Jan Sippola

TEKOÄLYN MAHDOLLISUUDET KYBERUHKIEN JATKUVAN VALVONNAN JA TORJUNNAN TYÖKALUNA

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2023

TIIVISTELMÄ

Sippola, Jan
Tekoälyn mahdollisuudet kyberuhkien jatkuvan valvonnan ja torjunnan työkaluna
Jyväskylä: Jyväskylän yliopisto, 2023, 39 s.
Tietojärjestelmätiede, Kandidaatintutkielma
Ohjaaja: Vuorinen, Jukka

Nopea teknologian kehitys on saanut ihmiset ja yhteiskunnat riippuvaisemmiksi entistä monimutkaisemmista järjestelmistä ja yhä useampi yritys ja organisaatio käyttääkin laajasti tietotekniikkaa eri toiminnoissaan. Tämä kehitys on tuonut mukanaan lisääntyneet, entistä monimutkaisemmat kyberturvallisuutta koskevat uhat, joihin vastaamiseen tekoälystä on haettu apua. Tämän kandidaatintutkielman tarkoituksena oli selvittää tekoälyn käyttömahdollisuuksia kyberturvallisuuden parantamiseksi erityisesti jatkuvan valvonnan alueella. Tutkielma toteutettiin joidenkin menetelmien osalta systemaattista katsausta lähestyvänä kuvailevana kirjallisuuskatsauksena, jossa analysoitiin monipuolisesti tutkimuksia ja artikkeleita tekoälystä, kyberturvallisuudesta, jatkuvasta valvonnasta ja tekoälyn käytöstä kyberturvallisuudessa. Tekoäly ja kyberturvallisuus ovat käsitteinä laajoja, joten niiden määritelmiä ja niihin liittyviä seikkoja käsiteltiin tutkielmassa aiheen kannalta olennaisin osin. Tutkielman tulokset osoittivat, että tekoälyllä on suuri potentiaali erityisesti kyberuhkien jatkuvassa valvonnassa ja uhkien torjunnassa, vaikkakin sen käyttöön liittyy myös haasteita. Kyberympäristön jatkuvassa valvonnassa tekoäly osoittautui erittäin hyödylliseksi, sillä sen avulla voidaan havaita tehokkaasti esimerkiksi poikkeamia normaalista toiminnasta ja ennaltaehkäistä mahdollisia tulevia hyökkäyksiä. Erilaisten tekoälysovellusten avulla voidaan vastata kehittyneisiin ja monipuolistuviin kyberuhkiin tehokkaammin kuin perinteisillä menetelmillä. Tehokkaammat keinot kyberuhkien torjunnassa ovatkin tärkeitä, sillä teknologian kehityksen myötä myös kyberuhat ovat entistä kehittyneempiä ja monimutkaisempia. Tekoäly näyttääkin olevan välttämätön apu tulevaisuuden kyberuhkien ennaltaehkäisemisessä ja torjumisessa, ja organisaatioilla ja yrityksillä onkin jo nykyään käytössä erilaisia tekoälysovelluksia kyberympäristön jatkuvassa valvonnassa.

Asiasanat: tekoäly, kyberturvallisuus, AI-algoritmi, kyberuhat, jatkuva valvonta, haittaohjelma, teollisuus

ABSTRACT

Sippola, Jan
The possibilities of artificial intelligence as a tool for continuous monitoring and mitigation of cyber threats
Jyväskylä: University of Jyväskylä, 2023, 39 pp.
Information Systems, Bachelor's Thesis
Supervisor: Vuorinen, Jukka

The rapid development of technology has made people and societies increasingly dependent on more complex systems, and an increasing number of companies and organizations widely use information technology in their operations. This development has brought about increased, more complex cybersecurity threats, for which artificial intelligence (AI) has been sought as a solution. The purpose of this bachelor's thesis was to explore the potential use of AI in improving cybersecurity, particularly in the area of continuous monitoring. The thesis was conducted partly as a systematic review, analyzing a diverse range of research papers and articles on AI, cybersecurity, continuous monitoring, and the use of AI in cybersecurity. AI and cybersecurity are broad concepts, so the thesis addressed their definitions and relevant aspects for the topic. The results of the thesis demonstrated that AI has great potential, especially in continuous monitoring and mitigating cyber threats, although its use also presents challenges. In the context of continuous monitoring in the cyber environment, AI proved to be highly beneficial, as it can effectively detect anomalies in normal operations and prevent potential future attacks. Various AI applications can respond to advanced and diversified cyber threats more effectively than traditional methods. More effective means of countering cyber threats are crucial, as technological advancements have led to increasingly sophisticated and complex cyber risks. AI appears to be an essential tool in preventing and combating future cyber threats, and organizations and companies already utilize various AI applications for continuous monitoring in the cyber environment.

Keywords: artificial intelligence, cybersecurity, AI algorithm, cyber threats, continuous monitoring, malware, industry

TAULUKOT

SISÄLLYS

1 JOHDANTO

Erilaisten teknologioiden nopea kehitys on mullistanut maailmaamme ja tehnyt yhteiskunnista ja ihmisistä riippuvaisia yhä monimutkaisemmista järjestelmistä. Esimerkiksi yhä useampi yritys ja organisaatio käyttää laajasti tietotekniikkaa eri toiminnoissaan (Culot ym., 2019). Tämä kehitys on kuitenkin tuonut mukanaan myös uusia haasteita, kuten uudenlaisia kyberturvallisuusuhkia, jotka ovat yhä monimutkaisempia ja vaikeampia havaita ja torjua (August ym., 2022). Viime aikoina kyberturvallisuudesta on Suomessakin tullut mielenkiintoinen ja näkyvä aihe, sillä se on noussut julkiseen keskusteluun yhä useammin muun muassa uutisten muodossa. Suomessa on uutisoitu näkyvästi vakavista tietovuodoista ja teollisuuden alalla tapahtuneista kyberhyökkäyksistä, joilla on ollut kalliitakin seurauksia (*Suomeen tehdään nyt aiempaa vaarallisempia kyberhyökkäyksiä*, 2022; *Kyberturvallisuuskeskus*, 2023; *Lunnaiden maksu rikollisille voi olla välttämätöntä tietomurroissa*, 2022).

Kyberturvallisuus onkin tullut yhä tärkeämmäksi alueeksi digitalisoituvassa maailmassa, jossa muun muassa Mattilan ym. (2020) mukaan tietoturvaongelmat ovat jokapäiväistyneet (Mattila ym., 2020). Kyberhyökkäykset voivatkin aiheuttaa merkittävää haittaa niin yksilöille, yrityksille, kuin yhteisöille laajemminkin, joten turvallisuuden parantaminen uusien teknologioiden avulla on tärkeää (*HS Järvenpää*, 2023; *Kyberhyökkäys pysäytti tuotannon Valtran tehtaalla*, 2022; *Vastaamon tietovuoto*, 2023). Uudenlaisiin monimutkaistuviin kyberuhkiin vastauksena on tekoälyä alettu käyttää yhä enemmän kyberturvallisuuden parantamiseksi. Tekoälyn käyttö tällä alueella ei ole kuitenkaan ollut täysin mutkatonta, ja sen käyttöön on huomattu liittyvän joitakin haasteita ja rajoitteita.

Tämä kandidaatintutkielma käsitteleekin tekoälyn käyttöä kyberturvallisuudessa ja erityisesti sen soveltamista jatkuvassa valvonnassa. Aihe on tieteellisessä tutkimuksessa suhteellisen tuore ja alaa leimaa jatkuva muutos, joten lisätutkimukselle tästä näkökulmasta nähdään olevan tarvetta (Ansari ym., 2022). Tutkielman tavoitteena on osaltaan tuoda esiin tekoälyn roolia kyberturvallisuudessa, sekä pohtia hieman sen käyttöön liittyviä mahdollisuuksia ja rajoituksia. Lisäksi mielenkiinnon kohteena tutkielmassa on hieman käsitellä tekoälyn käyttöönoton vaikutuksia ihmisen roolin kehittymiseen kyberturvallisuuden toteuttamisen kokonaisuudessa.

Tutkielma sisältää katsauksen myös tekoälyn määritelmään, sen käyttöön kyberturvallisuudessa, sekä erityisesti jatkuvan valvonnan merkitykseen alalla. Tekoälyn ajatellaan olevan karkeasti monimutkaisia algoritmeja ja ohjelmistoja, jotka jäljittelevät ihmiselle ominaisia kykyjä, kuten oppimista ja päättelyä. Kyberturvallisuus puolestaan ottaa määritelmässään huomioon perinteisen tietoturvallisuuden lisäksi ihmisen osana turvallisuuden kokonaisuutta.

Tekoälyn käytölle kyberturvallisuuden parantamiseksi jatkuvan valvonnan saralla nähdäänkin tämän tutkielman pohjalta suurta potentiaalia, sillä erilaiset tekoälysovellukset ovat osoittautuneet tehokkaiksi kyberuhkien havainnoinnissa ja torjumisessa verrattuna perinteisiin menetelmiin. Toisaalta tässä tutkielmassa nousee esiin tekoälyn käyttöön liittyviä haasteita, kuten luotettavuuteen ja ihmisten roolin kehitykseen liittyvät haasteet.

Tutkielmalle asetettiin lisäksi kaksi tutkimuskysymystä, joihin katsauksessa pyrittiin etsimään vastauksia:

- Miten tekoälyä voidaan hyödyntää kyberturvallisuuden parantamisessa jatkuvassa valvonnassa?
- Miten eri tekoälysovellukset vertautuvat toisiinsa kyberuhkien torjunnassa?

Tutkielma on toteutettu osittain kuvailevana kirjallisuuskatsauksena, joka lähenee joidenkin menetelmien osalta systemaattista katsausta. Lähdeaineistoa onkin pyritty keräämään ja hakemaan mahdollisimman erilaisista lähteistä ja julkaisukanavista kattavamman kokonaiskuvan ja monipuolisempien näkökulmien aikaansaamiseksi. Pääasiallisesti aineistoa on haettu esimerkiksi IEEE Xplore ja Jyväskylän yliopiston kirjaston JYKDOK tietokannoista, sekä Google Scholar-hakukoneesta, mutta myös eri organisaatioiden ja uutistoimistojen verkkosivustoja on hyödynnetty hakuprosessin aikana. Näiden lisäksi lähteenä on käytetty joitakin alan oppikirjoja ja muita relevanteiksi katsottuja julkaisuja, kuten muilta tutkielman aiheisiin perehtyneiltä saatuja tutkimuksia ja artikkeleita, sekä luentoja.

Aineiston hakemiseen tietokannoista ja hakukoneista on käytetty englanninkielisiä aiheeseen liittyviä hakusanoja, kuten *artificial intelligence, AI, cyber threats, malware, AI algorithm* ja *cyber security*, sekä niiden yhdistelmiä. Hakutuloksia tuli tällä menetelmällä satoja, joten tuloksia rajattiin lisäksi myös julkaisuvuoden ja tyypin mukaan. Näillä keinoilla hakutulokset rajautuivat noin 200 artikkeliin kussakin tietokannassa ja hakukoneessa. Hakutuloksista valittiin tiivistelmien perusteella sopivimmat 150 artikkelia tarkempaan analyysiin, jossa tutustuttiin kattavammin niiden sisältöön ja taustoihin. Tämän pohjalta tutkielman kannalta epärelevantti aineisto rajattiin pois. Epärelevanteiksi tuloksiksi katsottiin sellaiset raportit ja artikkelit, jotka eivät suurelta osin liittyneet tekoälyn hyödyntämiseen kyberturvallisuudessa tai olivat suurelta osin samankaltaisia jo lähdeaineistoon valittujen artikkelien kanssa. Hakutuloksien rajaamisessa priorisoitiin hieman myös artikkelien julkaisukanavien JUFO-portaalin tasoluokituksia, mutta artikkeleita on valittu myös luokittelemattomista tai matalammalle luokitelluista kanavista tarkoituksellisesti. Wiafen ym. (2020) mukaan alan tutkimusta on julkaistu pääosin suppeasti vain parissa alan julkaisussa, joten kattavamman

näkökulman vuoksi lähdeaineistoa on pyritty etsimään laajemmin. Tekoälyteknologioiden nopean kehityksen vuoksi lähdeaineistoon on lisäksi pyritty valitsemaan pääosin alle viisi vuotta vanhaa kirjallisuutta, mutta joidenkin teorioiden ja aiheiden kohdalla tätäkin vanhempi aineisto on katsottu asianmukaiseksi ja relevantiksi. Kokonaisuudessaan tässä tutkielmassa käytettyyn lähdeaineistoon näiden toimenpiteiden jälkeen valikoitui hieman alle 70 lähdettä.

Kyberturvallisuus ja tekoäly ovat molemmat ajankohtaisia ja laajoja aihealueita, joista on julkaistu jonkin verran tutkimusta ja artikkeleita eri tieteenaloilla. Tämä tutkielma erottuu muusta aiheeseen liittyvästä kirjallisuudesta kirjallisuuskatsauksena, joka käsittelee tekoälyn käyttömahdollisuuksia kyberturvallisuuden parantamiseksi erityisesti jatkuvan valvonnan alueella. Niinpä tämä tutkielma voi tuoda uusia ja hyödyllisiä näkökulmia tekoälyn soveltamisesta kyberturvallisuudessa kaikille aiheesta kiinnostuneille, kuten opiskelijoille, alalle pyrkiville tai oman organisaation kyberympäristön kehittämisestä kiinnostuneelle.

Tutkielma koostuu viidestä osasta. Johdannon jälkeen toisessa luvussa käsitellään tekoälyn määritelmää, sekä siihen liittyviä menetelmiä aiheen kannalta olennaisin osin. Kolmannessa luvussa käsitellään puolestaan kyberturvallisuuden määritelmää, käsitystä turvan toteutumisesta, kyberuhkia ja niiden torjuntaa, jatkuvaa valvontaa ja sen merkitystä kyberturvallisuudessa, sekä hieman teollisuussektorin erityispiirteitä kyberturvallisuuteen liittyen. Neljännessä luvussa nämä kaksi aluetta tuodaan yhteen ja tarkastellaan tekoälyn käyttöä kyberturvallisuudessa ja siihen liittyviä hyötyjä ja rajoitteita. Neljännessä luvussa paneudutaan myös tekoälyn käyttöön jatkuvassa valvonnassa tapausesimerkkien kautta ja käsitellään hieman tekoälyn käyttöönoton vaikutuksia kyberturvallisuusammattilaisten rooliin. Viimeisessä yhteenvetoluvussa pyritään luomaan synteesiä löydösten pohjalta, kuvailemaan tähän tutkielmaan liittyviä rajoitteita, sekä esittämään aiheen kannalta relevantteja jatkotutkimusaiheita.

