Käyttäjien manipulointi ja sen ehkäiseminen

Oulun yliopisto
Tietojenkäsittelytieteet
Jyrki Tihinen
2024

Tiivistelmä

Käyttäjien manipuloinnilla (eng. social engineering) tarkoitetaan rikollisten käyttämiä keinoja, joilla pyritään vaikuttamaan uhreihin niin, että he paljastavat arkaluontoisia tietoja tai toimivat muutoin oman etunsa vastaisesti. Seurauksen luvaton taho voi esimerkiksi saada pääsyn organisaation toimitiloihin, tietojärjestelmiin tai yksityishenkilön pankkitunnuksiin tai salasanoihin.

Tässä tutkielmassa tarkastellaan, millaisia eri käyttäjien manipuloinnin muotoja on olemassa ja kuinka niitä vastaan voi suojautua. Työssä keskitytään erityisesti sähköisten kanavien kautta tapahtuvaan manipulointiin.

Käyttäjien manipulointia voidaan tehdä sekä sähköisten kanavien kautta että fyysisesti. Hyökkääjä voi esimerkiksi lähettää organisaation jäsenille aidolta vaikuttavia sähköposteja, jotka sisältävät linkin hyökkääjän hallinnoimalle väärennetylle verkkosivustolle. Fyysisessä manipuloinnissa hyökkääjä saattaa esimerkiksi esittää huoltohenkilöä pukeutumalla huomioliiveihin ja kantamalla mukanaan työkaluja. Sitten hyökkääjä yrittää päästä sisään organisaation tiloihin luottamalla siihen, että ihmiset ystävällisyyttään pitävät hänelle ovia auki.

Käyttäjien manipulointi on nykyään erittäin yleinen ilmiö ja sen aiheuttamat vahingot ja taloudelliset tappiot ovat valtavat. Onnistuneen hyökkäyksen seurauksena organisaation toiminta saattaa kärsiä pitkään tai jopa lakata kokonaan. Yksityishenkilön sosiaalisen median tiliä saatetaan käyttää huijausten levittämiseen ja uskottavuuden lisäämiseen. Pankkitietonsa rikollisille menettäneen henkilö saattaa menettää rahansa pankkitililtään ja niiden takaisin saaminen voi olla hankala prosessi, eikä uhri välttämättä aina saa kaikkia menettämiään rahojaan takaisin.

Käyttäjien manipuloinnin hyökkäykset kehittyvät koko ajan ja saavat uusia muotoja. Esimerkiksi Covid-19-pandemian aikana esiintyi Covid-teemaisia kalasteluviestejä, joissa käytettiin hyväksi pandemian aiheuttamaa ahdistusta ja pelkoa. Pandemian aikana QR-koodien käyttö lisääntyi valtavasti ja rikolliset ovat valjastaneet nekin hyökkäystarkoituksiin.

Avainsanat:
käyttäjien manipulointi, tietoturva, kalastelu, phishing, vishing, quishing

Ohjaaja:

FT, Yliopistonlehtori, Elina Annanperä

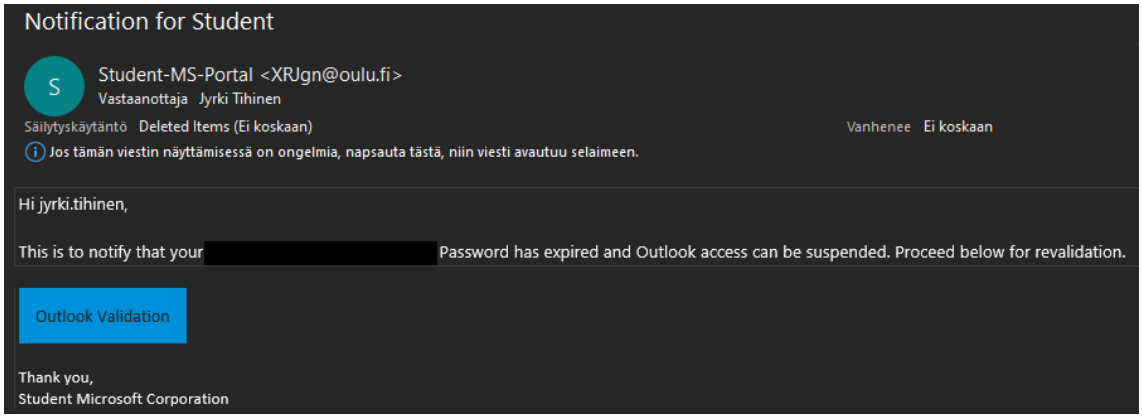
Sisällysluettelo

1. Johdanto

Tietoturva ja sen tutkimus ovat viime vuosina kehittyneet huimasti niiden saaman huomion ansiosta. Pelkkä teknologian kehittyminen ei kuitenkaan takaa sitä, että organisaation tai yksityishenkilön yksityiset tiedot olisivat turvassa. Usein tietoturvan heikoin osuus ovat sen käyttäjät (Mouton, Malan & Leenen, 2014).

Käyttäjien manipulointi (social engineering) tarkoittaa hyökkäystä, jossa hyödynnetään ihmismielelle tyypillisiä psykologisia ominaispiirteitä ja tehdään näistä ominaispiirteistä haavoittuvuuksia (Wang, Zhu & Sun, 2021). Näiden hyökkäysten tavoitteena on saada uhri paljastamaan arkaluontoista tietoa ja/tai saada uhri tekemään jotakin oman etunsa vastaista, kuten esimerkiksi asentamaan haittaohjelma (Abroshan, Devos & Laermans, 2021a). Koska kaikki tietotekniset järjestelmät ovat riippuvaisia käyttäjistä, taitava hyökkääjä voi käytännössä päästä sisään mihin tahansa järjestelmään (Wang ym., 2021).

Käyttäjien manipulointi on kiinnostava ilmiö, enkä ole itsekään säästynyt kalasteluviesteiltä. Esimerkiksi syyskuussa 2021 Oulun yliopisto joutui laajan tietojenkalastelukampanjan kohteeksi. Hyökkääjät saivat saaliikseen noin 900 käyttäjätunnuksen salasanaa (Uusitalo, 2021). Kuvassa 1 on esitetty kyseisessä hyökkäyksessä käytetty kalasteluviesti.



Kuva 1. Oulun yliopistoa vastaan käytetty huijausviesti syyskuulta 2021.

Kuvan 1 viestissä silmäänpistävintä on epämääräinen lähettäjän sähköpostiosoite. Se näyttäisi olevan peräisin oulu.fi -osoitteesta, mutta alkuosa ei ole mitään Oulun yliopiston ICT-palveluihin viittaavaa. Lisäksi viestin painikkeessa oleva hyperlinkki on kaikkea muuta kuin Oulun yliopiston viralliselle sivustolle johtava hyperlinkki.

Tämän tutkielman tavoitteena on löytää vastauksia seuraaviin kysymyksiin: Millaisia käyttäjien manipuloinnin muotoja on olemassa? Mikä on paras tapa suojautua manipuloinnilta? Millaisia teknisiä ratkaisuja ongelman torjuntaan on kehitetty? Kysymyksiin haettiin vastauksia tutustumalla aiheesta tehtyisin tutkimuksiin. Tässä työssä keskityn erityisesti sähköisten viestimien kautta tapahtuvaan manipulointiin ja sen ehkäisyyn organisaatioiden näkökulmasta.

Tutkielman alussa aihe esitellään lyhyesti ja selitetään tutkimusmenetelmä. Seuraavaksi tutustutaan aiheesta kertovaan kirjallisuuteen ja haetaan vastauksia yllä esitettyihin tutkimuskysymyksiin. Lopuksi esitetään kirjallisuuden pohjalta johdettuja päätelmiä ja esitetään mahdollisia jatkotutkimusaiheita.

2. Tutkimusmenetelmä

Tämän tutkielman materiaalit on haettu IEEE Xplorer ja Scopus -tietokannoista. Lisäksi lähteiden etsinnässä hyödynnettiin helmenkasvatus-menetelmää. Haussa löytyneiden kirjallisuuskatsausten lähdeluetteloista etsittiin työhön sopivia lähteitä. Lisäksi hyödynnettiin jonkin verran hakukoneiden suosittelemia samankaltaisia artikkeleita.

Lähteitä etsiessä käytettiin seuraavia hakutermejä: ”social engineering”, ”social engineering” AND phone, phishing, ”social engineering” AND impact, ”social engineering” AND COVID, “social engineering” AND “artificial intelligence”, “social engineering” AND training, “social engineering” AND gamification, quishing.

