tiedostoissa on luokiteltua dataa, joka on tallennettu Azuren pilveen, niin miksi Azure OpenAI:hin ei voida syöttää samaa dataa jos tekniikka pysyy samana?
- Miksi useissa organisaatioissa tekoälyä saa hyödyntää vain julkisen datan kanssa?

Lokaali tekoäly

- Tiedätkö mikä on pilvitekoäly ja mikä on lokaali tekoäly?
- Lokaali kielimalli ladataan käyttäjän kovalevylle, haut tehdään sieltä ilman nettiä eli data ei kierrä ulos laitteesta. Mielipide?
- Jos käyttäisimme lokaalia tekoälyä, tehdessämme luottamuksellisen tiedon kanssatöitä, mitä hyötyä ja haasteita siitä olisi?
- Onko lokaali tekoäly yksi ratkaisu chattibottien tietosuoja ongelmille?
- Jos pilvitekoälyä ei saa käyttää Secret datan kanssa, voisiko sitä käyttää lokaalissa tekoälyssä?
- Mitä riskejä on lokaalissa mallissa, miten riskejä voidaan minimoida?
- Miten Aalto IT:ssä päästään hyödyntämään tehokkaasti lokaalia tekoälyä?
- Voiko tehokkaan chatbotin saada jotenkin Esupportiin helpottamaan Service Deskin työtaakkaa?
- Jos yksilö suunnittelee pommin aallon tekoälyä hyväksikäyttäen, niin mikä on Aallon vastuu juridisesti, kun AI ACT kieltää laittomat tekoälyt?
- Kenellä on loppupeleissä vastuu, jos pilvitekoälystä tietoja vuotaa johonkin, Microsoftin, Aallon vai Yksilön?
- Mitä käyttäjän tulee ymmärtää tietoturvallisesta Azure OpenAI:n käytöstä