2 TEKOÄLY

Tässä luvussa pyritään aiempaan tutkimukseen ja kirjallisuuteen pohjautuen määrittelemään tekoälyn käsite. Lisäksi tässä luvussa perehdytään kyberturvallisuuteen olennaisesti liittyviin tekoälyn menetelmiin.

2.1 Tekoälyn määritelmä

Tekoäly (artificial intelligence, AI) on laaja käsite, jolle tieteellisessä kirjallisuudessa ei näytä olevan täysin yksiselitteistä määritelmää. Tekoäly voidaan ymmärtää lähestymistavasta riippuen hyvin monella eri tavalla, kuten Russell ja Norvig (2010) kirjassaan *Artificial Intelligence a Modern Approach* toteavat (Russell & Norvig, 2010). On kuitenkin nähtävissä, että tieteellisessä kirjallisuudessa pääosin vallitsee varsin pitävä konsensus siitä, mitä tekoälystä puhuttaessa tarkoitetaan.

Tekoälyn ajatellaan olevan erilaisten tietokonejärjestelmien ja ohjelmistojen kykyä suorittaa älykkäinä pidettyjä toimintoja, joita yleensä pidetään ihmisten kykyinä. Russellin ja Norvigin (2010) mukaan se on lisäksi eräänlainen tietojenkäsittelyn osa-alue, joka pyrkii kehittämään järjestelmiä ja algoritmeja, jotka kykenevät suorittamaan, ilman ihmisen suorittamaa ohjausta ja valvontaa, näitä älykkyyttä vaativia tehtäviä, kuten päätöksentekoa, oppimista ja ongelmanratkaisua (Russell & Norvig, 2010). Heidän määritelmänsä korostaa erityisesti tekoälyn kykyä suorittaa tehtäviä ilman ihmisen ohjausta ja valvontaa, mutta huomioi myös älykkyyden käsitteen. Tämän perusteella tekoälyn voidaankin ajatella pyrkivän jäljittelemään ihmisen älykkyyttä ja käyttämään sitä monimutkaisten tehtävien suorittamiseen.

Aihe nähdään tieteellisessä kirjallisuudessa kuitenkin hyvin laajana ja monimutkaisena, minkä vuoksi tekoälyn tarkka määritteleminen yksiselitteisesti voi olla haasteellista. Tekoälyä on pyritty aika ajoin määrittelemään paremmin ja kuvaavammin, joten erilaisia näkökulmia aiheeseen on olemassa. Esimerkiksi Goodfellow ym. (2016) korostavat määritelmässään tekoälyn kykyä oppia ja sopeutua uusiin tilanteisiin, kun taas esimerkiksi Zeadally ym. (2020), Pourjuavan

(2019) ja Pavithra ja Femilda Josephin (2020) painottavat sen kykyä käsitellä suuria määriä tietoa nopeasti ja tehokkaasti (Goodfellow ym., 2016; Pavithra & Femilda Josephin, 2020; Pourjavan, 2019; Zeadally ym., 2020).

Yksi selitys poikkeaville määritelmille ja lähestymistavoille voi olla se, että tekoäly voi perustua erilaisiin tekniikoihin ja menetelmiin, ja sen ajatellaankin koostuvan pienemmistä osa-alueista. Vähäkainun ja Neittaanmäen (2018) mukaan tekoäly-termiä voidaan pitää tietynlaisena ylätason käsitteenä tai niin sanottuna sateenvarjoterminä, joka pitää sisällään eri tekoälytekniikat ja -menetelmät, kuten koneoppimisen (machine learning, ML) ja syväoppimisen (deep learning, DL) (Vähäkainu & Neittaanmäki, 2018).

Tieteellisessä kirjallisuudessa tekoäly jaetaan usein myös sen kyvykkyyden ja ominaisuuksien perusteella kahteen päätyyppiin: heikko tekoäly ja vahva tekoäly. Fein ym. (2022) ja Goertzellin (2014) mukaan heikko tekoäly (narrow AI) on tietokonejärjestelmä, joka on suunniteltu suorittamaan vain tiettyjä rajoitettuja tehtäviä (Fei ym., 2022; Goertzel, 2014). Esimerkkejä heikosta tekoälystä ovat muun muassa kasvojentunnistus- ja puheentunnistusohjelmat. Vahva tekoäly (strong AI) heidän mukaansa tarkoittaa puolestaan tietokonejärjestelmää, joka kykenee suorittamaan älykkäitä toimintoja ihmisen tavoin (Fei ym., 2022; Goertzel, 2014). Vahvaa tekoälyä ei tämän tutkielman lähdeaineiston perusteella ole kuitenkaan vielä kehitetty, mutta sen kehittäminen näyttäisi olevan yksi tekoälytutkimuksen tärkeimmistä tavoitteista (Hamet & Tremblay, 2017; Russell & Norvig, 2010; Vähäkainu & Neittaanmäki, 2018; Wiafe ym., 2020).

Aiheen monitahoisuudesta ja laajuudesta huolimatta, tekoälyn voidaan lähdekirjallisuuden perusteella sanoa karkeasti olevan kehittyneitä ja monimutkaisia algoritmeja ja tietokoneohjelmistoja, jotka voivat oppia ja parantaa suorituskykyään kokemuksen perusteella. On kuitenkin tärkeää ottaa huomioon, että tekoäly kehittyy jatkuvasti, joten sen määritelmät voivat muuttua, sillä tekoälyn ominaisuuksien kehityksen myötä sen soveltamisalueetkin tullee laajenemaan.

2.2 Tekoälyn menetelmiä

Kuten edellä mainittiin, tekoäly koostuu erilaisista menetelmistä ja tekniikoista, joten tekoälystä puhuttaessa onkin olennaista hieman yksityiskohtaisemmin ymmärtää sen sisältämiä kokonaisuuksia ja olennaisimpia menetelmiä.

Tässä tutkielmassa käytetyssä lähdekirjallisuudessa esiintyy useita tekoälymenetelmiä ja niiden mahdollisia sovelluskohteita kyberturvallisuudessa. Näistä olennaisimmat on koottu taulukkoon (taulukko 1) helpottamaan eri menetelmien erojen ja kyberturvallisuuden sovelluskohteiden nopeaa hahmottamista. Tässä tutkielmassa menetelmä on katsottu olennaiseksi, mikäli se on mainittu useammassa kuin viidessä lähdeaineistona käytetyssä artikkelissa ja sille on osoitettu selkeä sovelluskohde kyberturvallisuudessa. Näillä perusteilla tärkeimmiksi voidaan katsoa edellä mainittujen koneoppimisen ja syväoppimisen lisäksi neuroverkot, tietokoneen näkö, luonnollisen kielen käsittely, sekä päättely ja suunnittelu. Näillä kaikilla on omia hieman toisistaan poikkeavia erityispiirteitä, joten

jokaisella niistä on myös yksilöllisiä sovelluskohteita muun muassa kyberturvallisuuden alueella.

TAULUKKO 1 Tekoälyn menetelmiä ja sovelluskohteita kyberturvallisuudessa

Tekoälyn menetelmä	**Kuvaus**	**Esimerkkisovellusalueet kyberturvallisuudessa**
Koneoppiminen	Koneoppiminen käyttää algoritmeja datan analysointiin ja oppimiseen	Haittaohjelmien tunnistaminen ja torjunta, haavoittuvuuksien ennustaminen (mm. Chellappan ym., 2018; Dua & Du, 2016; Wiafe ym., 2020)
Neuroverkot	Neuroverkot koostuvat tietokoneiden monimutkaisista yhteyksistä	Verkkoliikenteen analysointi, käyttäjätunnusten hallinta (mm. Fei ym., 2022; Goodfellow ym., 2016; Russell & Norvig, 2010)
Tietokoneen näkö	Tietokoneen näkö tarkoittaa tietokoneen kykyä tunnistaa kuvia ja niissä esiintyviä asioita	Tunnistaa mm. haitallisia kuvia ja videoita (mm. Goertzel, 2014; Nikiforakis ym., 2012; Russell & Norvig, 2010; Zeadally ym., 2020)
Luonnollisen kielen käsittely	Luonnollisen kielen käsittely auttaa tietokonetta ymmärtämään ja käsittelemään ihmisen kieltä	Huijausviestien havaitseminen ja torjunta (mm. Goertzel, 2014; Pourjavan, 2019; Vähäkainu & Neittaanmäki, 2018)
Syväoppiminen	Syväoppiminen käyttää syviä neuroverkkoja datan analysointiin ja siitä oppimiseen	Havaitsee uusia tietoturvariskejä ja haavoittuvuuksia (mm. Fei ym., 2022; Goodfellow ym., 2016; Pourjavan, 2019)
Päättely ja suunnittelu	Päättely ja suunnittelu käyttävät loogisia sääntöjä ja tietoa päätösten tekemiseen	Riskianalyysi, haittaohjelmien torjunta (mm. Goertzel, 2014; Nikiforakis ym., 2012; Pourjavan, 2019; Russell & Norvig, 2010; Zeadally ym., 2020)

Chellappan ym. (2018), Duan ja Dun (2016) ja Wiafen ym. (2020) mukaan koneoppiminen on yksi keskeisimmistä tekoälyn menetelmistä, joka mahdollistaa tietokonejärjestelmien oppimisen datasta ilman nimenomaista ohjelmointia. Koneoppimisen algoritmit tunnistavat heidän mukaansa malleja käytetystä datasta ja käyttävät näitä malleja uusien datanäytteiden ennustamiseen ja luokitteluun. Goodfellowin ym. (2016) ja Pourjuavanin (2019) mukaan syväoppiminen on vahvasti kytköksissä juuri koneoppimiseen, sillä syväoppiminen perustuu syvien neuroverkkojen käyttöön ja pyrkii hieman koneoppimisen tapaan oppimaan

automaattisesti monimutkaisia piirteitä ja abstraktioita suurista datamääristä useiden kerrostettujen neuroverkkojen avulla. Syväoppiminen onkin heidän mukaansa hyvin tehokas kuvien, äänen ja luonnollisen kielen käsittelyssä, joten sen lukuisten sovellusalojen joukkoon mahtuu muun muassa autonominen ajaminen. (Chellappan ym., 2018; Dua & Du, 2016; Goodfellow ym., 2016; Pourjavan, 2019; Wiafe ym., 2020.)

Syväoppimiseen liittyvät neuroverkot ovat myös erillinen tekoälyn menetelmä, joka pyrkii jäljittelemään ihmisen aivojen rakennetta ja toimintaa. Fein ym. (2022), Goodfellowin ym. (2016), sekä Russellin ja Norvigin (2010) mukaan neuroverkot koostuvat monista yksiköistä ja neuroneista, jotka ovat yhteydessä toisiinsa. Heidän mukaansa nämä yksiköt pystyvät suorittamaan monimutkaisia laskutoimituksia, joista on suurta hyötyä esimerkiksi koneoppimisen tapaan luokittelussa ja ennustamisessa. (Fei ym., 2022; Goodfellow ym., 2016; Russell & Norvig, 2010.)

Syväoppimiseen ja neuroverkkoihin liittyvä luonnollisen kielen käsittely puolestaan mahdollistaa tietokonejärjestelmien ymmärtää ihmisten käyttämää kieltä, sekä vastaavasti tuottaa luonnollista kieltä. Tämä on tärkeää esimerkiksi kielipohjaisissa hakukoneissa ja Chat-roboteissa, mutta siitä on hyötyä myös monella muulla alueella, kuten kyberturvallisuudessa roskapostisuodattimissa. (Goertzel, 2014; Pourjavan, 2019; Vähäkainu & Neittaanmäki, 2018.)

Tietokoneen näkö taas tarkoittaa sitä, että tietokonejärjestelmien on mahdollista tunnistaa kuvioita kuvista ja videoista. Tällainen kyky on tärkeää esimerkiksi kasvojentunnistuksessa tai automaattisessa ajoneuvonavigoinnissa. Päättely ja suunnittelu puolestaan mahdollistaa kyseisten järjestelmien tehdä nimensä mukaisesti päätöksiä ja suunnitella toimintaa esimerkiksi pelitekoälyissä tai muissa älykkäissä järjestelmissä. (Goertzel, 2014; Nikiforakis ym., 2012; Pourjavan, 2019; Russell & Norvig, 2010; Zeadally ym., 2020.)

Lähdekirjallisuuden perusteella nämä kaikki edellä mainitut ja taulukkoon (taulukko 1) nostetut tekniikat ja menetelmät ovatkin yleisesti käytössä tekoälyssä, joten sitä voidaan soveltaa ja käyttää kyberturvallisuuden alan lisäksi muun muassa Zeadallyn ym. (2020) ja Vähäkainun ja Neittaanmäen (2018) mukaan lukuisilla eri aloilla, kuten terveydenhuollossa, taloudessa, viihteessä ja koulutuksessa (Vähäkainu & Neittaanmäki, 2018; Zeadally ym., 2020). On kuitenkin hyvä tiedostaa, että tätä tutkielmaa varten koottu taulukko ei ole kaiken kattava, sillä tieteellisessä kirjallisuudessa esiintyy paljon muitakin tekoälymenetelmiä. Lisäksi on syytä noteerata joidenkin menetelmien päällekkäisyydet, kuten esimerkiksi syväoppimisen, koneoppimisen ja neuroverkkojen suhdetta tarkasteltaessa käy ilmi, minkä vuoksi tehty taulukko sopii lähinnä suurpiirteiseen tarkasteluun ja eri alueiden hahmottamiseen tämän tutkielman seuraavissa luvuissa käsiteltäviä seikkoja ajatellen.

3 KYBERTURVALLISUUS

Tässä luvussa pyritään aiempaan tutkimukseen ja kirjallisuuteen pohjautuen määrittelemään kyberturvallisuuden käsite, sekä kuvailemaan jatkuvan valvonnan merkitystä kyberturvallisuudessa. Lisäksi tarkastellaan hieman teollisuussektorin ominaispiirteitä kyberturvallisuuden näkökulmasta.