Hakutuloksia rajattiin julkaisuvuoden perusteella. Aluksi hakutuloksia rajattiin siten, että vanhimmat julkaisut olivat vuodelta 2000. Myöhemmin vanhimpia tuloksia rajattiin vuoteen 2010 ja uusimmat hakutulokset olivat vuodelta 2023. Löydetyistä artikkeleista valittiin ne, jotka ovat relevanteimpia työn rajaukseen, eli käyttäjien manipulointi ja sen ehkäisy.

3. Aiempi tutkimus

Tyypillinen piirre käyttäjien manipuloinnille on, että hyökkäyksissä käytetään hyväksi ihmismielen perusominaisuuksia, jotka taitavasti käytettynä voidaan muuttaa haavoittuvuuksiksi. Hyökkääjä saattaa esimerkiksi pyrkiä herättämään uhrissaan voimakkaita tunteita, kuten pelkoa, hämmennystä tai uteliaisuutta (Wang ym., 2021). Hyökkääjä saattaa esimerkiksi esiintyä auktoriteettina, kuten poliisina tai organisaation johtohahmona ja näin yrittää manipuloida uhriaan (Harris, Derakhshan & Carlsson, 2021). Toinen yleisesti käyttäjien manipuloinnin tekniikka on aikapaineen ja kiireen tunteen luominen. Näin pyritään hämäämään uhria niin, ettei hän kiinnitä huomiota huijauksen paljastaviin tunnusmerkkeihin, vaan taipuu hyökkääjän tahtoon (Abroshan ym., 2021a)

Käyttäjien manipuloinnin hyökkäykset voidaan jakaa kolmeen luokkaan hyökkäyksessä tapahtuvan viestinnän suunnan perusteella: kaksisuuntaiset hyökkäykset, yksisuuntaiset hyökkäykset ja epäsuorat hyökkäykset. Kaksisuuntaisessa hyökkäyksessä hyökkääjä ja uhri kommunikoivat keskenään esimerkiksi puhelimen välityksellä. Yksisuuntaisessa viestinnässä hyökkääjä on yhteydessä uhriinsa esimerkiksi sähköpostin välityksellä. Epäsuorassa hyökkäyksessä hyökkääjän ja uhrin välillä ei ole vuorovaikutusta, vaan hyökkääjä jättää hyökkäysvektorin, esimerkiksi haittaohjelmassa saastutetun USB-tikun, löydettäväksi kohdeorganisaation lähettyville (Mouton, Leenen & Venter, 2016).

Käyttäjien manipulointi on valtavan laaja ongelma. Esimerkiksi vuoden 2018 1. neljänneksellä Anti-Phishing Working Group (APWG) raportoi yli 260000 tietojenkalasteluhyökkäystä, joista lähes 40 % kohdistui maksupalveluorganisaatioihin. Vuotta aikaisemmin Kaspersky Lab kertoi raportissaan kalastelun vastaisen mekanisminsa torjuneen yli 246000000 hyökkäystä (Rao, Vaishnavi & Pais, 2020). Käyttäjien manipuloinnin taloudelliset vaikutukset ovat myös merkittävät: vuonna 2017 esitetyn arvion mukaan kyberrikollisuuden maailmanlaajuinen hinta on arviolta 500 miljardia Yhdysvaltain dollaria, josta kalastelun osuus on yli 90 %. Saman vuoden arvion mukaan kyberrikollisuuden hinta nousisi 2 biljoonaan dollariin vuoteen 2019 mennessä (Carella, Kotsoev & Truta, 2017).

Käyttäjän alttiutta manipuloinnille voidaan arvioida käyttäjän riskinottokäyttäytymisestä. Enemmän riskejä ottavat käyttäjät avaavat herkemmin kalastelusähköposteja ja klikkaavat niissä olevia linkkejä (Abroshan, Devos, Poels & Laermans, 2021c). Myös tietyt luonteenpiirteet, kuten itserakkaus ja ahneus ovat manipuloinnille altistavia ominaisuuksia (Wang ym., 2021). Myös sukupuolten välillä on havaittu eroja tietoturvakäyttäytymisessä ja alttiudessa joutua manipuloinnin uhriksi. Naiset saattavat olla hieman alttiimpia käyttäjien manipuloinnille kuin miehet (Abroshan, Devos, Poels & Laermans, 2021b).

Kalastelukampanjoiden määrä on lisääntynyt, ja perinteisten tietojenkalasteluviestien rinnalle on kehittynyt uudenlaisia COVID-19-teemaisia kalasteluviestejä, joissa saatetaan tarjota taudin ehkäisyyn ja hoitoon tarkoitettuja tuotteita. Tutkimuksessa havaitaan myös, että pandemian aiheuttama psykologiset vaikutukset, kuten stressi, ahdistus ja pelko, saavat ihmisiä helpommin lankeamaan kalasteluviesteihin (Abroshan ym., 2021a.)

Organisaatiotasolla onnistuneen hyökkäyksen vakavin seuraus on mainehaitta, josta kärsivät sekä käyttäjä itse että organisaatio. Muita organisaation toiminnan jatkuvuuden kannalta vakavia seuraamuksia ovat asiakasluottamuksen katoaminen, tietovuodot ja datan ja tietojärjestelmien saatavuuden rajoittuminen (Aldawood & Skinner, 2019.)

Manipulointihyökkäyksistä varoittavia toimintoja on jo olemassa, mutta niiden vaikutus on heikentynyt. Syitä tähän ovat viestintäkanavien runsas määrä ja niistä tuleva tiedon paljous. Käyttäjien voi olla vaikeaa erottaa oleellista tietoa informaatiotulvan keskeltä. Ajan myötä käyttäjät tottuvat liikaa turvallisuusilmoituksiin, jolloin ne helposti vain ohitetaan, eivätkä ne saa käyttäjiä reagoimaan tilanteeseen (Pasquali ym., 2023). Turvallisuusilmoituksissa saatetaan myös käyttää niin teknistä ja vaikeasti ymmärrettävää kieltä, että joillakin käyttäjillä on vaikeuksia ymmärtää niiden merkitystä (Sharevski, Devine, Pieroni & Jachim, 2022).

3.1 Sähköpostihyökkäykset

Sähköpostin kautta tapahtuva käyttäjien manipulointi useimmiten pyrkii kalastelemaan käyttäjien tietoja, kuten salasanoja tai pankkitietoja. Uhrille lähetetään sähköpostiviesti, joka sisältää linkin tietojen varastamiseen tarkoitetulle, luotettavan näköiselle sivustolle, ja viesti, joka houkuttelee uhria klikkaamaan linkkiä (Rao ym., 2020). Toinen tyypillinen hyökkäysmuoto on lähettää viestin mukana haitallinen liitetiedosto, joka avattaessa saastuttaa käyttäjän tietokoneen haittaohjelmalla, joka voi olla esimerkiksi käyttäjän tiedostot kryptaava kiristysohjelma (Carella ym., 2017). Tietojenkalastelukampanjoiden toteuttamiseen on olemassa valmiita työkaluja, joten ne ovat hyökkääjille helppo ja houkutteleva vaihtoehto (Rao ym., 2020).

Sähköpostin välityksellä tapahtuva kalastelu voidaan jakaa kolmeen eri kategoriaan. Phishing: uhri yritetään saada klikkaamaan sähköpostissa olevaa linkkiä, joka johtaa väärennetylle sivustolle, ja välittää sinne syötetyt tiedot hyökkääjälle. Spear phishing on räätälöidympi versio phishing-hyökkäyksestä. Sen tavoitteena on jonkin tietyn organisaation tietojen kalasteluun pyrkivä hyökkäys. Whaling on organisaation ylimmän johdon tietojen kalasteluun pyrkivä hyökkäys (Mattera & Chowdhury, 2021).

Yksi suurimmista syistä sähköpostihyökkäysten onnistumiselle ovat käyttäjien puutteelliset tiedot tietokonejärjestelmistä ja tietoturvasta. Käyttäjät eivät esimerkiksi ymmärrä URL-osoitteiden merkitystä, eivätkä erota väärennettyä ja aitoa sivustoa toisistaan URL-osoitteen perusteella tai osaa tulkita sähköpostin otsikkotietoja (Dhamija, Tygar & Hearst, 2006).

Kalastelusivustoille on tyypillistä, että niillä pyritään hämäämään uhria visuaalisilla keinoilla ja ohjaamaan uhrin huomio pois sivuston vaarallisuudesta kertovista tunnusmerkeistä. Visuaalisia hämäyskeinoja voivat olla esimerkiksi kirjainten korvaaminen sivuston osoitteessa samankaltaisilla kirjaimilla tai numeroilla. Sivustoilla voidaan myös näyttää väärennettyjä turvallisuuselementtejä, kuten salatusta https-yhteydestä ilmoittavaa lukon kuvaa (Dhamija ym., 2006).

3.2 Puhelinhyökkäykset

Puhelimen välityksellä tapahtuvaa käyttäjien manipulointia kutsutaan yleisesti englannin kielen termillä vishing (voice phishing) (Jones, Armstrong & Tornblad, 2020). Vishing-hyökkäykset voivat olla ennalta nauhoitettuja tai reaaliaikaisia, jolloin uhri on suoraan tekemisissä hyökkääjän kanssa. Nauhoitetuilla hyökkäyksillä on helpompi toteuttaa laajempia hyökkäyksiä, kun taas reaaliaikaisella hyökkäyksellä voidaan toteuttaa jotain tiettyä organisaatiota tai henkilöä vastaan räätälöity hyökkäys (Derakhshan, Harris & Behzadi, 2021).

Vishing-hyökkäysten tunnistaminen ja torjunta on vaikeampaa kuin sähköpostin välityksellä tapahtuvien hyökkäysten torjuminen. Sähköpostissa on otsikkotietoja ja muuta metadataa, joita voidaan käyttää vaarallisten sähköpostien tunnistamiseen. Puhelussa vastaavia tietoja ei ole. Lisäksi soittajan tiedot, kuten puhelinnumero, ovat helposti väärennettävissä (Derakhshan ym., 2021).

Vishing-hyökkäys on tyypillisesti sähköpostitse tapahtuvaa hyökkäystä tehokkaampi menetelmä. Puhelinhyökkäys voi vaikuttaa enemmän uhrin tunteisiin, jos hän on reaaliaikaisesti tekemisissä toisen ihmisen kanssa puhelimen välityksellä. Uhrille on helpompi luoda painetta toimia ja tehdä mitä hyökkääjä haluaa (Harris ym., 2021).

Jones ym. (2020) tutkivat, kuinka uhri saadaan suostuteltua hyökkääjän tahtoon vishing-hyökkäyksissä. Tutkimuksessa havaittiin, että hyökkääjä todennäköisesti esittää olevansa uhrin organisaation ulkopuolinen taho, jolla on kuitenkin oikeus päästä käsiksi arkaluontoiseen tietoon. Uhrin olisi helpompi todentaa oman organisaation sisäisiä hierarkioita kuin ulkoisen tahon. Tutkimuksessa havaitaan myös, että hyökkääjä pyrkii imitoimaan aitoa puhelua mahdollisimman paljon. Vishing-hyökkäyksissä ei yleensä esiinny elementtejä, kuten ystävällisyyttä tai flirttailua, joilla pyrittäisiin saamaan uhri pitämään hyökkääjästä. Ystävälliset eleet odottamattomassa puhelussa voisivat herättää uhrissa epäilyksiä, joten hyökkääjät useimmiten välttävät niitä.

Jones ym. (2020) havaitsevat tutkimuksessaan myös merkittävimmän eron sähköpostitse ja puhelimitse tapahtuvien hyökkäysten välillä. Puhelinhyökkäykset ovat räätälöidympiä ja vaativat spesifimpiä suostuttelukeinoja. Hyökkääjä voi painottaa erityisesti tottelemisesta seuraavia hyötyjä uhrille ja uhrin organisaatiolle. Tämä on vaikeampaa toteuttaa perinteisillä kalastelusähköposteilla pelkillä kuvilla ja teksteillä.

3.3 QR-koodit

QR-koodit ovat olleet olemassa jo pidemmän aikaa, mutta niiden käyttö lisääntyi merkittävästi COVID-19-pandemian aikana. Syntyi tarve nopealle ja kosketusvapaalle tavalle jakaa tietoa ja URL-osoitteita, ja QR-koodit vastaavat tähän tarpeeseen täydellisesti. Myös rikolliset havaitsivat tämän. QR-koodien välityksellä tapahtuva kalastelu, eli quishing, ei myöskään saanut kovin suurta huomiota ennen pandemiaa (Sharevski ym., 2022).

QR-koodien välityksellä toimivat hyökkäykset ovat toimintaperiaatteeltaan varsin samankaltaisia kuin sähköpostitse tapahtuvat hyökkäykset. Uhri yritetään saada avaamaan haitallinen linkki ja syöttämään tietonsa hyökkääjän hallinnoimalle kalastelusivustolle tai asentamaan laitteelleen haittaohjelman. Haitallisia QR-koodeja voidaan levittää sekä fyysisesti että digitaalisesti. Hyökkääjä voi esimerkiksi liimata haitallisen QR-koodin sisältävän tarran näkyville julkiselle paikalle, kuten esimerkiksi julisteeseen tai pysäköintiautomaattiin, ja odottaa pahaa-aavistamattomien uhrien skannaavan sen. Haitallisia QR-koodeja voidaan levittää myös sähköpostin välityksellä, jolloin vastaanottajan haitallisen linkkien suodatus ei välttämättä tunnista viestiä haitalliseksi, vaan päästää sen läpi (Sharevski ym., 2022).

QR-koodihyökkäykset ovat kehittyneet ovelammiksi ajan myötä. Aluksi hyökkääjät liimasivat haitallisia koodeja alkuperäisten koodien päälle. Seuraavaksi hyökkääjät oppivat sulauttamaan haitallisia koodeja aitojen koodien sisään, jolloin koodia skannatessa haitallinen koodi tulee skannatuksi ensin. Myöhemmin hyökkääjät keksivät

keinon muokata aidon koodin pikseleitä ja näin saada haitallinen koodi upotettua aitoon koodiin ilman koodin fyysistä muokkausta (Sharevski ym., 2022).

Tietoa QR-koodeihin liittyvistä riskeistä on ollut saatavilla varsin rajoitetusti, eikä älypuhelimissa ole ollut tehokkaita teknisiä ratkaisuja haitallisten koodien tunnistamiseksi. Vasta viime aikoina esimerkiksi Yhdysvaltain liittovaltion poliisi FBI on ryhtynyt aktiivisesti tiedottamaan kansalaisia QR-koodeihin liittyvistä riskeistä. (Sharevski ym., 2022).

3.4 Käyttäjien manipuloinnin ehkäisy

Tietyt luonteenpiirteet altistavat ihmisiä käyttäjien manipuloinnille. Abroshan ym. (2021a) kehittävät tutkimuksessaan menetelmän, jolla pyritään kartoittamaan organisaation henkilöstön alttiutta manipuloinnille ja vähentämään sitä. Ratkaisu koostuu kolmesta vaiheesta: käytöksen mittaaminen, riskin pisteytys ja niiden ehkäisy ja torjunta. Ensimmäisessä vaiheessa valitaan käytettävä asteikko, kuten esimerkiksi Dohmenin riskinottomittari tai Koronavirusahdistusasteikko. Toisessa vaiheessa määritetään työntekijöiden alttius käyttäjien manipuloinnille laskemalla ensimmäisessä vaiheessa saadut pisteet. Tässä voidaan käyttää apuna koneoppimista. Ehkäisy- ja torjuntavaiheessa tarjotaan työntekijöille sopivaa koulutusta. Se voi olla tietoisuutta lisäävää koulutusta joko luokkahuoneessa tai verkossa. Toinen mahdollisuus on psykologinen ehkäisy, kuten esimerkiksi kognitiivinen käytösterapia, jota voidaan hyödyntää muun muassa pelon ja ahdistuksen lieventämiseen.

3.4.1 Sähköpostihuijausten ehkäisy

Rao ym. (2020) vertailevat tutkimuksessaan erilaisia sähköpostihyökkäysten torjuntaan kehitettyjä teknisiä toteutuksia. Mustien ja valkoisten listojen avulla voidaan määrittää kielletyt ja sallitut IP- ja URL-osoitteet. Käytössä listat ovat tehokas keino karsia liikennettä haitallisille sivustoille, mutta niiden heikkous ovat täysin uudet kalastelusivustot, eli nollapäivähyökkäykset, jotka eivät ole vielä päätyneet mustalle listalle. URL-luokittelutekniikat etsivät URL-osoitteista kalastelun paljastavia elementtejä, kuten epäilyttäviä binääriominaisuuksia ja mustalla listalla olevia sanoja. Sisällönsuodatustekniikat tutkivat sivuston lähdekoodia ja sieltä kalasteluun viittaavia piirteitä.