3.1 Kyberturvallisuuden määritelmä

Kyberturvallisuus (cyber security) on tekoälyn tapaan kohtuullisen laaja käsite, jonka tarkasta määrittelystä on olemassa muun muassa Craigen ym. (2014) mukaan monenlaisia hieman toisistaan poikkeavia näkemyksiä (Craigen ym., 2014). Tutkielmassa käytetyn aineiston perusteella kyberturvallisuudella tarkoitetaan kuitenkin pääasiassa toimintaa, joka kattaa kaikki ne toimenpiteet, joiden tarkoituksena on suojata tietotekniikkaan liittyviä järjestelmiä, verkkoja, laitteita, ohjelmistoja ja tietoja tietoturvauhilta, joita kohdistuu järjestelmiin yhden tai useamman mahdollisen hyökkäysvektorin kautta. Kyberturvallisuus nähdäänkin laajasti monitieteellisenä alana, joka yhdistää tietotekniikan, tietoturvan, tiedonhallinnan ja riskienhallinnan, sekä juridiset seikat. (Craigen ym., 2014; *Cybersecurity and Cyberwar*, ei pvm.; Ozkaya, 2019; Padallan, 2019; Zeadally ym., 2020.)

Fisherin (2016) mukaan kyberturvallisuuden käsitteen määrittelyssä on tiettyjä haasteita, sillä kyberturvallisuus sekoitetaan usein muiden aihetta sivuavien konseptien, kuten tietoturvallisuuden kanssa (Fischer, 2016). Myös Von Solmsin ja Van Niekerkin (2013) mukaan näillä kahdella käsitteellä on toki paljon yhteneväisyyksiä, mutta niillä on huomattaviakin eroja. Heidän mukaansa kyberturvallisuutta pidetään tietoturvallisuutta paljon laajempana käsitteenä, sillä se ottaa huomioon myös ihmisen mahdollisena uhkana turvallisuudelle (von Solms & van Niekerk, 2013). Arkisessa keskustelussa näillä ei kirjallisuuden perusteella nähdä kovin suurta merkitystä, mutta virallisessa keskustelussa tarkempien merkitysten kanssa on syytä olla tarkkana erityisesti silloin, kun kysymyksessä on lakiin, vastuukysymyksiin tai sopimuksiin liittyviä keskusteluita.

Ozkayan (2019), Padallan (2019), Zeadallyn (2020) ja monien muiden alan tutkijoiden mukaan kyberturvallisuudessa on kuitenkin pohjimmiltaan tavoitteena suojata tietotekniikkajärjestelmät, tietoverkot, laitteet, tiedot, organisaatiot ja yksittäiset käyttäjät erilaisilta tietoturvaan liittyviltä riskeiltä, kuten tahattomalta tai tarkoitukselliselta vahingoittamiselta, luvattomalta pääsyltä verkkoon tai järjestelmiin, urkinnalta, haittaohjelmilta, tietomurroilta, palvelunestohyökkäyksiltä, identiteettivarkauksilta ja tietojen kalastelulta. Heidän mukaansa kyberturvallisuuden avulla pyritään varmistamaan tietoturvallinen käyttö ja luotettava toiminta tietotekniikkaa hyödyntävissä järjestelmissä. (Craigen ym., 2014; Ozkaya, 2019; Padallan, 2019; von Solms & van Niekerk, 2013; Zeadally ym., 2020.)

Näin ollen kyberturvallisuus kattaa muun muassa Nikiforakiksen (2012) ja Wiafen ym. (2020) mukaan laajan kirjon erilaisia teknisiä, organisatorisia ja inhimillisiä toimenpiteitä, jotka auttavat suojaamaan tietojärjestelmiä ja tietoverkkoja. Näitä toimenpiteitä ovat esimerkiksi tietoturvasovellusten käyttö, vahvojen salasanojen käyttö, käyttäjien koulutus, riskien arviointi ja tietoturvakäytäntöjen noudattaminen. (Hamet & Tremblay, 2017; Nikiforakis ym., 2012; Wiafe ym., 2020.)

3.1.1 Turvan toteutuminen

Kyberturvallisuudesta puhuttaessa on tietenkin kiinnostavaa pohtia sitä, että milloin jokin asia koetaan olevan turvassa tai mitä turvallisuus itseasiassa on. Tieteellisessä kirjallisuudessa kyberturvallisuudesta ja tietoturvallisuudesta puhuttaessa usein keskustellaan tiedon luottamuksellisuudesta, eheydestä ja saatavuudesta (confidentiality, integrity, availability, CIA). Kirjallisuudessa näytetään ajattelevan niin, että kun tieto on luottamuksellista, eheää ja helposti saatavilla, eli CIA toteutuu, niin tietoturva toteutuu. (Craigen ym., 2014; Lezzi ym., 2018; Vuorinen & Tetri, 2012.)

Vuorinen ja Tetri (2016) sivuavat tätä aihetta artikkelissaan *Paradoxes in information security*, jonka mukaan luottamuksellisuus ja eheys perustuu pitkälti ideaan siitä, että vain valtuutetulla käyttäjällä pitäisi olla pääsy tietoon ja sen muokkaamiseen. Eheys liittyy myös tiedon muuttumattomuuteen, eli tiedot eivät saisi muuttua, ellei valtuutettu käyttäjä itse niitä muuta. Saatavuus puolestaan viittaa saavutettavuuteen, eli tiedon tulisi olla saatavilla ja muokattavissa silloin, kun valtuutettu käyttäjä sitä tarvitsee. Vuorisen ja Tetrin (2016) mukaan turvallisuus muodostuukin jollain tapaa suhteista ja omistajuudesta, eli kysymys on siitä, kuka hyväksytty käyttäjä on, kenellä on omistajuus tietoon ja mitä valtuutettu käyttäjä aikoo tiedolla tehdä. Toisin sanoen heidän ajatuksensa mukaan turvallisuus on pitkälti yhteydessä käyttäjien haluihin, joten on mahdotonta tosiasiallisesti sanoa pelkästään itse tiedostoa tarkastelemalla, onko jokin tiedosto turvassa vai ei. (Vuorinen & Tetri, 2016.)

Vuorinen ja Tetri (2012) pohtivat turvallisuuden olemusta hieman laajemmin toisessa artikkelissaan *The Order Machine – The Ontology of Information Security* eräänlaisen konemallin kautta, mutta päällimmäisenä ajatuksena he tuovat esiin tietynlaista kontrollin ja järjestyksen merkityksellisyyttä turvallisuuden suhteen. Heidän mallissaan ikään kuin pyritään pitämään jonkinlaista järjestystä

ja kontrollia yllä sisäisesti ja sitä kautta taistelemaan ulkoista kaaoksellisuutta vastaan. Ulkoiset voimat voivat kuitenkin pahimmassa tapauksessa jollakin tavalla häiritä tätä sisäistä järjestystä, jolloin turvallisuus voi järkkyä. Artikkelin esimerkkejä mukaillen voitaisiin ajatella, että olisimme esimerkiksi talon sisällä, jonka ovet ovat lukossa. Tällöin sisätiloissa voidaan ajatella vallitsevan jonkinlainen järjestys, eli turva. Kun ulko-ovi avataan, altistuu sisätilat ulkomaailman kaaoksellisuudelle ja ohikulkijat voivat halutessaan pyrkiä sisälle ja hakea pöydällä lojuvan postin, eli häiritä vallitsevaa järjestystä tai pääsyä tietoon, jolloin tietoturvallisuus on jossain määrin menetetty. (Vuorinen & Tetri, 2012.)

Tältä pohjalta tarkasteltuna kyberturvallisuudessa voisi olla kyse juuri tällaisesta Vuorisen ja Tetrin (2012, 2016) ajatuksen mukaisesta epäjärjestyksen ulkona pitämisestä ja pyrkimyksestä erilaisten turvatoimien avulla säilyttää järjestys ja kontrolli tietyillä alueilla oman organisaation hallussa. Toisaalta Vuorinen ja Tetri (2012, 2016) nostavat esiin seikan, jossa itse turvatoimet voivat häiritä sisäistä järjestystä ja CIA:n toteutumista, sillä esimerkiksi edellä mainitussa esimerkissä talon omistaja joutuisi avaamaan oven lukon, avaamaan oven erikseen ja sitten vasta pääsisi käsiksi päivän postiin, jolloin posti ei ole enää niin helposti saatavilla verrattuna siihen, että ovi olisi valmiiksi auki. Hieman samankaltainen ilmiö tapahtuu, kun käyttäjät joutuvat kirjautumaan vaikkapa yrityksen järjestelmiin, eli Vuorisen ja Tetrin näkemyksen mukaan sisäinen järjestys ja informaation virta häiriintyy. (Vuorinen & Tetri, 2012, 2016) Tältä pohjalta voitaisiin loogisesti päätellä itse kyberturvatoimienkin sisältävän uhkia ja riskejä tietoturvan toteutumisen kannalta, sillä miksei kyseisen oven lukon voida ajatella olevan altis rikkoutumiselle, jolloin valtuutettu käyttäjäkään ei pääse enää postiin käsiksi (CIA ei toteudu).

3.1.2 Kyberuhat ja niiden torjunta

Jotta edellisessä luvussa kuvattu kontrolli ja turva voitaisiin säilyttää, on aiheellista tutustua kyberympäristössä ilmeneviin uhkiin ja eri mahdollisuuksiin torjua niitä.

Craigen ym. (2014), Ozkayan (2019) ja Uman ja Padmavathin (2013) mukaan kyberuhat ovat tietoteknisiin järjestelmiin ja verkkoihin kohdistuvia tarkoituksellisia toimia, jotka pyrkivät aiheuttamaan vahinkoa tai häiriötä. Heidän mukaansa kyberuhkien kirjo on laaja, ja ne voivat kohdistua erilaisiin kohteisiin, kuten yksittäisiin tietokoneisiin, tietoverkkoihin, yritysten tai organisaatioiden järjestelmiin, kriittiseen infrastruktuuriin ja valtioiden turvallisuuteen liittyviin kohteisiin. Lähdekirjallisuuden perusteella kyberuhkien muodot ja vaikutukset voivat vaihdella huomattavasti ja ne voivatkin aiheuttaa esimerkiksi taloudellisia tappioita liiketoiminnassa, yksityisyyden menettämisen yksittäisen henkilön kohdalla, muita henkilökohtaisia vahinkoja tai jopa vaarantaa kansallisen turvallisuuden. (Craigen ym., 2014; *Cybersecurity Guide for SMEs*, 2021; Ozkaya, 2019; Uma & Padmavathi, 2013.)

Kyberuhkien torjunnan osalta Uman ja Padmavathin (2013) mukaan on hyvin tärkeää pystyä luokittelemaan erilaisia kyberuhkia, jotta saavutettaisiin tietoisuus erilaisista kyberhyökkäyksistä ja niiden toimintatavoista. Heidän mukaansa tarkan luokittelun avulla voidaan puolustautua asianmukaisesti

erityyppisiä kyberuhkia ja –hyökkäyksiä vastaan. (Uma & Padmavathi, 2013.) Kyberuhkien luokittelu näyttää kuitenkin tieteellisen kirjallisuuden perusteella olevan haasteellista juuri siksi, että niiden muodot ja tavoitteet voivat vaihdella suuresti. Tässä tutkielmassa käytetyssä lähdekirjallisuudessa esiintyy kuitenkin joitakin yleisiä tapoja luokitella kyberuhkia, joista on koostettu taulukko (taulukko 2), missä ilmenee lähdekirjallisuudessa yleisimmin mainitut luokittelutavat ja esimerkit kyberuhista luokittain. Kyberuhkia on luokiteltu kirjallisuudessa esimerkiksi hyökkäysmenetelmän ja -kohteen mukaan. (Fischer, 2016; Heim & Wessel, 2023; Ozkaya, 2019; Pavithra & Femilda Josephin, 2020; Tang ym., 2023.)

On kuitenkin tärkeää huomioida, että kyberuhat voivat olla myös hyvin monimutkaisia ja monitahoisia, jolloin niiden luokittelu yhteen kategoriaan voi Uman ja Padmavathin (2013) mukaan olla hankalaa tai jopa mahdotonta. Niinpä taulukossa (taulukko 2) mainittujen luokittelutapojen lisäksi tieteellisessä kirjallisuudessa on olemassa myös paljon muita tapoja luokitella kyberuhkia. (Uma & Padmavathi, 2013.)

TAULUKKO 2 Kyberuhkien luokitteluesimerkit

Kyberuhkien luokittelu	**Näkökulma**	**Esimerkit**
Hyökkäysmenetelmän mukaan	Millä tavalla hyökkäys tapahtuu?	Haittaohjelmat, tietomurrot, tietojenkalastelu, palvelunestohyökkäys (DoS)(Fischer, 2016; Heim & Wessel, 2023; Tang ym., 2023; Uma & Padmavathi, 2013)
Vaikutuksen mukaan	Millaisia vaikutuksia hyökkäys aiheuttaa?	Tietojen varastaminen tai tuhoutuminen, palvelun estyminen, verkon tai järjestelmän kaatuminen (Fischer, 2016; Heim & Wessel, 2023; Tang ym., 2023; Uma & Padmavathi, 2013)
Kohteen mukaan	Mikä/kuka on hyökkäyksen kohteena?	Yksittäiset tietokoneet, yritykset, kriittinen infrastruktuuri, valtiot (Fischer, 2016; Heim & Wessel, 2023; Tang ym., 2023; Uma & Padmavathi, 2013)
Tavoitteen mukaan	Mikä on hyökkääjän motiivi?	Taloudellisen hyödyn tavoittelu, vakoilu, sabotointi, poliittinen agendojen edistäminen (Fischer, 2016; Heim & Wessel, 2023; Tang ym., 2023; Uma & Padmavathi, 2013)

Kuinka näiltä uhilta voidaan sitten suojautua? Perinteisin menetelmin kyberuhkien torjunnassa käytetään useita erilaisia tekniikoita ja menetelmiä riippuen

uhkien luonteesta. Lezzin ym. (2018) tekemän tutkimuksen perusteella voidaan näistä tärkeimpänä pitää tietoturvaohjelmistoja, jotka ovat keskeinen osa tietoturvan perusratkaisua. Tietoturvaohjelmistoja käytetään tietokonejärjestelmien suojaamiseksi haittaohjelmilta, viruksilta ja muilta vastaavilta uhilta. Tällaisia ohjelmistoja ovat esimerkiksi palomuurit, virustorjuntaohjelmistot ja edellä mainitut roskapostin suodattimet. (Lezzi ym., 2018; Ozkaya, 2019; Padallan, 2019.)