Rao ym. (2020) kehittävät tutkimuksessaan CatchPhish-nimisen sovelluksen, joka kykenee tunnistamaan kalastelusivuston pelkän linkin perusteella. Käyttäjä syöttää tarkistettavat URL-osoitteet ohjelmaan, ohjelma lähettää osoitteet palvelimelle, jossa ne tutkitaan tarkistuslistan ja TF-IDF (Term Frequency-Inverse Document Frequency) -menetelmällä. Lopuksi ohjelma kertoo tarkistuksen tulokset käyttäjälle. Sovelluksen käyttäminen on nopeaa, koska tekstimuotoisten analyysien suorittaminen on nopeampaa kuin sivuston tai sen lähdekoodin analysoiminen. Sovelluksen käyttäminen myös lisää turvallisuutta, kun epäilyttävää sivustoa ei tarvitse avata.

Pelkät tekniset ratkaisut eivät kuitenkaan riitä sähköpostihyökkäysten torjumiseen. Käyttäjien kouluttaminen on olennaisessa osassa niiden ehkäisemisessä. Carella ym. (2017) tutkivat, millainen koulutus on tehokkainta näiden hyökkäysten torjumisessa. Tutkimuksen aikana muodostetaan kolme ryhmää, joista ensimmäisen ryhmän, eli kontrolliryhmän jäsenet eivät saa koulutusta, toisen ryhmän jäsenet saavat koulutusta avattuaan tutkijoiden lähettämän kalasteluviestin linkin, ja kolmas ryhmä saa opetusta

luokkahuoneessa. Tutkimuksessa havaitaan, että luokkahuoneessa saadun koulutuksen teho alkaa vähetä viikkojen sisällä koulutuksesta. Viestien kautta saatu valistus vähensi klikkauksia huomattavasti enemmän. Tutkimuksessa todetaan, että organisaation pitää työntekijöiden testaamisen lisäksi käyttää negatiivista rangaistusta, kuten tutkimuksen toinen ryhmä sai. Näin saadaan käyttäjät paremmin ymmärtämään varovaisuuden ja valppauden tärkeys linkkejä avattaessa.

3.4.2 Puhelinhuijausten ehkäisy

Myös puhelimen kautta tapahtuvien hyökkäysten tärkein torjuntakeino on koulutus. Tämä pätee erityisesti niihin organisaatioiden työntekijöihin, jotka ovat tekemisissä arkaluontoisten tietojen kanssa, ja joilla ei ole erityisen hyvää tietämystä monimutkaisista kalasteluhyökkäyksistä (Jones ym., 2020). Työntekijät, jotka eivät oleta joutuvansa kalasteluhyökkäyksen uhriksi ja eivät ymmärrä käsittelemänsä tiedon merkitystä ja arvoa organisaatiolle ja hyökkääjälle, ovat alttiimpia kalasteluhyökkäyksille ja saattavat luovuttaa tietoja hyökkääjille varsin helposti (Mouton ym., 2016).

Mouton ym. (2016) ovat luoneet käyttäjien manipuloinnin havaitsemiseen tarkoitetun mallin (Social Engineering Attack Detection Model, SEADM). Ensimmäinen versio oli tarkoitettu puhelinkeskusten työntekijöille kaksisuuntaisen viestinnän tilanteisiin. Mallin seuraava versio (SEADMv2) on jatkokehitetty siten, että sitä voidaan käyttää myös yksisuuntaisen ja epäsuoran manipuloinnen tunnistamiseen. Tutkimuksen kirjoittajat ovat myöhemmin toteuttaneet malliinsa perustuvan sovelluksen (Social Engineering Prevention Training Tool, SEPTT) Android-laitteille. Kirjoittajien omissa tutkimuksissa sovellus vähensi huomattavasti kaksisuuntaisten ja epäsuorien hyökkäysten onnistumismääriä. Yksisuuntaisia hyökkäyksiä vastaan sovellus ei kuitenkaan toiminut suunnitellusti (Mouton, Teixeira & Meyer, 2017).

Derakhshan ym. (2021) toteavat tutkimuksessaan, että vaikka puhelinhyökkäysten tarkka sisältö vaihtelee, hyökkäyksissä esiintyy samankaltaisuuksia ja piirteitä, joista hyökkäyspuhelut voidaan tunnistaa. Näitä tunnusmerkkejä kutsutaan huijaustunnusmerkeiksi (scam signature) ja niitä voidaan käyttää hyökkäysten tunnistamiseen samoin kuin haittaohjelmien tunnusmerkkejä käytetään haittaohjelmien tunnistamiseen. Tutkimuksessa huijaustunnusmerkit määritellään sarjaksi ilmaisuja, jotka ovat uniikkeja erityyppisille huijauksille ja joilla hyökkääjä pyrkii saavuttamaan tavoitteensa.

Derakhshan ym. (2021) kehittävät hyökkäysten torjuntaan tarkoitetun työkalun Anti-Social Engineering Tool (ASsET). Työkalun algoritmi kykenee käsittelemään luonnollista kieltä ja tunnistamaan keskustelusta eri huijausten tunnusmerkkejä. Tutkijoiden omassa testissä työkalu toimii hyvin tehokkaasti. Tutkijoiden testissä algoritmi tunnistaa yli 90 % huijauksista, eikä aiheuta paljon vääriä positiivisia tuloksia.

3.4.3 QR-koodihuijausten ehkäisy

Sharevski ym. (2022) selvittävät tutkimuksessaan, kuinka alttiita QR-koodien käyttäjät ovat kalastelulle, mitkä turvaominaisuudet auttavat käyttäjiä tunnistamaan haitalliset QR-koodit, ja kuinka QR-koodikalastelututkimuksen tuloksia voidaan hyödyntää QR-koodihyökkäysten vastaisessa koulutuksessa. Tutkimusta varten tutkijat kehittävät ensin asteikon nimeltään *Quishing Awareness Scale (QAS)*, jolla voidaan mitata käyttäjän tietoturvavalveutuneisuutta. Asteikko perustuu jo olemassa olevaan *Security Behavior*

Intention Scale (SeBIS) asteikkoon, jolla mitataan käyttäjien valveutuneisuutta sähköpostikalasteluun. Asteikko koostuu tietoturvakäyttäytymistä mittaavista kysymyksistä, joihin vastataan numerolla väliltä 1–5, jossa 1 on ”ei koskaan” ja 5 on ”aina”. Tutkimusta varten tutkijat luovat Yhdysvaltain liittovaltion tartuntatautiviranomaisen *CDC:n* virallisia tiedotteita jäljittelevän julisteen, jossa houkutellaan käyttäjiä osallistumaan digitaalisen rokotepassin testaamiseen. Julisteessa oleva QR-koodi ohjaa uhrin tutkijoiden luomalle kalastelusivustoa jäljittelevälle verkkosivustolle, jossa käyttäjälle tarjotaan eri tapoja kirjautua sivustolle, luoda uusi tili kirjautumista varten tai jatkaa ilman kirjautumista, jos käyttäjä huomasi painaa lisää kirjautumisvaihtoehtoja -painiketta sivun alaosassa. Todellisuudessa käyttäjien tunnuksia ei kerätty. Seuraavaksi käyttäjälle esitettiin kysymyksiä käyttäjän valitsemaan kirjautumisvaihtoehtoon, QR-koodien käyttöön ja quishing-tietoisuuteen liittyen. Tulosten perusteella tutkijat pyrkivät luomaan ohjeistuksia QR-koodihyökkäysten vastaiseen koulutukseen.

Sharevski ym. (2022) havaitsevat tutkimuksensa tuloksista, että Facebook-tunnuksilla kirjautuneiden henkilöiden QAS-tulokset ovat matalimmat, eli heidän quishing-tietoisuutensa on heikoin, kun taas kirjautumisen ohittaneiden henkilöiden QAS-tulokset olivat korkeimmat, eli heidän quishing-tietoisuutensa oli tutkimukseen osallistuneiden paras. Tutkijat havaitsevat tuloksista myös, että auktoriteettina tai viranomaisena esiintyminen toimii tehokkaasti myös QR-koodihyökkäyksissä.

Sharevski ym. (2022) tutkimuksen vastaajista hieman yli puolet kertoo käyttäneensä sosiaalisen median tunnuksia kirjautumiseen helpomman kirjautumisen takia. Huolestuttava havainto tuloksissa on se, että osallistujista alle 5 % epäili voivansa joutua tietojenkalastelun uhriksi.

QR-koodihyökkäysten ehkäisystä Sharevski ym. (2022) toteavat, että QR-koodihyökkäysten vastainen koulutus on vielä alkutekijöissään ja monilla kouluttajillakin on vaikeuksia tunnistaa ilmiötä QR-koodien laajojen käyttömahdollisuuksien ja ilmiön uutuuden takia. Tutkijat kirjoittavat, että koulutuksessa tulisi faktojen lisäksi selittää koko hyökkäyksen rakenne, eli millä verukkeella uhri saadaan skannaamaan haitallinen koodi ja luovuttamaan tietonsa kalastelusivustolle. He luovat tutkimuksen pohjalta kuuden kohdan taulukon, jossa opastetaan tunnistamaan quishing ilmiönä ja siitä varoittavat tunnusmerkit.