Toinen lähdeaineiston perusteella yleisesti tärkeänä pidetty menetelmä on käyttäjien koulutus ja tietoisuuden lisääminen kyberuhkiin liittyen. Kuten Mosteanu (2020) ja monet muut ovat todenneet, käyttäjät ovat usein heikoin lenkki tietoturvaketjussa, joten käyttäjien tietoisuuden ja osaamisen kehittäminen on erityisen tärkeää. Osaamisen ja tietoisuuden kasvattamiseen liittyy esimerkiksi niin salasanakäytäntöjen ja tietojen jakamisen ohjeistaminen, kuin käyttäjämanipulaation (social engineering) tunnistamisen opettaminen. (August ym., 2022; Mosteanu, 2020; Ozkaya, 2019.)

Kolmas perinteinen menetelmä on Augustin ym. (2022) ja Lezzin ym. (2018) mukaan tietoturva-aukkojen paikkaaminen. Tietoturva-aukot ovat ohjelmistojen tai järjestelmien haavoittuvuuksia, joiden kautta hyökkääjä voi päästä käsiksi tietokonejärjestelmään. Näiden haavoittuvuuksien paikkaaminen vaatii säännöllistä päivitysten ja korjausten asentamista, jotta tietojärjestelmät pysyisivät turvallisempina. (August ym., 2022; Lezzi ym., 2018.) Hieman Vuorisen ja Tetrin (2012, 2016) näkemystä tietoturvan häiriintymisestä sivuten voidaan ajatella myös ohjelmistojen päivityksiin liittyvän aina paradoksaalisesti potentiaalisia vaaroja, mikäli ne aiheuttavat ennakoimattomia ongelmia kompleksisissa järjestelmissä (Vuorinen & Tetri, 2012, 2016). Esimerkiksi Adams (2010) kirjoitti erään McAfeen virustorjuntaohjelmiston version ongelmista, kun kyseinen ohjelmisto luuli Windowsin prosessiin liittyvää osaa haittaohjelmaksi, minkä seurauksena tämän version vaikutuksenalaisena olleet koneet ajautuivat uudelleenkäynnistysten silmukkaan (Adams, 2010). Tässä tapauksessa CIA ei luonnollisesti toteudu, sillä tiedot eivät ole käyttäjän saatavilla koneen kaatuessa, joten tietoturvan voidaan sanoa siltä osin pettäneen, vaikka tarkoitus oli tietoturvapäivityksen osalta päinvastainen.

Neljäntenä erittäin tärkeänä varautumiskeinona uhkia vastaan kirjallisuudessa pidetään tietojen varmuuskopiointia, joka mahdollistaa tietojen nopean palauttamisen vahinkotilanteessa. Tiedostojen varmuuskopioinnin avulla voidaankin merkittävästi vähentää haittoja, joita kyberhyökkäyksen toteutuessa voi esimerkiksi liiketoiminnalle tietojen kadotessa aiheutua. (August ym., 2022; Lezzi ym., 2018; Ozkaya, 2019.)

3.2 Jatkuva valvonta ja sen merkitys kyberturvallisuudessa

Jatkuvalla valvonnalla (continuous monitoring) tarkoitetaan useiden tutkijoiden mukaan (esim. Ahmed ym., 2019; Alkahtani & Aldhyani, 2022; Khanna ym., 2018) tietojärjestelmän tai tietoverkon jatkuvaa seurantaa, jonka avulla pyritään havaitsemaan tietoturvaan liittyvät ongelmat tai poikkeavuudet mahdollisimman

varhaisessa vaiheessa. Sen ymmärretään olevan jatkuvaa tiedonkeruuta ja analyysiä kyberturvallisuuden näkökulmasta, jotta turvallisuusriskeihin reagointi olisi mahdollisimman nopeaa ja tehokasta. Jatkuva valvonta on lähdekirjallisuuden perusteella käytännön toimia ja menetelmiä, joissa tietojärjestelmiä, järjestelmissä olevia tietoja, verkkoliikennettä ja muita tietoliikenteen osia seurataan jatkuvasti näiden turvallisuushäiriöiden ja tietoturvaongelmien (esim. haittaohjelmat) osalta. Jatkuvan valvonnan menetelmät perustuvat usein automaattisiin järjestelmiin ja ohjelmistoihin, jotka keräävät ja analysoivat tietoa järjestelmän tilasta ja toiminnasta. (Ahmed ym., 2019; Alkahtani & Aldhyani, 2022; Khanna ym., 2018; Wiafe ym., 2020.)

Muun muassa Wiafen ym. (2020) ja Ahmedin ym. (2019) mukaan kyberturvallisuudessa jatkuvan valvonnan merkitys on huomattava, sillä kyberhyökkäykset voivat tapahtua erittäin nopeasti ja yllättäen, eikä niiden vaikutuksia pystytä aina ennakoimaan tai arvioimaan etukäteen. Siksi onkin tärkeää kyetä reagoimaan uhkiin mahdollisimman varhaisessa vaiheessa, ennen kuin mahdollinen hyökkäys ehtii aiheuttamaan merkittäviä vahinkoja tai vääristämään järjestelmien toimintaa. (Ahmed ym., 2019; Fischer, 2016; Wiafe ym., 2020.)

Jatkuvan valvonnan menetelmiä voidaan Ansarin ym. (2022) mukaan soveltaa laajasti erilaisissa tietojärjestelmissä ja verkkojen turvallisuuden varmistamisessa. Esimerkiksi verkkoliikenteen monitorointi, lokitietojen analysointi, virustorjunnan päivitykset ja haavoittuvuuksien skannaus ovat Aljuhanin (2021) ja Ansarin ym. (2022) mukaan tärkeitä osia jatkuvan valvonnan prosesseissa. (Aljuhani, 2021; Ansari ym., 2022.)

Toki jatkuvan valvonnan toteuttamiseen liittyy Ahmedin ym. (2019) mukaan sen hyötyjen lisäksi myös haasteita. Jatkuvan valvonnan haasteena pidetään yleisesti muun muassa sen käsittelemää suurta tietomäärää ja sen hallintaa, erilaisten järjestelmien monimutkaisuutta, vääriä hälytyksiä ja hälytysten kokonaismäärää. Ahmedin ym. (2019) mukaan väärät hälytykset uhkien havainnoinnissa voivat pahimmillaan hukuttaa oikeat hälytykset alleen. Näiden lisäksi he alleviivaavat jatkuvan valvonnan vaativan paljon resursseja ja henkilöstöä, jotka pystyvät tarvittaessa tulkitsemaan kerättyä tietoa ja reagoimaan mahdollisiin uhkiin ja hälytyksiin. (Ahmed ym., 2019.)

3.3 Kyberturvallisuus teollisuudessa

Kyberuhista ja jatkuvasta valvonnasta puhuttaessa tämän tutkielman lähdeaineiston perusteella näyttää siltä, että yksi keskeisimmistä uhkasektoreista kyberturvallisuuden alalla on teollisuussektori, joka näyttää olevan erityisen altis kyberhyökkäyksille ja teollisuusvakoilulle. Teollisuussektorin kyberturvallisuuteen kohdistuu aineiston perusteella huomattavia haasteita, sillä monimutkaiset valvomo- ja ohjausjärjestelmät, sekä laitteistojen ja ohjelmistojen vanhentuminen ovat kasvattaneet kyseisen sektorin haavoittuvuutta ja mielenkiintoa kyberrikollisuuden näkökulmasta. (Lezzi ym., 2018; Mattila ym., 2020; National Institute of Standards and Technology, 2018; Vartolomei & Avasilcăi, 2020.) Teollisuus on monessa yhteiskunnassa, kuten Suomessa, yhä nykypäivänäkin erittäin tärkeä

alue, joten tämänkin tutkielman kannalta on hyvä tarkastella lyhyesti teollisuuden kyberturvallisuuteen liittyviä erityispiirteitä. (Mattila ym., 2020; *Suomen teollisuus rakennemuutoksen pyörteissä*, 2017.)

Tarkasteltaessa teollisuussektorin tietoturvariskejä, yksi merkittävä huolenaihe muun muassa Culotin ym. (2019) ja Lezziin ym. (2018) mukaan on teollisuuslaitteiden ja -järjestelmien haavoittuvuudet. Heidän mukaansa tämä on johtanut huoleen teollisuusautomaation käyttöönoton turvallisuudesta ja kyvystä puolustautua kyberhyökkäyksiä vastaan, erityisesti viime vuosien aikana teollisuusautomaation järjestelmiin kohdistuvien hyökkäysten määrän kasvun seurauksena. Kyberturvallisuus nähdäänkin yhä tärkeämpänä teollisuussektorilla, sillä useat teollisuusyritykset ovat hyvin riippuvaisia tänä päivänä tietotekniikasta ja muista sähköisistä järjestelmistä. Erityisen haavoittuvaisina lähdeaineiston perusteella nähdään kriittinen infrastruktuuri, kuten sähköverkot, vesihuoltojärjestelmät, lentoliikenteen hallintajärjestelmät ja muut vastaavat järjestelmät. (Culot ym., 2019; ”Cyber War”, 2011; Lezzi ym., 2018.)

Yksi mielenkiintoinen uhkavektori teollisuuden alalla on muun muassa Culotin ym. (2019), Khorshedin ym. (2012) ja Aljuhanin (2021) mukaan niin sanotun teollisuuden neljännen vallankumouksen (Industry 4.0) seurauksena yleistyneet teollisuuden Internet of Things-järjestelmät (IIoT), jotka liittyvät teolliseen automaatioon ja tehdasautomaatioon. Kazançoğlu ja Özkan-Özen (2021) mukaan tälle Industry 4.0 ”vallankumoukselle” on tyypillistä siirtyminen täysin digitaalisiin tuotantotapoihin ja -ympäristöihin. IIoT-laitteet keräävätkin dataa reaaliaikaisesti, mikä auttaa yrityksiä tehostamaan prosessejaan ja optimoimaan tuotantoa. IIoT-laitteet voivat olla kuitenkin hyvin alttiita kyberhyökkäyksille, mikä voi johtaa vakaviin seurauksiin, kuten tuotantokatkoksiin, laitteiden vioittumiseen tai jopa henkilövahinkoihin. (Aljuhani, 2021; Culot ym., 2019; Khorshed ym., 2012; Ozkan-Ozen & Kazancoglu, 2021; Uma & Padmavathi, 2013.)

Culotin ym. (2019), Gajekin ym. (2021) ja Lezzin ym. (2018) mukaan IIoT-järjestelmiin liittyvät riskit ovat moninaisia ja ne voivat johtua esimerkiksi laitteiden huonosta suojaustasosta, puutteellisista tietoturvakäytänteistä tai vanhentuneista ohjelmistoista. Jotta IIoT-järjestelmät voidaan pitää suojattuna ja turvallisina, on heidän mukaansa tärkeää toteuttaa kattavaa tietoturvapolitiikkaa, joka sisältää esimerkiksi laitteiden suojausmekanismit, säännölliset tietoturvapäivitykset, jatkuvan valvonnan ja henkilöstön koulutuksen. (Culot ym., 2019; Gajek ym., 2021; Lezzi ym., 2018.)

Toinen merkittävä pinnalla oleva hyökkäyspolku teollisuuden alalla näyttäisi aineiston perusteella olevan niin sanotut ransomware-hyökkäykset (kiristyshaittaohjelma), joissa kyberrikolliset salaavat uhrin tiedostot ja vaativat lunnaita salauksen purkamisesta. Tämä nähdään erityisen vaaralliseksi teollisuuden alalla, sillä tietojen ja tiedostojen salaus ulkopuolisen toimesta voi vaikuttaa merkittävästi tuotantolaitteisiin ja -prosesseihin, mikä puolestaan johtaa vakaviin häiriöihin tuotannossa ja aiheuttaa tarpeettomia kustannuksia liiketoiminnalle. Esimerkiksi vuonna 2019 Norjan alumiinitehdas Norsk Hydro joutui ransomware-hyökkäyksen kohteeksi, mikä vaikutti merkittävästi yhtiön tuotantoon aiheuttaen joidenkin arvioiden mukaan noin 40 miljoonan dollarin kustannukset. (August ym., 2022; Bécue ym., 2021; Bello ym., 2021; Culot ym., 2019; *Cyber-Attack on Hydro*, 2020.)

Teollisuusautomaation järjestelmiin on kohdistunut lukuisia muitakin hyökkäyksiä, joista tunnetuimpien joukossa lienee Stuxnet ja Triton-hyökkäykset. Suxnet-hyökkäyksen tavoitteena oli vahingoittaa Iranin ydinteknologiaohjelmaa, ja se onnistuikin hyökkäämään Siemensin valmistamaan SCADA-järjestelmään muuttaen sen toimintaa niin, että Iranin ydinkeskus kärsi merkittäviä vaurioita. Lyhykäisyydessään tämä tapahtui ohjelmoimalla vääriä komentoja PLC-kontrollereihin, jotka ohjasivat sentrifugeja toimimaan väärällä tavalla ylikuormittaen niitä. Kyseinen hyökkäys oli vakava, mutta toisaalta se paljasti ja toi yleiseen tietoon teollisuusautomaation järjestelmien turvallisuusriskejä. (Gajek ym., 2021; Ozkaya, 2019.)

Triton-hyökkäys on toinen esimerkki vastaavanlaisesta tapauksesta, mikä oli suunnattu saudiarabialaisen öljynjalostamon turvajärjestelmiin. Se oli suunniteltu manipuloimaan turvajärjestelmää niin, että se kykenisi sammuttamaan teollisuuslaitteiden turvallisuutta valvovan järjestelmän. Mikäli tämä hyökkäys olisi onnistunut, olisi se voinut puolestaan johtaa vakavaan onnettomuuteen. (*Triton Is the World's Most Murderous Malware, and It's Spreading*, ei pvm.)

Näiden hyökkäysten kohdalla on juurikin yritetty, ja Stuxnetin tapauksessa onnistuttu, vaikuttaa luvussa 3.1.1 sivuttuun organisaation sisäiseen järjestykseen ja kontrolliin, ottamalla toimintoja ja prosesseja hyökkääjän omaan hallintaan, jolloin CIA:n toteutuminen häiriintyy.