Sharevski ym. (2022) havaitsevat selviä ongelmia nykyisissä selainten turvaominaisuuksissa ja -ilmoituksissa. Tutkimuksen osallistujista vain yksi oli kertonut havainneensa vaaran merkin avatessaan sivuston. Koska turvailmoitusten vaikutus on vähentynyt, tutkijat kehittävät konsepteja uudenlaisista ilmoitusta, joilla varoitetaan käyttäjää mahdollisesta vaarasta. Ilmoitukset suunnitellaan niin, että ne keskeyttävät käyttäjän työnkulun ja kertovat selkeästi, mitä on tapahtumassa. Tutkimusta varten kehitetään neljä erilaista varoitusnäkymää ja niiden käytettävyys ja vaikuttavuus testataan testiryhmällä. Tuloksista havaitaan, että vastaajan QAS-tulos korreloi parhaaksi katsotun vaihtoehdon kanssa. Keskivertoa matalamman QAS-tuloksen saaneet henkilöt suosivat koko ruudun peittävää punaista varoitusnäkymää, joka osoittautuu myös kaikista vaihtoehdoista suosituimmaksi. Korkeamman QAS-tuloksen saaneet vastaajat suosivat mieluummin väritykseltään neutraalimpaa ilmoitusta, joka kuitenkin erottuu selkeästi muista näytölle tulevista push-ilmoituksista.

3.4.4 Robottien käyttö manipuloinnin torjunnassa

Koska tietoturvauhista varoittavien ilmoitusten teho on vähentynyt, Pasquali ym. (2023) tutkivat, kuinka robotteja voidaan hyödyntää turvailmoitusten välittämisessä. Robotit ovat vielä suhteellisen uusi ilmiö, eivätkä käyttäjät ole ehtineet turtua niihin. Aiemmissa tutkimuksissa on myös havaittu, että robotit voivat auttaa ihmisiä keskittymään paremmin ja ne toimivat sosiaalisina tekijöinä. Fyysisen robotin läsnäolo voi tehostaa varoitusviestien vaikuttavuutta. Tutkimuksen tavoitteena on selvittää, kuinka robotin tulisi olla vuorovaikutuksessa käyttäjänsä kanssa, jotta se voisi suojella käyttäjäänsä manipulointihyökkäyksiltä.

Tutkimustaan varten Pasquali ym. (2023) värväävät 19 koehenkilöä, jotka jaetaan tutkimusta varten kahteen ryhmään: looginen suostuttelu, jonka päätöksiin vaikutetaan loogisin argumentein ja affektiivinen ryhmä, jonka päätöksiin vaikutetaan tunteisiin vetoavin keinoin. Osallistujat laitetaan pelaamaan videopeliä. Pelin tavoitteena on, että valuuttana ja terveyspisteinä toimiva kvanttienergia ei pääse loppumaan ja että pelaajan pelin sisäinen henkilöllisyys ei paljastu pelin aikana. Pelin on tarinavetoinen ja pelaajat reagoivat pelin aikana 21 eri tilanteeseen kysymyksiin vastaamalla. Tilanteet on laadittu niin, että niissä esiintyy manipuloinnissa yleisesti käytettyjä keinoja, kuten auktoriteettina esiintyminen, kiire, niukkuuden tunne ja lahjan saaminen henkilötietoja vastaan. Tutkimuksessa käytetään Furhat-pöytärobottia, jolla on selkeät ihmiskasvot ja joka kykenee esittämään eri ilmeitä ja tunnetiloja. Se toimii pelissä pelaajan apulaisena, joka toimii pelin tarinan kertojana ja yrittää vaikuttaa pelaajien päätöksentekoon. Tutkimuksessa mitataan, kuinka robotti onnistuu vaikuttamaan pelaajien päätöksiin (eksplisiittinen reaktio), kuinka osallistujien kasvojen ilmeet muuttuvat pelin aikana ja miten he liikuttavat kursoria tietokoneen ruudulla (implisiittinen reaktio).

Pasquali ym. (2023) tutkimuksen tuloksista käy ilmi, että robotti kykenee vaikuttamaan pelaajien päätöksiin, vaikka ne olisivat pelaajan edun vastaisia. Myös aiemmissa tutkimuksissa on havaittu, että ihmiset ovat taipuvaisia luottamaan robotteihin riskialttiissa tilanteissa. Tutkimuksessa selviää, että loogisesta ja affektiivisesta vaikutustavasta looginen tapa on tehokkaampi. Tutkijat arvelevat, että looginen suostuttelu sai pelaajat pohtimaan tilannetta ja tekemään ratkaisuja loogisen ajattelun perusteella. Käyttäjän valintojen ennustaminen kasvojen ilmeiden ja hiiren liikkeiden perusteella vaatii lisää tutkimusta. Tämän tutkimuksen perusteella kasvojen ilmeet eivät ole yhteydessä valintojen ennustamiseen, mutta tulos saattaa johtua tutkimuksen etätoteutuksesta.

3.4.5 Koulutuksen haasteet

Aldawood ym. (2019) tutkivat käyttäjien manipuloinnin vastaista koulutusta ja sen haasteita. Tutkimuksessa todetaan, että koulutuksen suunnitteluun ja toteutukseen liittyy haasteita, jotka voidaan jakaa kuuteen eri lähteeseen: bisnesympäristö, sosiaalinen, valtiollinen, organisatorinen, taloudellinen ja henkilökohtainen. Tutkimuksessa havaitaan myös, että suurin haaste manipuloinnin vastaiselle koulutukselle organisaatioissa on koulutukseen vaadittavan rahoituksen puute. Organisaatiot eivät tunnista käyttäjien manipuloinnin aiheuttamia riskejä, eikä niillä ole juurikaan haluja käyttää rahaa muuhun kuin suoraan liiketoimintaan, joten manipuloinnin vastaiseen koulutukseen ei varata määrärahoja. Tutkimuksessa ehdotetaan, että organisaatiot testaavat työntekijöidensä valveutuneisuutta, ja kohdentavat koulutusta työntekijöiden tarpeiden mukaan. Asiaan vähemmän perehtyneet saavat yleisluontoisempaa koulutusta ja enemmän perehtyneille tarjotaan spesifimpää koulutusta.

Aldawood ym. (2019) mukaan toinen merkittävä haaste koulutuksen toteutuksessa on koordinointi tiimin jäsenten kesken. Organisaatioissa tulee kehittää suunnitelma, josta selviää, millaisia eri vastuualueita tiimin jäsenillä on uhkien hallinnassa. Tämä suunnitelma tulee myös tehdä mahdollisimman helposti havainnoitavaksi, ja tulee varmistaa, että kaikki tiimin jäsenet saavat tietoa siitä. Tietoturva-asiantuntijoiden tulee myös seurata käyttäjien manipuloinnin hyökkäysten evoluutiota ja kehittää organisaation tietoturvakoulutusta sen mukaisesti.

Aldawood ym. (2019) tutkimuksessa havaitaan myös, että perinteinen tietoturvakoulutus sisältää usein ryhmässä tehtäviä aktiviteetteja. Seurauksena saattaa syntyä tiimin sisäisiä riippuvuussuhteita, joka vaikeuttaa työntekijän toimintaa itsenäisessä työskentelyssä ja saattaa altistaa käyttäjien manipuloinnille. Ratkaisuna tähän toimii esimerkiksi tiedostuskampanja, jossa rohkaistaan työntekijöitä ryhtymään itsenäisesti toimeen manipuloinnin estämiseksi. Tutkimuksessa mainitaan myös, että manipuloinnin vastaisessa koulutuksessa tulee painottaa, että hyökkäykset eivät rajoitu vain koulutuksessa käytettyjen esimerkkien kaltaisia hyökkäyksiä. Hyökkääjät saattavat käyttää jotain täysin odottamatonta hyökkäystekniikkaa, kuten esimerkiksi käänteispsykologiaa. Koulutuksessa tulee käyttää monipuolisesti ja vaihtelevasti esimerkkejä elävästä elämästä.

3.5 Koulutuksen pelillistäminen

Käyttäjien manipuloinnin vastainen koulutus voidaan toteuttaa luentomuotoisen opetukena, mutta se ei ole kovin tehokasta ilman käytännön osuutta. Kouluttava organisaatio voi järjestää koulutettavalle organisaatiolle koulutustarkoituksessa tehtävän manipulointikampanjan, mutta niiden haasteena useimmissa tapauksissa ovat lailliset ja eettiset seikat (Hafner ym., 2023). Kyberturvallisuuskoulutus voi käyttäjien mielestä olla vain ikävystyttävä pakko, johon ei jakseta panostaa. Tämän korjaamiseksi on kehitetty erilaisia manipuloinnista valistavia pelejä koulutuksen interaktivoimiseksi ja kiinnostavuuden lisäämiseksi. Näistä peleistä puhuttaessa voidaan käyttää termiä hyötypeli (eng. serious game) (Aladawy, Beckers & Pape, 2018).

Pelillistämisellä tarkoitetaan pelin kaltaisten ominaisuuksien lisäämistä johonkin, jota ei tyypillisesti pidetä pelinä, kuten esimerkki tästä on koulutus. Pelin kaltaisia ominaisuuksia voivat olla esimerkiksi pisteytys ja erilaiset haasteet. Koulutus voi tapahtua sekä digitaalisessa muodossa, että fyysisesti lautapelien muodossa (Hafner, Wutz, Pöhn & Hommel, 2023).

3.5.1 Persuaded

Aladawy ym. (2018) ovat aiemmin tutkineet sosiaalipsykologian ja kyberturvallisuuden välisiä riippuvuussuhteita käyttäjien manipuloinnin näkökulmasta ja kartoittaneet psykologisia metodeja manipuloinnin ehkäisemiseksi. Tutkimuksessa havaittiin koulutuksessa esiintyvä toistuva puute: koulutuksessa ei painoteta riittävästi haastavien tilanteiden toistuvuutta. Ratkaisuksi tähän ongelmaan tutkijat kehittävät oman opetuspelin nimeltään *Persuaded*. Tutkimuksen tavoitteiden pohjalta tutkijat määrittävät opetuspelilleen seuraavat vaatimukset: helppo opittavuus, helppo pelattavuus, uudelleenpeluuarvo, pelaajan rooli ja tekstikonteksti. Jotta peli olisi mahdollisimman suurelle joukolle helposti lähestyttävä, sen pitää olla mahdollisimman yksinkertainen ja helposti ymmärrettävä. Pelaajan roolin tulee olla hyökkäyksen uhri, jotta pelaaja saa mahdollisimman todenmukaisen käsityksen omasta osuudestaan

manipulointiprosessissa. Peliskenaariot esitetään tekstimuodossa, jotta ne tukevat hyökkäys-puolustusskenaariota mahdollisimman tehokkaasti. Tutkijat päätyvät kehittämään pasianssin kaltaiseen, yksin pelattavaan, tietokoneella pelattavaan korttipeliin.

Aladawy ym. (2018) testaavat pelin toimivuutta koeryhmällä, joka koostuu 21 19–35-vuotiaasta yliopisto-opiskelijoista ja akateemisella alalla työskentelevästä henkilöstä. Osallistujista lähes kaikki käyttävät tietokonetta päivittäin ja puolet päivittäin työtarkoituksiin. Osallistujat vastasivat kyselyyn sekä ennen että jälkeen pelin pelaamisen. Kyselyn tuloksilla kartoitetaan osallistujien alttiutta manipuloinnille ja pelin vaikutuksia siihen. Tulosten mukaan osallistujat ovat pelin pelaamisen jälkeen huomattavasti vähemmän alttiita skenaarioissa, joissa heille tarjotaan tuntematonta tallennusmediaa (baiting) ja jossa tuntematon henkilö yrittää saada heiltä pääsyä lukittuihin tiloihin (tailgating). Pelissä oli kaksi erilaista phishing skenaariota. Toisessa ystävä pyytää uhrilta rahaa ja toisessa uhri koitetaan saada avaamaan tuntematon liitetiedosto. Näissä skenaarioissa pelin pelaamisella ei ollut suurta vaikutusta.

Tutkimukseen osallistuneille teetettyjen kyselyjen mukaan Aladawy ym. (2018) saavuttavat pelille asetetut tavoitteet. Peli tutustutti pelaajat uusiin manipuloinnin ja tietoturvahyökkäysten muotoihin, joita he eivät osanneet edes harkita ennen pelaamista. Pelin vaikeustaso vaikuttaa olevan juuri sopiva: peliä kuvaillaan helposti ymmärrettäväksi, mutta se pakottaa silti pelaajan pohtimaan eteen tulevia tilanteita. Pelin uudelleenpeluuarvo saavutettiin lisäämällä peliin sattumanvaraisuutta lisääviä elementtejä. Sattumanvaraisuus ei ollut kaikkien tutkimukseen osallistuneiden mieleen, mutta ainakin yksi osallistuja totesi sen simuloivan hyvin oikeaa elämää. Koskaan ei voi etukäteen tietää, mitä tuleman pitää.

3.5.2 TASEP

Hafner ym. (2023) toteavat, että vaikka kyberturvallisuustaitoja kehittävien pelien määrä on viime vuosina kasvanut, erityisesti käyttäjien manipulointia vastaan kouluttavia pelejä ei vielä ole paljoa tarjolla. He kehittävät tutkimuksessaan suosittuun Dungeons & Dragons -pöytäroolipeliin pohjautuvan roolipelin nimeltään TASEP, joka tulee sanoista *Tabletop As Social Engineering Prevention*. Heidän tavoitteenansa on luoda realistinen ja opettavainen peli, jossa pelaajat pääsevät käyttämään luovuuttaan ja oppimaan, kuinka hyökkääjät suunnittelevat manipulointikampanjoita. Peli on kuitenkin tarkoitus pitää mahdollisimman yksinkertaisena, jotta se on helppo ymmärtää ja pelin aloittaminen on vaivatonta. Pelin pelaamiseen tarvitaan useamman henkilön muodostama joukkue, joka toimii yhdessä pelin tavoitteiden saavuttamiseksi. Normaalin pöytäroolipelin tavoin pelissä on myös yksi pelin vetäjä *(game master),* joka selittää pelaajille heidän tehtävänsä, kertoo pelaajille pelin tarinan ja hoitaa skenaariossa muiden kuin pelaajien hahmojen liikkeet. Pelin vetäjän tehtävä on pitää peli realistisena ja hän voi tarvittaessa estää pelaajia tekemästä tiettyjä toimintoja, jotka eivät sovi pelin ja skenaarion henkeen.

Peliä varten Hafner ym. (2023) rakentavat peliä varten Lego-palikoista kaksi mallia. Yksi kuvastaa pieniä ja keskisuuria organisaatioita ja toinen suuria organisaatioita. Pienten ja keskisuurten yritysten malli sisältää yhden toimistorakennuksen, jossa kohdeorganisaatolla on muutama toimistohuone ja taukohuone. Suuren organisaation mallissa on yksi tai useampi aidalla ympäröity rakennus mahdollisesti tarkempi kulunvalvonta, pysäköintialue, sosiaalitiloja ja kokoushuoneita. Pelaajat eivät pelin alussa näe rakennusten sisälle, mutta pelin edetessä ja tiedon karttuessa pelaajat pääsevät tutkimaan rakennusten sisäosia. Molempiin malleihin voidaan lisätä myös valinnaisia

elementtejä, kuten valvontakameroita ja vartijoita. Molempiin malleihin kehitetään myös kolme peliskenaariota, joissa määritellään pelaajien tavoite, heidän taustansa ja käytettävissään olevat resurssit. Mallit sisältävät värikoodattuja esineitä, joita pelaajat voivat tutkia, kuten esimerkiksi työpöydät ja kaapistot. Myös ovet rakennuksen sisällä on värikoodattu. Vihreät ovet ovat avoimia ja muiden ovien kulkuoikeuksia on rajoitettu.

Koska kyseessä on pöytäroolipeli, Hafner ym. (2023) pelaajat luovat omat hahmonsa ja allokoivat niille taitopisteitä eri kykyihin. Pelihahmojen kyvyt jaetaan kolmeen pääluokkaan: sosiaaliset taidot, kuten manipulointi ja sopeutuminen, tekniset taidot, kuten väärentäminen ja ohjelmointi, sekä murtautumistaidot, kuten tiirikointi ja näpistely. Kun pelaaja kehittää tarpeeksi jotakin kykyä, hän saa käyttöönsä jonkin kykyyn liittyvän esineen. Esimerkiksi taitava manipuloija saa käyttöönsä lahjukset ja taitava tiirikoija murtautumistarvikkeita.