Ransomwaren ja IIoT-järjestelmien lisäksi teollisuussektoriin liittyy lähdeaineiston perusteella vahvasti myös muita hyökkäyspolkuja, kuten esimerkiksi tietojärjestelmien heikkoudet ja henkilöstöön kohdistuvat hyökkäykset, käyttäjien manipulointi ja vanha, mutta toimiva USB-hyökkäyspolku. Näiden seikkojen vuoksi onkin erittäin tärkeää, että teollisuuden yritykset ymmärtävät kyberympäristöön liittyvät uhkakuvat ja ryhtyvät toimiin välttääkseen kyberhyökkäysten mahdollisen toteutumisen seuraamukset. (Culot ym., 2019; Gajek ym., 2021; Lezzi ym., 2018; Tang ym., 2023.)

4 TEKOÄLY KYBERTURVALLISUUDESSA

Useat tutkimukset (esim. Parisi, 2019; Wiafe ym., 2020) ovat korostaneet tekoälyn merkitystä kyberuhkien torjunnassa. Erityisesti jatkuvan valvonnan mahdollistavat tekoälypohjaiset ratkaisut ovat osoittautuneet muun muassa Khannan ym. (2018), Russellin ja Norvigin (2010) ja Sonin (2020) mukaan tutkimuksissa erittäin tehokkaiksi (Khanna ym., 2018; Russell & Norvig, 2010; Soni, 2020). Toisaalta, vaikka tekoälyllä onkin paljon potentiaalia kyberuhkien torjunnan saralla, on sen käytössä havaittu myös selkeitä haasteita, joita käsitellään tarkemmin tämän luvun lopussa (Ansari ym., 2022; Chellappan ym., 2018; Fischer, 2016).

Kokonaisuudessaan tekoälyn soveltaminen kyberturvallisuudessa näyttäisi tieteellisen kirjallisuuden perusteella olevan edelleen suhteellisen uusi ilmiö, vaikkakin tutkimusalana rajusti kasvava. Tekoälyn sovelluskohteet kyberturvallisuudessa vaihtelevat erilaisten tehtävien suorittamisesta aina monimutkaisten järjestelmien hallintaan, ja sitä voidaankin käyttää esimerkiksi tietoturvatietojen keräämiseen, havaitsemiseen ja analysointiin, käyttäjätunnusten hallintaan, verkkoliikenteen hallintaan, tietojen salaamiseen ja suojaamiseen, sekä erilaisten tietoturvauhkien automaattiseen torjuntaan. (Alkahtani & Aldhyani, 2022; Mosteanu, 2020; Wiafe ym., 2020.)

Kyberturvallisuuteen tekoälyn nähdäänkin tuovan parempaa ja tarkempaa suojaa verrattuna perinteisiin menetelmiin ilman tekoälyn tuomaa apua. Kuitenkin, kuten kaikki tietojenkäsittelymenetelmät, myös tekoälysovellukset ovat muun muassa Chellappan ym. (2018) mukaan alttiita virheille ja hyökkäyksille. Tästä syystä erilaiset tekoälyä hyödyntävät ratkaisut on suunniteltava ja toteutettava huolellisesti. (Bécue ym., 2021; Chellappan ym., 2018; Fischer, 2016.)

Niinpä tässä luvussa käsitellään ensin hieman tekoälyn hyötyjä ja haittoja kyberturvallisuudessa, minkä jälkeen paneudutaan tekoälyn sovelluskohteisiin ja käytössä oleviin sovelluksiin kyberturvallisuudessa ja sen jatkuvassa valvonnassa. Lisäksi pohditaan hieman kyberturvallisuusammattilaisten roolin kehitystä tekoälyn käyttöönoton pohjalta, sillä edellisissäkin luvuissa nousi esiin huoli resurssien riittävyydestä kyberturvallisuudesta vastaavan henkilöstönkin osalta. Julkiseen keskusteluun on lisäksi tämän tutkielman kirjoittamisen aikana noussut ihmisten huoli tekoälyn yleistymisen vaikutuksista työllisyysnäkymiin. (esim. Eloundou ym., 2023; Högmander, 2019; Ozkan-Ozen & Kazancoglu, 2021.)

4.1 Tekoälyn hyödyt kyberturvallisuudessa

Tekoälyn käytöllä on tieteellisessä kirjallisuudessa tunnistettu olevan useita etuja ja paljon potentiaalia kyberturvallisuuden parantamisessa ja kyberuhkien torjunnassa. Muun muassa Ahmed ym. (2019) ja Ansarin ym. (2022) mukaan tekoäly voi auttaa havaitsemaan ja estämään uhkia nopeammin ja tarkemmin kuin perinteisesti kyberturvallisuudessa käytetyt menetelmät. Tämä johtuu heidän mukaansa pääosin tekoälyn kyvystä käsitellä suuria määriä dataa ja tietoa, sekä kyvystä havaita poikkeamia toimintaympäristössä reaaliajassa. Tekoäly voi myös parantaa merkittävästi uhkien havaitsemisen ja vastatoimien automatisointia, mikä puolestaan nopeuttaa uhkiin reagointia ja vähentää inhimillisten virheiden riskiä toiminnassa. (Ahmed ym., 2019; Ansari ym., 2022; Soni, 2020.)

Bellon ym. (2021), Debarin ym. (2000) ja Aljuhanin (2021) mukaan erityisesti ennakoivan tekoälyn sovellukset kyberuhkien torjunnassa voivat tarjota merkittäviä etuja, sillä ne voivat auttaa havaitsemaan uhkia ennen kuin ne kykenevät aiheuttamaan vahinkoa hyökkäyksen kohteelle, mikä luonnollisesti tarjoaa mahdollisuuden ennakoivalle puolustautumiselle. Heidän mukaansa esimerkiksi koneoppimismenetelmiä voidaan käyttää tehokkaasti automaattiseen haittaohjelmien havaitsemiseen ja poistamiseen, sekä anomalioiden havaitsemiseen verkkoliikenteessä, jotka voivat viitata kyberhyökkäykseen tai sen uhkaan. Neuroverkkoihin perustuvat tekoälymenetelmät nähdään kirjallisuudessa puolestaan suureksi hyödyksi erityisesti tuntemattomien uhkien havaitsemisessa. (Aljuhani, 2021; Ansari ym., 2022; Bello ym., 2021; Debar ym., 2000.)

Näiden ennakointiin perustuvien hyötyjen lisäksi merkittäväksi eduksi Ansarin ym. (2022) mukaan nähdään tekoälyn kyky oppia jatkuvasti uusista kyberuhista ja kyky sopeutua niihin nopeasti (Ansari ym., 2022). Tällainen ominaisuus ja kyky mahdollistaakin nopean reagoinnin pyrittäessä estämään uusia, ennen näkemättömiäkin uhkia aiheuttamasta vahinkoa. Toisaalta tekoälystä voi näiden osa-alueiden lisäksi Mosteanun (2020) mukaan olla vielä suurempi apu myös tietoturvan hallinnassa, kuten resurssien hallinnassa, kirjautumisten valvonnassa ja riskien hallinnassa. (Mosteanu, 2020.)

4.2 Tekoälyn haasteet ja rajoitteet kyberturvallisuudessa

Vaikka tekoälyn nähdään tutkielman lähdeaineiston perusteella tarjoavan lukuisia mahdollisuuksia kyberturvallisuuden parantamiseksi, on sillä havaittu olevan myös selkeitä haasteita ja rajoitteita, jotka on hyvä ottaa huomioon. Tekoälyn haasteet kyseisessä kontekstissa liittyvät aineiston mukaan erityisesti sen saamien tietojen ja datan laatuun, määrään ja monimuotoisuuteen, tekoälyn väärinkäyttöön, inhimillisiin virheisiin ja eettisiin ja juridisiin kysymyksiin. (Ansari ym., 2022; Bello ym., 2021; Debar ym., 2000; Zeadally ym., 2020.)

Alkahtanin ja Aldhyanin (2022) ja monien muiden tutkijoiden mukaan tekoälyn suorituskyky yleisesti riippuu erityisesti siitä, kuinka hyvin se on opetettu ja koulutettu käyttämään dataa ja tietoja. Tieteellisessä kirjallisuudessa on

nostettu esiin myös huoli tekoälyn saaman materiaalin laadusta, sillä jos se on huonoa, voi se helposti saada tekoälyn tekemään virheellisiä johtopäätöksiä. Se on tietysti seikka, joka voi heikentää sen tehokkuutta kyberuhkien torjunnassa ja olla itsessäänkin uhka turvallisuudelle. Dongin ym. (2018) mukaan on olemassa useita tekoälyalgoritmeja, jotka nimenomaan käyttävät tätä haavoittuvuutta hyväkseen ja hyökkäävät esimerkiksi opetusmateriaalia vastaan tai syöttämällä saastunutta dataa tekoälylle, aiheuttaen pitkäaikaisia haittoja tekoälyn toimintaan. (Alkahtani & Aldhyani, 2022; Dong ym., ei pvm.; Khorshed ym., 2012; Soni, 2020; Uma & Padmavathi, 2013.)

Myös Ahmed ym. (2019) ja Khanna ym. (2018) yhtyvät huoleen datan määrästä ja monimuotoisuudesta. Khanna ym. (2018) mukaan käytettävissä olevan datan rajallisuus ja yksipuolisuus voi merkittävästi heikentää tekoälyn tarkkuutta ja tehokkuutta. Ahmed ym. (2019) korostavat näiden lisäksi myös tekoälyssä käytettyjen algoritmien tehokkuutta. Tieteellisessä kirjallisuudessa puhutaankin paljon tekoälyalgoritmien yli- ja alisovitteisuuksista, jotka voivat vaikuttaa huomattavasti tekoälyn suoriutumiseen myös kyberuhkien torjunnassa. Esimerkiksi liian ylisovitteinen algoritmi voi tarttua liian herkästi näennäisiin kyberuhkiin ja sitä kautta täyttää organisaatiossa tietoturvallisuudesta vastaavien työpöydän niin sanotuista vääristä hälytyksistä. Tällainen ilmiö puolestaan sotii muun muassa Mosteanun (2020) ajatusta tehokkaammasta resurssien hallinnasta ja yleistä kirjallisuudessa esiintyvää ajatusta tekoälyn vähentämästä työkuormasta vastaan. Toisaalta näitäkin ongelmia vastaan on kehitelty ja kehitetään uusia ratkaisuja. Esimerkiksi tekoälyteknologioiden parissa työskentelevä, yksi nykypäivän viitatuimmista tietokonetieteilijöistä, Ilya Sutskever kollegoineen esittää joidenkin ylisovitteisten Convultional Neural Network (CNN)-algoritmien vaihtoehtoiseksi parannuskeinoksi ”dropout” -menetelmää, joka pyrkii eliminoimaan sen negatiiviset vaikutukset ja säästämään GPU:lta (Graphics Processing Unit) vaadittavaa suorityskykyä. (Ahmed ym., 2019; Ansari ym., 2022; Khanna ym., 2018; Mosteanu, 2020; Pavithra & Femilda Josephin, 2020; Sutskever ym., 2017.)

Näin ollen on hyvä muistaa myös se, että tekoälyteknologiat ovat ihmisten suunnittelemia ja ylläpitämiä. Siksipä tekoälyjärjestelmien suunnittelussa ja ylläpidossa voi tapahtua virheitä, jotka voivat aiheuttaa vakaviakin turvallisuusriskejä myöhemmässä käytön vaiheessa. Esimerkiksi Aljuhanin (2021) mukaan erityisesti koneoppimisen opettamiseen käytettävät koulutuspaketit voivat olla nykypäivänä jo vanhentuneeksi luokiteltuja. Tällaisten vanhentuneiden pakettien käyttö koulutuksessa voidaan laskea myös inhimilliseksi tai tiedostetuksi virheeksi tekoälyn käyttöönoton, kehityksen ja suunnittelun vaiheessa. (Ahmed ym., 2019; Aljuhani, 2021; Khanna ym., 2018; Mosteanu, 2020.)

Lisäksi tekoälyn käyttöön liittyy vahvasti myös eettisiä ja juridisia kysymyksiä, joita on hyvä ottaa huomioon. Tieteellisen kirjallisuuden pohjalta mielenkiintoisia juridisia ja eettisiä kysymyksiä tekoälyn käytön näkökulmasta ovat esimerkiksi yksityisyydensuoja, vastuu- ja vahingonkorvauskysymykset ongelmatilanteissa, sekä vielä hieman puutteellinen tekoälyyn liittyvä lainsäädäntö. Voutilaisen (2023) mukaan tekoälyyn liittyvään lainsäädäntöön on toki tulossa päivityksiä EU-tasollakin (Euroopan unioni) vuoden 2023 loppuun mennessä, mutta teknologian nopean kehityksen myötä uusikin lainsäädäntö voi olla

hieman jälkijättöistä. Niinpä tähän liittyviä kysymyksiä on hänen mukaansa syytä pohtia perusteellisesti tekoälyn käyttöönoton yhteydessä. (Goertzel, 2014; Russell & Norvig, 2010; Soni, 2020; Voutilainen, 2023.)

Vaikka tekoäly kykenee hämmästyttävällä tavalla sekä tunnistamaan useita tunnettuja uhkia, että ennustamaan uusia, voi lähdeaineiston perusteella uusien ja täysin tuntemattomien uhkien havaitseminen nykyisillä ratkaisuilla olla kuitenkin vielä haasteellista. Rajoitteisiin lasketaan myös edellä mainittuihin eettisiin kysymyksiin liittyvä tekoälyn toiminnan läpinäkyvyyden puute ja kyvyttömyys käsitellä kontekstuaalista tietoa. Tekoälyn toiminnasta puuttuu kirjallisuuden perusteella myös täysin ihmiselle ominainen intuitio, josta voisi olla hyötyä monessa eri suhteessa kyberturvallisuuden kannalta, erityisesti uusien uhkien tunnistamisessa. Lisäksi itse tekoälyn käyttöönottoon näyttää liittyvän useita erilaisia haasteita. Haasteena tai esteenä tekoälyn käyttöönottoon varsinkin pienemmissä organisaatioissa nähdään yleisesti teknologioiden korkeat kustannukset, tekniset haasteet ja käyttäjien vastustus. Tekoäly voi aiheuttaa muun uuden teknologian käyttöönoton tavoin henkilökunnassa yleistä vastustusta ja jopa pelkoa. (Hamet & Tremblay, 2017; Mosteanu, 2020; Soni, 2020.)

Näiden rajoitteiden huomioonottamattomuus voi muun muassa Ansarin ym. (2022), Parisin (2019) ja Vartolomein ja Avasicain (2019) mukaan tulla tekoälyä käyttävän organisaation kannalta kalliiksi ja olla monin tavoin haitallista, joten niiden vaikutus on otettava huomioon. Näiden seikkojen vuoksi onkin tärkeää, että tekoälypohjaisia ratkaisuja kehitetään ja arvioidaan jatkuvasti, jotta niiden tarkkuus ja luotettavuus voitaisiin varmistaa. (Ansari ym., 2022; Parisi, 2019; Vartolomei & Avasilcăi, 2020.)

4.3 Kyberuhkien torjunnassa käytettäviä tekoälyratkaisuja

Keskeisimpinä tekoälyn sovelluskohteina kyberturvallisuuden alalla voidaan tutkielman lähdeaineiston pohjalta pitää ennakoivaa analytiikkaa, käyttäjän tunnistusta, haittaohjelmien torjuntaa ja tietojen salausta. Teollisuuden kontekstissa tekoälyllä on nähty olevan potentiaalia näiden lisäksi myös verkkoliikenteen skannauksessa ja fyysisen turvallisuuden valvonnassa.

Ennakoivassa analytiikassa tekoälyn avulla kerätään ja analysoidaan suuria tietomääriä ja sen pohjalta tunnistetaan poikkeuksellisia toimintoja, joita ei perinteisin menetelmin tavallisesti havaittaisi. Käyttäjien tunnistuksessa puolestaan sen avulla tunnistetaan käyttäjien tietokoneisiin ja verkkoihin liittyviä ongelmia, kuten tuntemattomia laitteita tai kirjautumisyrityksiä. Haittaohjelmien torjunnassakin kyseisen teknologian hyödyntämistä jo sivuttiin, joten sen hyödyt virusten, troijalaisten ja matojen havaitsemisessa ovat kiistattomat. Verkkoturvallisuuden näkökulmasta taas muun muassa Aljuhanin (2021) mukaan tekoälyllä on suuri hyöty verkkohyökkäysten, kuten DoS-hyökkäysten (Denial of Service) ja verkkoskannausten havainnoinnissa ja torjunnassa. Myös tietojen salaukseen tekoälystä on suuri apu, sillä sen avulla voidaan suojata tietoja ja dataa salauksella ja autentikoinnilla. (Aljuhani, 2021; Khanna ym., 2018; Pavithra & Femilda Josephin, 2020; Wiafe ym., 2020.)

Kuitenkin yksi merkittävimmistä ja tärkeimmistä tekoälyn sovelluskohteista näyttäisi kirjallisuuden perusteella olevan haittaohjelmien torjunta ja luokittelu, joissa käytetään yleensä koneoppimismenetelmiä. Esimerkiksi kahta tunnettua tekoälyalgoritmia RandomForest (RF) ja DecisionTree (DT) on käytetty hyvin menestyksekkäästi haittaohjelmien havaitsemisessa ja luokittelussa, joista RF-algoritmia käyttämällä on päästy jopa 99,4 % tarkkuuteen haittaohjelmien luokittelussa. (Debar ym., 2000; Fei ym., 2022; Pavithra & Femilda Josephin, 2020.) Toisaalta tietoturvan toteutumisen näkökulmasta voidaan ajatella tällaisen tarkkuuden olevan sinänsä riittämätöntä, mikäli suojaus kestää esimerkiksi 199 ensimmäistä kertaa, mutta 200. kerralla suojaus pettää ja tietoturvauhka toteutuu.

Edellä mainittujen sovelluskohteiden lisäksi tekoälysovelluksia on kuitenkin useita erilaisia ja niitä kehitetään jatkuvasti lisää, ja eri teknologioiden tehokkuus ja tarkkuus vaihteleekin eri tehtävissä. Useat tutkimukset ovat vertailleet eri tekoälysovellusten suorituskykyä ja osoittaneet, että jokaisella niistä on omat vahvuutensa ja heikkoutensa. Esimerkiksi edellä mainitut DT- ja RF-algoritmit ovat osoittaneet erittäin hyvää suorituskykyä nimenomaan haittaohjelmien havaitsemisessa ja luokittelussa, kun taas CNN-algoritmit ovat osoittautuneet hyviksi kuvien ja kuvioiden tunnistuksen saralla. (Alkahtani & Aldhyani, 2022; Chellappan ym., 2018; Pavithra & Femilda Josephin, 2020; Sutskever ym., 2017.)

4.3.1 Tekoälyn sovelluskohteet jatkuvassa valvonnassa

Tietojärjestelmissä ja -verkoissa tapahtuukin jatkuvasti monenlaista liikennettä ja aktiviteettia, eikä kaikkea voida valvoa manuaalisesti. Mosteanun (2020) mukaan verkkohyökkäysyritysten maantieteellinen kasvu kuormittaa IT-alan ammattilaisia, jotka joutuvat lähes päivittäin tekemisiin hakkerointiyritysten lisääntymisen kanssa. Tämän vuoksi organisaatiot käyttävät hänen mukaansa jatkuvan valvonnan ja uhkien havaitsemisen työkaluja, jotka pystyvät automaattisesti havaitsemaan epätavallisia aktiviteetteja verkoissa ja järjestelmissä, sekä hälyttämään niistä vastuuhenkilöitä. Mosteanun (2020) toteuttaman kyselyn mukaan 44 % kyselyyn vastanneista globaaleista yrityksistä käyttää jo tai ovat halukkaita käyttämään tekoälyä IT-osastoillaan tietoturvahäiriöiden seuraamiseen, havainnointiin ja estämiseen. Hänen mukaansa organisaatioilla onkin kasvava halu helpottaa tietoturvaan liittyviä työtehtäviä tekevän henkilöstön työkuormaa lisääntyneiden kyberuhkien ja -hyökkäysten vuoksi. (Mosteanu, 2020.)

Perinteiset työkalut tällä alueella seuraavat verkkoliikennettä ja tietojärjestelmien lokitietoja ja analysoivat niitä automaattisesti. Mikäli ne havaitsevat epätavallista toimintaa, kuten luvattomia sisäänkirjautumisyrityksiä tai tietojen poistoyrityksiä, ne voivat hälyttää organisaation vastuuhenkilöitä ja käynnistää automaattisia toimenpiteitä, kuten käyttäjätilin sulkemisen tai tietojen varmuuskopioinnin. Näillä toimilla pyritään ylläpitämään organisaation sisäinen järjestys ja tietoturvan toteutuminen, eli varmistetaan tiedon muuttumattomuus, eheys ja vain oikeutettujen henkilöiden pääsy tietoon. (Aljuhani, 2021; Khorshed ym., 2012; Lehto, 2022; Mosteanu, 2020; Vuorinen & Tetri, 2012.)

Jatkuvan valvonnan ja uhkien havaitsemisen työkalut voivat luonnollisesti sisältää myös tietynlaisia tekoälyalgoritmeja, jotka pystyvät tunnistamaan entistä monimutkaisempia uhkia. Koneoppiminen voi esimerkiksi auttaa tunnistamaan

tuntemattomia haittaohjelmia tai käyttäytymismalleja, jotka poikkeavat normaalista. Teollisuudessa Khorshedin ym. (2012) mukaan juuri tällaisten työntekijöiden käyttäytymistä seuraavista jatkuvan valvonnan tekoälyratkaisuista voi olla suuri hyöty poikkeavan toiminnan havaitsemisessa, kuten epätavallisten tiedostojen avaamisessa tai tiedostonsiirrosta epäilyttävistä lähteistä, mikä voi olla heidän mukaansa merkki esimerkiksi tietomurrosta tai haittaohjelmien leviämisestä yrityksen verkkoon (Khorshed ym., 2012). Tällaiset työkalut ovat hyvin tärkeitä kyberuhkien torjunnassa, sillä ne mahdollistavat entistä nopeamman reagoinnin uhkiin ja haittojen leviämisen minimoimisen. Lisäksi näiden sovellutusten avulla organisaatioiden nähdään pystyvän keräämään tietoa ja oppimaan aiemmista hyökkäyksistä ja sitä kautta mahdollistamaan paremman valmiuden ehkäistä tulevia hyökkäyksiä. (Ahmed ym., 2019; Pourjavan, 2019; Wiafe ym., 2020.)

Tämän tutkielman lähdeaineiston perusteella voidaankin korostaa tekoälyn suurta hyötyä erityisesti jatkuvassa valvonnassa. Aineiston perusteella sitä käytetäänkin laajasti ennen kaikkea tietomurtojen ja haittaohjelmien havainnointiin, potentiaalisten uhkien ennakoimiseen ja haavoittuvuuksien hallintaan, sillä tekoäly voidaan valjastaa tutkimaan myös omien järjestelmien haavoittuvuuksia, sekä ehdottamaan korjaustoimenpiteitä ennen mahdollista hyökkäystä.

4.3.2 Case-esimerkkejä tekoälyn käytöstä jatkuvassa valvonnassa

Maailmalla näyttää olevan käytössä useita tekoälyteknologioihin perustuvia tietoturvaohjelmistoja, joita käytetään nimenomaan jatkuvassa valvonnassa. Esimerkkejä käytössä olevista sovelluksista ovat muun muassa Darktrace, Cylance ja McAfeen:n tuottama palvelu.

Darktrace on tekoälypohjainen järjestelmä, jonka sanotaan käyttävän koneoppimisen menetelmää havaitakseen ja torjuakseen erilaisia kyberuhkia. Sen keskeisiä ominaisuuksia näyttää markkinointimateriaalin perusteella olevan itseoppiva käyttäytyminen, joka mahdollistaa kyberhyökkäysten nopean havaitsemisen, sekä kyky sopeutua uusiin uhkiin. Darktracea käyttävät monet tunnetut organisaatiot, kuten Airbus, BillaBong, Las Vegasin kaupunki ja McLaren Group. (*Darktrace | Cyber Security That Learns You*, ei pvm.; *Darktrace Customer Stories*, ei pvm.)

Cylance on myös tekoälypohjainen tietoturvaohjelmisto, jonka kerrotaan käyttävän hyödyksi koneoppimista, mutta myös luonnollisen kielen prosessointia (NLP) haittaohjelmien ja muiden kyberuhkien tunnistamisessa. Cylance kykenee tunnistamaan uusia ja kehittyviä uhkia, mikä heidän markkinointimateriaalinsa mukaan tekee siitä erittäin tehokkaan jatkuvan valvonnan alueella. (*Cylance AI from BlackBerry*, ei pvm.)

Myös McAfeen tekoälypohjaiset ratkaisut käyttävät koneoppimista ja käyttäytymisanalytiikkaa uhkien havainnointiin. Tämän lisäksi kyseisen teknologian mainitaan käyttävän tietoa verkossa tapahtuvasta toiminnasta uhkien havaitsemiseksi ennen niiden toteutumista. McAfee väittää suojanneensa tuotteillaan jopa yli 500 miljoonaa laitetta globaalisti. (*Antivirus, VPN, Identity & Privacy Protection*, ei pvm.)

Riippumatonta tutkimusta näyttäisi kuitenkin olevan hankala löytää eri yhtiöiden tietoturvaratkaisuiden osalta, joten objektiivista käsitystä tuotteiden

toimivuudesta tai tehokkuudesta ei voida muodostaa. Näiden perusteella voidaan kuitenkin todeta, että tällaisia tekoälypohjaisia teknologioita on olemassa ja niitä käytetään liiketoiminnan ja kyberympäristön suojaamiseksi jatkuvassa valvonnassa.

4.4 Tekoälyn käytön vaikutus kyberturvallisuusammattilaisten rooleihin

Kokonaisuutena tekoälyn tuloa osaksi kyberturvallisuuden ratkaisuja tarkasteltaessa voidaan tieteellisessä kirjallisuudessa havaita muutamia kasvavia trendejä ja lieveilmiöitä. Suurimpana näistä varmastikin näyttäytyy ihmisen rooli kyberturvallisuuden toteuttamisessa.

Jatkuva toiminnan tehostamisen tarve näyttää kasvattavan tarvetta automatisoida prosesseja ja toimintoja, ja samalla yhä monimutkaisemmat uhkakuvat vaativat yhä monimutkaisempia suojautumiskeinoja. Tekoäly näytteleekin tässä isoa roolia, ja sen suhde ihmisen suorittamaan työhön näyttäytyy mielenkiintoisella tavalla ristiriitaisena jo julkisessa keskustelussakin. Esimerkiksi tekoäly-yhtiö OpenAI:n tekemässä tutkimuksessa selvitettiin LLM-mallien, eli kielimallien vaikutusta työmarkkinoihin ja siinä havaittiin kyseisen tekoälymenetelmän vaikuttavan suoraan erityisesti toistuviin työtehtäviin, mutta muun muassa myös sellaisiin työtehtäviin, joissa vaaditaan esimerkiksi analytiikkaa, matematiikkaa, ohjelmointia ja kirjoittamista. (Culot ym., 2019; Eloundou ym., 2023; Gonçalves ym., 2022; Mosteanu, 2020; Vartolomei & Avasilcăi, 2020.)

Tekoälyä onkin pitkään kehitetty suoriutumaan tehtävistä, joita ihmiset ovat perinteisesti tehneet ja se pystyy suoriutumaan monista tehtävistä jo ihmistä paremmin. Tästä onkin syntynyt pelko siitä, että viekö tekoäly ihmisten työpaikat. Toisaalta taas nopea kehitys ajaa monia alueita siihen, että ilman tekoälyä ei voida vastata kilpailuun ja suoriutumaan vaaditulla tavalla erinäisistä tehtävistä, kuten kyberturvallisuuden ylläpidosta uudenlaisia uhkia vastaan. (Goertzel, 2014; Hamet & Tremblay, 2017; Ozkan-Ozen & Kazancoglu, 2021.)

Näin ollen on mielenkiintoista tutkia myös ihmisen roolin kehitystä tekoälyn käyttöönoton seurauksena kyberturvallisuuden osalta, joten tämän tutkielman lähdekirjallisuuden pohjalta on kuvattu taulukossa (taulukko 3) eri kyberturvallisuuden toteuttamisen ulottuvuuksien suhteen ihmisen roolin mahdollisia muutoksia. Taulukossa jaotellaan tehtäviä ulottuvuuden kautta sen mukaan, miten ihmisen rooli voi heikentyä tekoälyn käyttöönoton yhteydessä, mutta toisaalta myös sen mukaan miten ihmisen rooli voi paradoksaalisesti vahvistua. Kokonaisuudessaan kuitenkin esimerkiksi Högmanderin (2019) mukaan näyttää siltä, että ihmistä vielä tarvitaan erityisesti kyberturvallisuuden strategian toteuttamisessa ja haastavimpien asiantuntijatehtävien hoitamisessa. Toisaalta hänen mukaansa tekoäly helpottaa huomattavasti työkuormaa toistuvien ja valvovien tehtävien osalta. (Högmander, 2019.)