Hafner ym. (2023) testaavat peliä kahdella opiskelijoista muodostuvasta vähintään kolmen pelaajan ryhmällä, joista toinen koostuu tietokonealan opiskelijoista ja toinen humanististen alojen opiskelijoista. Pelin vetäjä toimii käyttäjän manipulointiin perehtynyt henkilö ja yksi henkilö toimii tarkkailijana. Ennen pelin aloitusta pelaajat vastaavat monivalintakysymyksistä koostuvaan kyselyyn, jolla kartoitetaan heidän tietämystänsä käyttäjien manipuloinnista. Pelin loputtua pelaajat täyttävät palautelomakkeen ja antoivat suullista palautetta.

Hafner ym. (2023) peli saa molemmilta testiryhmiltä varsin hyvät arvosanat. Peli toimii hyvänä tapana perehtyä käyttäjien manipulointiin ilmiönä. Peli vaikuttaa todenmukaiselta niin hahmojen kuin skenaarion puolesta ja pelin kulku on sujuvaa. Molemmat testiryhmät pelasivat saman kampanjan, jossa he hyökkäävät pieneen startup-yritykseen. Skenaario on pelaajien mielestä turhan helppo. Tutkijat havaitsevat pelissä selvän oppimisvaikutuksen, vaikka peli ei onnistunut muuttamaan pelaajien asenteita manipulointia kohtaan. Tutkijat arvelevat, että asenteiden muuttamiseksi vaaditaan useampia pelikertoja eri skenaarioilla.

3.5.3 Quishing ja pelillistäminen

Myös Sharevski ym. (2022) pohtivat tutkimuksessaan pelillistämisen mahdollisuuksia quishingin torjunnassa. QR-koodeja käytettäessä pelin ei tarvitse rajoittua vain digitaalisiin ympäristöihin, vaan sitä voidaan helposti laajentaa fyysiseen ympäristöön. Tällaisesta pelistä voidaan käyttää termiä *Alternate Reality Game (ARG)*. Pelaaja voi pelin alussa omaksua esimerkiksi tutkijan aseman, joka etsii ympäristössään olevia QR-koodeja, skannaa niitä, arvioi luotettavuutta ja etsii mahdollisia merkkejä QR-koodien peukaloinnista. Tutkijat toteavat myös, että tämän kaltainen peli on mahdollista laajentaa myös pidempikestoiseksi quishingin vastaiseksi koulutukseksi organisaation sisällä. Organisaation tiloihin voidaan asettaa näkyville tai henkilöstölle voidaan silloin tällöin lähettää sähköpostitse QR-koodeja. Henkilöstön tulee tunnistaa harhaanjohtavat koodit tai päätyä esimerkiksi tietoturvan tärkeydestä muistuttavalle sivustolle.

4. Pohdinta

Tämän työn tavoitteena oli tutkia, millaisia eri käyttäjien manipuloinnin muotoja on olemassa ja kuinka niitä vastaan voidaan suojautua. Työtä tehdessä selvisi myös, kuinka yleistä manipulointi on ja kuinka suurta vahinkoa sillä saadaan aikaan maailmanlaajuisesti.

Käyttäjien manipulointi voidaan siis jakaa kolmeen eri kategoriaan: kaksisuuntainen, yksisuuntainen ja epäsuora vaikuttaminen. Kaksisuuntaisessa vaikuttamisessa hyökkääjä ja uhri ovat vuorovaikutuksessa keskenään esimerkiksi puhelimen välityksellä. Yksisuuntaisessa vaikuttamisessa hyökkääjä lähettää uhrille esimerkiksi sähköposti- tai tekstiviestin. Epäsuorassa vaikuttamisessa hyökkääjä ja uhri eivät ole suorassa vuorovaikutuksissa keskenään.

Sähköpostin välityksellä tapahtuva tietojen kalastelu, eli phishing, on hyvin yleinen käyttäjien manipuloinnin muoto. Huijauskampanjoiden valmistaminen on valmiilla työkaluilla varsin helppoa ja viestejä voidaan lähettää pienellä vaivalla suuria määriä. Vaikka phishing ei ole enää ilmiönä kovin uusi, se on edelleen toimiva hyökkäyskeino. Vishing taas on puhelimitse tapahtuvaa kalastelua. Siinä on helpompi luoda uhrille aikapainetta. QR-koodien suosio on kasvanut hurjasti koronavuosien aikana, eikä se jäänyt rikollisilta huomaamatta. QR-koodien kautta tapahtuvasta kalastelusta käytetään termiä quishing. QR-koodien suosio on kasvanut valtavasti koronpandemian aikana, koska ne tarjoavat helpon, nopean ja kosketusvapaan tavan välittää dataa. Tämä ei ole jäänyt rikollisilta huomaamatta, vaan QR-kooditkin on valjastettu rikollisiin tarkoituksiin. QR-koodien välityksellä tapahtuvat hyökkäykset ovat vielä varsin uusi asia, eikä niistä ole vielä paljonkaan tietoa koulutusmateriaaleissa.

Tärkein keino käyttäjien manipuloinnin ehkäisemiseksi on koulutus. Käyttäjille on kerrottava manipuloinnin eri muodoista ja niille ominaisista piirteistä. Lisäksi käyttäjien on ymmärrettävä, että käytännössä mitä tahansa viestintäkanavaa voidaan käyttää hyväksi manipulointitarkoituksiin. Käyttäjien on tiedostettava, että kuka tahansa voi joutua manipuloinnin uhriksi ja että kaikilla on hallussaan tietoa, joka on arvokasta varkaille. Varsinkin organisaatioiden tulee kiinnittää huomiota työntekijöidensä valistukseen. Eräs keino tähän on lähettää työntekijöille sähköpostiviestejä, joissa on linkki manipuloinnista valistavaan materiaaliin. Samankaltaista tietoiskuja tulisi kehittää myös puhelimitse ja QR-koodien kautta tapahtuvaa manipulointia vastaan.

Toimivan koulutuksen kehittäminen on haastavaa. Koulutuksen kehittäjän on hyvä olla tietoinen eri lähteistä, jotka asettavat rajoituksia ja haasteita koulutuksen kehittämiselle. Suurin haaste koulutukselle on määrärahojen vähyys organisaatioissa. Kuten tavalliset käyttäjät, myös organisaatioiden johdossa olevien tulisi ymmärtää, kuinka suuri riski käyttäjien manipulointi voi olla organisaatiolle, ja kuinka paljon halvemmaksi koulutus voi tulla kuin mahdollinen onnistunut hyökkäys. Onnistunut hyökkäys on aina mainehaitta organisaatiolle, joka vaikuttaa negatiivisesti sen liiketoimintaan. Pahimmassa tapauksessa koko organisaation olemassaolo voi olla uhattuna onnistuneen hyökkäyksen seurauksena.

Koulutuksessa tulee myös kiinnittää huomiota siihen, mitkä ovat ajankohtaisia ilmiöitä käyttäjien manipuloinnissa ja kehittää koulutusta sen mukaisesti. Lisäksi tulee kiinnittää huomiota myös siihen, että koulutus ei muutu pelonlietsonnaksi. Pelko manipuloinnin uhriksi joutumisesta voi olla myös hyökkäyksen uhriksi joutumiselle altistava tekijä.

Yksi hyvä keino koulutuksen tehostamiseksi on koulutuksen pelillistäminen. Se tukee käyttäjien manipuloinnin vastaisen teoriaopetuksen sisäistämistä. Pelit ovat turvallisia eikä niihin liity tietoturvakouluttajan koulutustarkoituksessa järjestämään hyökkäyskampanjaan liittyviä eettisiä ongelmia. Tässä työssä tarkasteltiin kahta hyvin erilaista opetuspeliä. Toinen oli tietokoneella yksin pelattava korttipeli ja toinen pöytäroolipeli, joka vaatii useamman pelaajan ja pelin ohjaajan. Tulevaisuudessa varmasti nähdään lisää tutkimuksia aiheesta ja uusia koulutustarkoituksiin soveltuvia pelejä. On hyvä asia, että kehitetään monipuolisesti erilaisia pelejä. Esimerkiksi tässä työssä tarkasteltu pöytäroolipeli on erittäin opettavainen, mutta voi joidenkin ihmisten mielestä tuntua liian aikaa vievältä ja liimaa uuden opettelua. Tietokoneella pelattava korttipeli voi tässä tapauksessa olla parempi vaihtoehto. Aikaisemminkin pöytäroolipelejä pelanneille puolestaan pöytäroolipeli voi olla huomattavasti mielenkiintoisempi vaihtoehto.