TAULUKKO 3 Ihmisen roolin muutos tekoälyn käyttöönoton yhteydessä

Ulottuvuus	Ihmisen roolin heikentyminen	Ihmisen roolin vahvistuminen
Hälytykset	Hälytysten automatisointi vähentää ihmisten tarvetta valvoa tietoturvaa manuaalisesti (Eloundou ym., 2023; Mosteanu, 2020; Vartolomei & Avasilcăi, 2020)	Ihmisen rooli vahvistuu, jos järjestelmä luo liian paljon hälytyksiä ja ihmisten täytyy arvioida, ovatko ne todellisia uhkia (Ahmed ym., 2019; Högmander, 2019)
Analyysi	Analyysin automatisointi vähentää ihmisten tarvetta tarkastella yksittäisiä tietoturvahälytyksiä manuaalisesti (Eloundou ym., 2023; Gonçalves ym., 2022; Mosteanu, 2020; Vartolomei & Avasilcăi, 2020)	Ihmisen rooli vahvistuu, jos järjestelmä ei kykene tunnistamaan uusia uhkia ja ihmiset joutuvat tekemään manuaalista analyysiä (Ahmed ym., 2019; Högmander, 2019; Mosteanu, 2020)
Reagointi	Automatisoidut järjestelmät voivat nopeasti reagoida uhkiin, mikä vähentää ihmisten tarvetta toimia nopeasti (Eloundou ym., 2023; Mosteanu, 2020; Vartolomei & Avasilcăi, 2020)	Ihmisen rooli vahvistuu, kun automatisoidut järjestelmät eivät pysty reagoimaan tiettyihin uhkiin ja ihmiset joutuvat tekemään manuaalista reagointia (Ahmed ym., 2019; Högmander, 2019; Mosteanu, 2020)
Opetus	Tekoäly voi oppia tunnistamaan uusia uhkia ilman ihmisten apua (Culot ym., 2019; Mosteanu, 2020; Wiafe ym., 2020)	Ihmisen rooli vahvistuu, kun ihmiset ovat vastuussa tekoälyn opettamisesta ja varmistavat sen oikeellisuuden (Ahmed ym., 2019; Alkahtani & Aldhyani, 2022; Mosteanu, 2020)
Suunnittelu	Automatisointi voi helpottaa tietoturvan suunnittelua ja vähentää ihmisten tarvetta suunnitella manuaalisesti (Eloundou ym., 2023; Mosteanu, 2020; Zeadally ym., 2020)	Ihmisen rooli vahvistuu, kun suunnitteluprosessissa tarvitaan ihmisten asiantuntemusta ja kokemusta (Högmander, 2019; Mosteanu, 2020; Zeadally ym., 2020)
Testaus	Automatisointi voi vähentää ihmisten tarvetta testata tietoturvaratkaisuja manuaalisesti (Mosteanu, 2020; Wiafe ym., 2020; Zeadally ym., 2020)	Ihmisen rooli vahvistuu, kun tietoturvaratkaisuja täytyy testata monimutkaisissa ympäristöissä, joissa automatisointi ei riitä (Ahmed ym., 2019; Högmander, 2019; Mosteanu, 2020; Zeadally ym., 2020)

5 YHTEENVETO

Tämän tutkielman tarkoituksena oli tutkia tekoälyn käyttöä ja sen mahdollisuuksia kyberturvallisuudessa, erityisesti kyberuhkien jatkuvan valvonnan osalta. Tutkielma toteutettiin kirjallisuuskatsauksena, jossa lähteenä käytettiin pääosin tieteellisiä artikkeleita ja raportteja, mutta myös alan oppikirjoja, verkkosivustoja ja uutislähteitä.

Tutkielman luvussa 2 määriteltiin tekoälyn käsite ja kuvailtiin lähdeaineistoon pohjautuen tekoälyyn liittyviä menetelmiä tutkielman aiheen kannalta olennaisin osin. Luvussa 3 puolestaan määriteltiin kyberturvallisuuden käsite ja käsiteltiin turvan toteutumista, kyberuhkia ja niiden torjuntaa, jatkuvaa valvontaa ja sen merkitystä kyberturvallisuudelle. Lisäksi kolmannessa luvussa tutustuttiin lyhyesti kyberturvallisuuden ominaispiirteisiin teollisuussektorilla. Luvussa 4 paneuduttiin tekoälyn käyttöön kyberturvallisuudessa siihen liittyvien hyötyjen ja haasteiden kautta. Neljännessä luvussa käsiteltiin myös kyberuhkien torjunnassa käytettäviä tekoälyratkaisuja case-esimerkkien kautta, sen sovelluskohteita jatkuvassa valvonnassa ja tekoälyn käyttöönoton mahdollisia vaikutuksia kyberturvallisuusammattilaisten rooleihin.

Tälle kirjallisuuskatsaukselle asetettiin aluksi kaksi tutkimuskysymystä, joihin pyrittiin löytämään vastauksia. Nämä kysymykset olivat:

- Miten tekoälyä voidaan hyödyntää kyberturvallisuuden parantamisessa jatkuvassa valvonnassa?
- Miten erilaiset tekoälysovellukset vertautuvat toisiinsa kyberturvallisuudessa?

Kirjallisuuskatsauksen perusteella voidaan todeta, että tekoälyn käytöllä kyberturvallisuudessa nähdään erittäin suuri potentiaali erityisesti jatkuvan valvonnan alueella, sekä sen kyvyssä reagoida nopeasti erilaisiin uhkiin kyberympäristössä. Verrattuna perinteisiin kyberuhkien torjuntamenetelmiin tekoälysovellutukset kykenevät parhaimmillaan tarkempaan ja nopeampaan uhkien havainnointiin, luokitteluun ja torjuntaan reaaliajassa. Tekoälyn myötä erilaiset jatkuvan valvonnan työkalut kykenevätkin oppimaan ja sopeutumaan uusiin ja muuttuviin uhkiin ilman tarvetta erillisille ohjelmistopäivityksille. Tekoäly kykenee

myös huomattavan suurten tietomäärien nopeaan analysointiin, joka mahdollistaa esimerkiksi anomalioiden tai muun epäilyttävän toiminnan havaitsemisen esimerkiksi tietoverkoissa tai organisaatioiden järjestelmissä ennen mahdollisen tietoturvariskin toteutumista.

Tekoälyn käyttöön kyberturvallisuudessa liittyy katsauksen perusteella kuitenkin myös huomioon otettavia haasteita. Esimerkiksi käytettävän tekoälyalgoritmin opetukseen käytettävän materiaalin laatuun on kiinnitettävä erityistä huomiota, sillä lähdeaineiston mukaan useat tekoälyn opetuspaketit voidaan luokitella nykypäivänä jo vanhentuneiksi. Tämä on myös uudenlainen tietoturvariski itsessäänkin, mutta lähdekirjallisuuden perusteella on toisaalta olemassa myös tekoälyalgoritmeja, jotka hyökkäävät juurikin tekoälyn opetusmateriaalia vastaan aiheuttaen vakaviakin mahdollisia seurauksia myöhemmässä käytön vaiheessa. Lisäksi on osattava valita ja säätää oikeanlainen tekoälyalgoritmi riittävän tarkasti kuhunkin käyttötarkoitukseen. Uuden teknologian pääasiallisena tarkoituksena näyttää olevan automatisoida toimintaa myös kyberturvallisuudessa, mutta esimerkiksi liian tarkkojen algoritmien käyttö voi paradoksaalisesti lisätä henkilöstön työkuormaa ylimääräisten hälytysten eli ”red flagien” tarkastuksessa.

Toisin sanoen tekoälysovellukset kyberturvallisuudessa voivat kestää ja sopeutua paremmin ennakoimattomiin ja uusiin kyberuhkiin kuin perinteiset menetelmät, jotka ovat näiltä osin riippuvaisia uusista ohjelmistopäivityksistä. Lisäksi tekoälysovellukset voivat havaita ja reagoida paremmin CIA-järjestystä uhkaaviin elementteihin esiohjelmoituja järjestelmiä paremmin, mutta toisaalta niihin sisältyy riski toimia itse tätä CIA-järjestystä vastaan. Mikäli tekoäly tarttuu liian tiukasti poikkeavuuksiin ja aiheuttaa ylimääräisiä hälytyksiä tai ryhtyy toimimaan normaaleja toimintoja vastaan, voidaan sen sanoa Vuorisen (2012, 2016) mukaan olevan yksi tietoturvaamisen paradoksi (Vuorinen & Tetri, 2012, 2016).

Eri tekoälysovelluksilla voidaankin sanoa katsauksen perusteella olevan eroja keskenään, vaikkakin eri tekoälymenetelmien hyödyntäminen ja niiden toiminta voi olla hyvinkin päällekkäistä tai samankaltaista. Nykyiset heikon tekoälyn sovellukset näyttävät kuitenkin olevan kykeneviä osittain vain yksittäisiin tehtäviin, joten eri sovelluksilla on erityyppisiä vaikutuksia muun muassa kyberturvallisuusammattilaisten työhön. Pääsääntöisesti tämän katsauksen perusteella trendi vaikuttaisi suuntauttavan ihmisen suorittamia työtehtäviä kohti haastavampia strategisia, suunnitteluun ja tekoälyn opettamiseen liittyviä työtehtäviä. Näitä trendejä ja tässä luvussa aiemmin käsiteltyjä tutkielman tuloksia on pyritty nostamaan alla olevaan taulukkoon (taulukko 4), jonka tarkoituksena on auttaa muodostamaan yhtenäisempi ja selkeämpi kuva tutkielman pääasiallisista tuloksista.

Kokonaisuudessaan voidaan sanoa tämän tutkielman olleen kohtuullisen onnistunut, sillä lähdeaineiston analyysissä löydettiin vastauksia asetettuihin tutkimuskysymyksiin. Pohdittaessa tutkielman tuloksia on kuitenkin syytä huomioida tutkielman tekemiseen liittyvät rajoitteet.

Tämän tutkielman haasteina ja rajoitteina voidaan pitää käytetyn aineiston rajallisuutta ja kuvailevan kirjallisuuskatsauksen käyttöä tutkimusmenetelmänä. Tekoälyn soveltamista kyberturvallisuudessa käsittelevä tutkimus on vielä suhteellisen uutta, joten tietoa ja kokemusta sen käytöstä näyttää olevan rajoitetusti,

minkä lisäksi alaa leimaa jatkuva muutos. Osa aihetta koskevasta tutkimuksesta ja kirjallisuudesta on lisäksi maksumuurin takana, mikä rajoitti joihinkin artikkeleihin pääsyä tämän tutkielman aineiston keruuprosessissa. Tämän tutkielman kannalta tämä seikka rajasi karkeasti noin kymmenen aihetta koskevaa potentiaalista lähdettä pois käytetystä aineistosta. On myös syytä ottaa huomioon esimerkiksi tekoälyä kehittävien ja käyttävien organisaatioiden ja yritysten mahdollinen haluttomuus antaa kattavaa tietoa kyberhyökkäyksiin ja niiden torjumiseen kehitettyjen ja käytössä oleviin tai olemattomiin sovelluksiin liittyen.

TAULUKKO 4 Tutkielman pääasialliset löydökset

Aihe	**Pääasialliset löydökset**
Tekoälyn hyödyt	Tehokkaampi kyberuhkien havainnointi ja torjunta
	Jatkuvan valvonnan alueella parempi poikkeamien havaitseminen ja hyökkäysten ennaltaehkäisy verrattuna perinteisiin menetelmiin
	Tehokkaampi vastaaminen kehittyneisiin kyberuhkiin, ei yhtä riippuvainen päivityksistä kuin perinteiset ratkaisut
Tekoälyn haasteet	CIA-järjestykseen, eettisiin kysymyksiin ja luotettavuuteen liittyvät haasteet
	Ei voida vielä täysin luottaa tekoälyyn yksinään kyberturvallisuuden päätöksenteossa
Tekoälyn sovellukset	Potentiaalia turvallisuuskriittisten järjestelmien valvonnassa ja uhkien havainnoinnissa
	Tehokkaammat keinot kyberuhkien torjumiseen
	Useita sovelluksia jo käytössä
Tekoälyn vaikutukset	Kyberturvallisuusammattilaisten rooliin odotettavissa muutoksia
	Nopeampi reagointi ja parannettu kyberuhkien torjunta, helpottaa ihmisten työkuormaa

Tekoälyn käyttöön kyberturvallisuudessa liittyvä tutkimus on joka tapauksessa tärkeää ja ajankohtaista, joten jatkotutkimukselle näyttää nopean kehityksen myötä olevan tarvetta tulevaisuudessakin. Tämän tutkielman pohjalta jatkotutkimusehdotuksena olisi toteuttaa empiirinen tutkimus tekoälyn kehittyneistä menetelmistä kyberturvallisuuden jatkuvassa valvonnassa ja toteuttaa se

eräänlaisena vertailevana analyysina algoritmeista ja niiden soveltuvuudesta kyberuhkien havainnointiin ja torjuntaan. Tutkimuksessa voitaisiin vertailla esimerkiksi syväoppimisen, neuroverkkojen ja luokittelumenetelmien algoritmeja, sekä arvioida niiden soveltuvuutta kyberuhkien havainnointiin ja torjuntaan. Tämän avulla saataisiin lisää ymmärrystä siitä, mitkä algoritmit ovat tehokkaimpia erityisesti jatkuvan valvonnan kannalta ja millaisia tietoja ja resursseja niiden käyttö edellyttäisi käytännössä. Lisäksi olisi mielenkiintoista tutkia lisää ihmisen ja tekoälyn yhteistyön kehitystä kyberturvallisuuden jatkuvassa valvonnassa.

LÄHTEET

Adams, D. (23.4.2010). *McAfee Identifies Core Windows File as Malicious – OSnews*. https://www.osnews.com/story/23196/mcafee-identifies-core-windows-file-as-malicious/

Ahmed, A., Krishnan, V. V. G., Foroutan, S. A., Touhiduzzaman, Md., Rublein, C., Srivastava, A., Wu, Y., Hahn, A. & Suresh, S. (2019). Cyber Physical Security Analytics for Anomalies in Transmission Protection Systems. *IEEE Transactions on Industry Applications*, *55*(6), 6313–6323. https://doi.org/10.1109/TIA.2019.2928500

Aljuhani, A. (2021). Machine Learning Approaches for Combating Distributed Denial of Service Attacks in Modern Networking Environments. *IEEE Access*, *9*, 42236–42264. https://doi.org/10.1109/ACCESS.2021.3062909

Alkahtani, H. & Aldhyani, T. H. H. (2022). Artificial Intelligence Algorithms for Malware Detection in Android-Operated Mobile Devices. *Sensors*, *22*(6), 2268. https://doi.org/10.3390/s22062268

Ansari, M. F., Dash, B., Sharma, P. & Yathiraju, N. (2022). The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review. *IJARCCE*, *11*(9). https://doi.org/10.17148/IJARCCE.2022.11912

Antivirus, VPN, Identity & Privacy Protection. (ei pvm.). McAfee. Noudettu 11. maaliskuuta 2023, osoitteesta https://www.mcafee.com/

August, T., Dao, D. & Niculescu, M. F. (2022). Economics of Ransomware: Risk Interdependence and Large-Scale Attacks. *Management Science*, *68*(12), 8979–9002. https://doi.org/10.1287/mnsc.2022.4300

Bécue, A., Praça, I. & Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: challenges and opportunities. *Artificial Intelligence Review*, *54*(5), 3849–3886. https://doi.org/10.1007/s10462-020-09942-2

Bello, I., Chiroma, H., Abdullahi, U. A., Gital, A. Y., Jauro, F., Khan, A., Okesola, J. O. & Abdulhamid, S. M. (2021). Detecting ransomware attacks using intelligent algorithms: recent development and next direction from deep learning and big data perspectives. *Journal of Ambient Intelligence and Humanized Computing*, *12*(9), 8699–8717. https://doi.org/10.1007/s12652-020-02630-7

Chellappan, S., Cheng, W. & Li, W. (toim.). (2018). *Wireless Algorithms, Systems, and Applications* (Vsk. 10874). Springer International Publishing. https://doi.org/10.1007/978-3-319-94268-1

Craigen, D., Diakun-Thibault, N. & Purse, R. (2014). Defining Cybersecurity. *Technology Innovation Management Review*.

Culot, G., Fattori, F., Podrecca, M. & Sartor, M. (2019). Addressing Industry 4.0 Cybersecurity Challenges. *IEEE Engineering Management Review*, *47*(3), 79–86. https://doi.org/10.1109/EMR.2019.2927559

Cyber war: the next threat to national security and what to do about it. (2011). *Choice Reviews Online*, *48*(05), 48-2963-48–2963. https://doi.org/10.5860/CHOICE.48-2963

Cyber-attack on Hydro. (2020). https://www.hydro.com/en/media/on-the-agenda/cyber-attack/

Cybersecurity and Cyberwar : What Everyone Needs to Know®. (ei pvm.). Noudettu 10. maaliskuuta 2023, osoitteesta https://web-s-ebscohost-com.ezproxy.jyu.fi/ehost/ebookviewer/ebook/ZTAwMHh3d19fNjU3NjI5X19BTg2?sid=617accb7-f0f1-4e66-9dab-1e48eed05750@redis&vid=0&format=EB&rid=1

Cybersecurity guide for SMEs - 12 steps to securing your business. (2021). [Report/Study]. ENISA. https://www.enisa.europa.eu/publications/cybersecurity-guide-for-smes

Cylance AI from BlackBerry. (ei pvm.). Noudettu 11. maaliskuuta 2023, osoitteesta https://www.blackberry.com/us/en/products/cylance-endpoint-security/cylance-ai

Darktrace | Cyber security that learns you. (ei pvm.). Noudettu 11. maaliskuuta 2023, osoitteesta https://darktrace.com/

Debar, H., Mé, L. & Wu, S. F. (toim.). (2000). *Recent advances in intrusion detection: third international workshop, RAID 2000, Toulouse, France, October 2-4, 2000: proceedings*. Springer.

Dong, Y., Zhu, P., Liu, Q., Chen, Y. & Xun, P. (ei pvm.). *Degrading Detection Performance of Wireless IDSs Through Poisoning Feature Selection*.

Dua, S. & Du, X. (2016). *Data Mining and Machine Learning in Cybersecurity*. CRC Press.

Eloundou, T., Manning, S., Mishkin, P. & Rock, D. (2023). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models* (arXiv:2303.10130). arXiv. http://arxiv.org/abs/2303.10130

Fei, N., Lu, Z., Gao, Y., Yang, G., Huo, Y., Wen, J., Lu, H., Song, R., Gao, X., Xiang, T., Sun, H. & Wen, J.-R. (2022). Towards artificial general intelligence via a multimodal foundation model. *Nature Communications*, *13*(1), 3094. https://doi.org/10.1038/s41467-022-30761-2

Fischer, E. A. (2016). *Cybersecurity Issues and Challenges: In Brief*.

Gajek, S., Lees, M. & Jansen, C. (2021). IIoT and cyber-resilience: Could blockchain have thwarted the Stuxnet attack? *AI & SOCIETY*, *36*(3), 725–735. https://doi.org/10.1007/s00146-020-01023-w

Goertzel, T. (2014). The path to more general artificial intelligence. *Journal of Experimental & Theoretical Artificial Intelligence*, *26*(3), 343–354. https://doi.org/10.1080/0952813X.2014.895106

Gonçalves, M. J. A., Link to external site, this link will open in a new window, Silva, A. C. F. da & Ferreira, C. G. (2022). The Future of Accounting: How Will Digital Transformation Impact the Sector? *Informatics*, *9*(1), 19. https://doi.org/10.3390/informatics9010019

Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning*. MIT Press.

Hamet, P. & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism*, *69*, S36–S40. https://doi.org/10.1016/j.metabol.2017.01.011

Heim, T. N. & Wessel, R. A. (2023). The Various Dimensions of Cyberthreats: (In)consistencies in the Global Regulation of Cybersecurity. *Anales de Derecho*, *40*, 39–65.

HS Järvenpää | Voimakas kyberhyökkäys pysäytti oppilaitoksen järjestelmät lähes kuukauden ajaksi. (10.3.2023). Helsingin Sanomat. https://www.hs.fi/kaupunki/jarvenpaa/art-2000009445588.html

Högmander, J. (11.1.2019). *Tekoäly on mullistanut tietoturvan – ihmistäkin tarvitaan vielä*. F-Secure Blog. https://blog.f-secure.com/fi/tekoaly-mullistanut-tietoturvan-ihmistakin-tarvitaan-viela/

Khanna, K., Panigrahi, B. K. & Joshi, A. (2018). AI-based approach to identify compromised meters in data integrity attacks on smart grid. *IET Generation, Transmission & Distribution*, *12*(5), 1052–1066. https://doi.org/10.1049/iet-gtd.2017.0455

Khorshed, Md. T., Ali, A. B. M. S. & Wasimi, S. A. (2012). A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. *Future Generation Computer Systems*, *28*(6), 833–851. https://doi.org/10.1016/j.future.2012.01.006

Kyberhyökkäys pysäytti tuotannon Valtran tehtaalla – kiristyshaittaohjelmat yleistyvät, yritykset varautuvat yhä paremmin. (10.5.2022). Yle Uutiset. https://yle.fi/a/3-12439169

Kyberturvallisuus | Suomeen tehdään nyt aiempaa vaarallisempia kyberhyökkäyksiä. (25.10.2022). Helsingin Sanomat. https://www.hs.fi/kotimaa/art-2000009150023.html

Kyberturvallisuuskeskuksen viikkokatsaus - 10/2023 | Kyberturvallisuuskeskus. (1.3.2023). https://www.kyberturvallisuuskeskus.fi/fi/ajankohtaista/kyberturvallisuuskeskuksen-viikkokatsaus-102023

Lehto, M. (2022). *View of APT Cyber-attack Modelling: Building a General Model*. https://papers.academic-conferences.org/index.php/iccws/article/view/36/18

Lezzi, M., Lazoi, M. & Corallo, A. (2018). Cybersecurity for Industry 4.0 in the current literature: A reference framework. *Computers in Industry*, *103*, 97–110. https://doi.org/10.1016/j.compind.2018.09.004

Lunnaiden maksu rikollisille voi olla välttämätöntä tietomurroissa – tietoturva-asiantuntija: Toinen vaihtoehto on konkurssi. (12.8.2022). Yle Uutiset. https://yle.fi/a/3-12573399

Mattila, J., Ali-Yrkkö, J. & Seppälä, T. (2020). *Kyberuhat yleistyvät – Miten Suomen yritykset pärjäävät? 93*.

McLaren - Darktrace Customer Stories. (ei pvm.). Noudettu 11. maaliskuuta 2023, osoitteesta https://darktrace.com/customers/mclaren

Mosteanu, N. R. (2020). Artificial Intelligence and Cyber Security – A Shield against Cyberattack as a Risk Business Management Tool – Case of European Countries. *Calitatea: Acces La Success*, *21*(175), 148–156.

National Institute of Standards and Technology. (2018). *Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1* (NIST CSWP 04162018; s. NIST CSWP 04162018). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.CSWP.04162018

Nikiforakis, N., Invernizzi, L., Kapravelos, A., Van Acker, S., Joosen, W., Kruegel, C., Piessens, F. & Vigna, G. (2012). You are what you include: large-scale evaluation of remote javascript inclusions. *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, 736–747. https://doi.org/10.1145/2382196.2382274

Ozkan-Ozen, Y. D. & Kazancoglu, Y. (2021). Analysing workforce development challenges in the Industry 4.0. *International Journal of Manpower*, *43*(2), 310–333. https://doi.org/10.1108/IJM-03-2021-0167

Ozkaya, E. (2019). *Cybersecurity: the Beginner's Guide: A Comprehensive Guide to Getting Started in Cybersecurity*. Packt Publishing, Limited. http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=5781046

Padallan, J. O. (2019). *Cyber Security*. https://web-p-ebscohost-com.ezproxy.jyu.fi/ehost/ebookviewer/ebook/ZTAwMHh3d19fMjMyNDMyN19fQU41?sid=731bf455-d1ab-4356-b54b-50168d2cc9bb@redis&vid=0&format=EB&rid=1

Parisi, A. (2019). *Hands-On Artificial Intelligence for Cybersecurity: Implement Smart AI Systems for Preventing Cyber Attacks and Detecting Threats and Network Anomalies*. Packt Publishing, Limited. http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=5847212

Pavithra, J. & Femilda Josephin, J. S. (2020). Analyzing Various Machine Learning Algorithms for the Classification of Malwares. *IOP Conference Series: Materials Science and Engineering*, *993*(1), 012099. https://doi.org/10.1088/1757-899X/993/1/012099

Pourjavan, S. (2019). Definitions: Machine learning, deep leerning and AI understanding. *Acta Ophthalmologica*, *97*(S263). https://doi.org/10.1111/j.1755-3768.2019.8214

Russell, S. J. & Norvig, P. (2010). *Artificial Intelligence a Modern Approach, Third Edition*. https://scholar.alaqsa.edu.ps/9195/1/Artificial%20Intelligence%20A%20Modern%20Approach%20%283rd%20Edition%29.pdf%20%28%20PDF Drive%20%29.pdf

Soni, V. D. (2020). Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3624487

Sutskever, I., Hinton, G. E. & Krizhesky, A. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Tang, B., Wang, J., Qiu, H., Yu, J., Yu, Z. & Liu, S. (2023). Attack Behavior Extraction Based on Heterogeneous Cyberthreat Intelligence and Graph Convolutional Networks. *Computers, Materials & Continua*, *74*(1), 235–252. https://doi.org/10.32604/cmc.2023.029135

Tavaroista palveluihin – Suomen teollisuus rakennemuutoksen pyörteissä. (9.11.2017). Euro ja talous. https://www.eurojatalous.fi/fi/2017/artikkelit/tavaroista-palveluihin-suomen-teollisuus-rakennemuutoksen-pyorteissa/

Triton is the world's most murderous malware, and it's spreading. (ei pvm.). MIT Technology Review. Noudettu 12. maaliskuuta 2023, osoitteesta https://www.technologyreview.com/2019/03/05/103328/cybersecurity-critical-infrastructure-triton-malware/

Uma, M. & Padmavathi, G. (2013). *A Survey on Various Cyber Attacks and Their Classification*.

Vartolomei, C. & Avasilcăi, S. (2020). Digitalization concept: Cyber-risks and damages for companies in adhered industries. *IOP Conference Series: Materials Science and Engineering*, *898*(1), 012044. https://doi.org/10.1088/1757-899X/898/1/012044

Vastaamon tietovuoto | Ex-toimitusjohtaja tietoturva-aukosta oikeudessa: "Vaikea keksiä järkevää selitystä". (10.3.2023). Helsingin Sanomat. https://www.hs.fi/kotimaa/art-2000009440043.html

von Solms, R. & van Niekerk, J. (2013). From information security to cyber security. *Computers & Security*, *38*, 97–102. https://doi.org/10.1016/j.cose.2013.04.004

Voutilainen, T. (31.3.2023). *Informaatio ja tietotekniikkaoikeus, Luento 5*. https://moodle.jyu.fi/pluginfile.php/1267875/mod_resource/content/0/Informaatio-%20ja%20tietotekniikkaoikeus%20Jyv%C3%A4skyl%C3%A4%2031.3.2023_jakelu.pdf

Vuorinen, J. & Tetri, P. (2012). The Order Machine – The Ontology of Information Security. *Journal of the Association for Information Systems*, *13*(9), 695–713. https://doi.org/10.17705/1jais.00306

Vuorinen, J. & Tetri, P. (2016). Paradoxes in Information Security. *IEEE Potentials*, *35*(5), 36–39. https://doi.org/10.1109/MPOT.2016.2569740

Vähäkainu, P. & Neittaanmäki, P. (2018). Tekoäly terveydenhuollossa. *Informaatioteknologian tiedekunnan julkaisuja*, *45/2018*. https://jyx.jyu.fi/bitstream/handle/123456789/57682/1/978-951-39-7360-5.pdf

Wiafe, I., Koranteng, F. N., Obeng, E. N., Assyne, N., Wiafe, A. & Gulliver, S. R. (2020). Artificial Intelligence for Cybersecurity: A Systematic Mapping of Literature. *IEEE Access*, *8*, 146598–146612. https://doi.org/10.1109/ACCESS.2020.3013145

Zeadally, S., Adi, E., Baig, Z. & Khan, I. A. (2020). Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity. *IEEE Access*, *8*, 23817–23837. https://doi.org/10.1109/ACCESS.2020.2968045