Käyttäjien manipuloinnin torjuntaan on kehitetty myös monenlaisia teknisiä ratkaisuja. Tutkimukset osoittavat, että siihen tarvittavaa teknologiaa on olemassa ja sitä on mahdollista jatkokehittää. Tekniset ratkaisut mitä luultavammin tulevat aluksi organisaatioiden käyttöön ja toivottavasti myöhemmin myös yksityishenkilöille ja kuluttajille helposti saataville. Tässä työssä tutkailtiin myös yhtä tutkimusta, jossa tutkittiin sosiaalisten robottien käyttöä manipuloinnin torjunnassa. Ajatus robotista varoittamassa tietoturvauhista on erittäin mielenkiintoinen ja jännittävä, mutta luultavasti mitään tällaista ei tulla näkemään ihan lähitulevaisuudessa. Kuten tutkimuksessakin todettiin, aihe vaatii vielä jatkotutkimuksia. Tämän kaltainen tuote luultavasti tulisi aluksi lähinnä yrityskäyttöön, mutta on toki mahdollista, että se tulisi ajan kuluessa myös kuluttajien saataville.

Tekoäly tuo myös oman lisänsä käyttäjien manipulointiin. Sen avulla hyökkääjät kykenevät luomaan entistä ovelampia ja aidomman oloisia hyökkäyksiä. Tekoälyn avulla voidaan tehtailla kalasteluviestejä ja sitä voidaan käyttää apuna haittaohjelmien ja kalastelusivustojen luomisessa. Puhelinhyökkäyksissä tekoälyä voidaan käyttää generoimaan ääntä. Jo nyt on uutisoitu tapauksista, joissa hyökkääjä on ensin kouluttanut tekoälyn puhumaan uhrin lapsen äänellä sosiaalisesta mediasta löytyvien videoiden avulla. Hyökkääjä soittaa uhrille ja pyytää lapsen äänellä lähettämään rahaa rikollisten hallinnoimalle pankkitilille. Tekoälyn avulla kyetään myös toteuttamaan entistä uskottavampia toimitusjohtajahuijauksia.

Tekoälyä voidaan myös hyödyntää manipulointihyökkäysten torjunnassa. Tekoälyä voidaan hyödyntää esimerkiksi väärennettyjen sähköpostien ja verkkosivujen tunnistamiseen. Luultavasti tullaan myös näkemään tekoälysovellutuksia, joka kykenee tunnistamaan tekoälyllä luodun äänen ja osaa varoittaa mahdollisista huijauspuheluista.

Aiheesta on tehty valtava määrä tutkimusta ja lisää tehdään jatkuvasti. Jatkotutkimusaiheita tälle työlle voisivat olla esimerkiksi tekoälyn luomat uhat ja torjuntamahdollisuudet käyttäjien manipuloinnissa, manipulointia ehkäisevien algoritmien tehokkuuden vertailu, paremman koulutuksen ja pelillistämisen kehittäminen.

5. Yhteenveto

Tässä tutkielmassa perehdyttiin Käyttäjien manipuloinnin eri muotoihin ja sen torjuntaan. Aiemmista tutkimuksista käy ilmi, että ilmiö on erittäin yleinen ja sen aiheuttaa vuosittain suurta vahinkoa niin organisaatioille kuin yksityishenkilöille.

Käyttäjien manipulointi voidaan kolmeen kategoriaan kaksisuuntaiseen, yksisuuntaiseen ja epäsuoraan. Kaksisuuntaisessa manipuloinnissa hyökkääjä ja uhri ovat vuorovaikutuksessa keskenään esimerkiksi puhelimen välityksellä. Yksisuuntaisessa manipuloinnissa hyökkääjä viestittää uhrille. Epäsuorassa hyökkäyksessä hyökkääjän ja uhrin välillä ei ole vuorovaikutusta (Mouton ym., 2016). Käyttäjien manipuloinnissa tavoitteena on herättää uhrissa voimakkaita tunteita, kuten pelkoa, hämmennystä tai uteliaisuutta (Harris ym., 2021).

Sähköpostitse tapahtuvaa manipulointia kutsutaan termillä phishing. Tällöin hyökkääjä lähettää uhrille sähköpostiviestin, joka sisältää linkin rikollisten hallinnoimalle, aidon näköiselle verkkosivustolle, jonne uhri houkutellaan syöttämään tietonsa (Rao ym., 2020). Viesteissä saatetaan myös levittää haitallisia liitetiedostoja, joiden avaaminen saastuttaa uhrin päätelaitteen haittaohjelmalla (Carella ym., 2017).

Puhelimen kautta tapahtuvista hyökkäyksistä käytetään termiä vishing. Näissä hyökkäyksissä on tyypillistä, että hyökkääjä esiintyy kohdeorganisaation ulkopuolisena tahona, jolla on tarve päästä käsiksi organisaaton arkaluontoiseen tietoon. Puhelinhuijauksissa ei tyypillisesti yritetä saada uhria pitämään hyökkääjästä ystävällisillä eleillä, vaan kiireen tunne ja nopeiden toimien vaatiminen ovat yleisiä. (Jones ym., 2020).

COVID-19-pandemian aikaan QR-koodien käyttö lisääntyi helppona ja kosketusvapaana tapana jakaa tietoja ja linkkejä. Tämän havaitsivat myös rikolliset, ja kalastelusivustoille johtavien QR-koodien määrä lisääntyi valtavasti. Ilmiö on vielä suhteellisen uusi, eikä siitä valistavaa materiaalia ole vielä paljoa tarjolla (Sharevski ym., 2022).

Abroshan ym. (2021b) ovat tutkineet COVID-19-pandemian vaikutusta Käyttäjien manipulointiin. Tutkimuksessa todettiin, että pandemian aikana käyttäjien manipulointi on yleistynyt ja perinteisten phishing-hyökkäysten rinnalle on tullut uusia COVID-19-teemaisia hyökkäyksiä. Tutkimuksessa todettiin myös, että pandemian aiheuttama stressi, pelko ja ahdistus ovat lisänneet hyökkäysten onnistumismääriä.

Sähköpostihuijausten torjuntaan on kehitetty useampaan eri tekniikkaan perustuvia teknisiä ratkaisuja. Rao ym. (2020) kehittivät tutkimuksessaan uuden URL-osoitteiden analysointiin perustuvan huijaustenestojärjestelmän. Sen etuja ovat nopeus ja turvallisuus.

Puhelinhyökkäysten torjuminen teknisesti on vaikeampaa. Mouton ym. (2016) kehittivät käyttäjien manipuloinnin torjuntaan tarkoitetun mallin, jonka toinen versio sopii kaikkien kolmen erityyppisten huijausten torjuntaan. Myöhemmässä tutkimuksessa Mouton ym. (2017) kehittivät malliin perustuvan puhelinsovelluksen. Tutkijoiden omassa tutkimuksessa puhelinsovelluksen käyttö vähensi kaksisuuntaisten ja epäsuorien hyökkäysten määriä merkittävästi.

Derakhshan ym. (2021) ovat kehittäneet puhelimen välityksellä tapahtuvien hyökkäysten tunnistamiseen ja torjuntaan soveltuvan työkalun nimeltään Anti-Social Engineering Tool

(ASsET). Se on luonnollisen kielen prosessointiin perustuva algoritmi, joka kykenee tunnistamaan hyökkäyksille tyypillisiä huijaustunnusmerkkejä. Tutkijoiden omassa testissä algoritmi toimi erittäin tehokkaasti.

Koulutus on tehokkain tapa torjua käyttäjien manipulointia. Tehokkaan koulutuksen järjestäminen ei kuitenkaan ole täysin ongelmatonta. Aldawood ym. (2019) ovat tutkimuksessaan määrittäneet koulutusta vaikeuttavia tekijöitä. Suurin ongelma on organisaatioiden haluttomuus järjestää koulutusta, koska manipuloinnin aiheuttamia uhkia ei aina täysin ymmärretä. Koulutuksen tulee olla myös rakenteeltaan sellaista, että tietoturva toimii kaikkien organisaation jäsenten välillä, eikä vain niiden henkilöiden välillä, joiden kanssa koulutuksessa harjoiteltiin.

Koulutuksen pelillistäminen on mainio tapa tehostaa manipuloinnin vastaista koulutusta. Siinä koulutettavat pääsevät tutustumaan erilaisiin manipulointimenetelmiin, kuinka rikolliset laativat manipulointikampanjoita ja miten manipulointiyritykset voidaan tunnistaa. Pelillistämiseen ei myöskään liity eettisiä ongelmia, joita liittyy esimerkiksi kouluttajan koulutusmielessä toteuttamaan manipulointikampanjaan (Hafner ym., 2023). Manipulointia vastaan kouluttavia pelejä on jo olemassa jonkin verran ja tulevaisuudessa nähdään luultavasti lisää ja erilaisia. Tässä työssä tutustuttiin kahteen hyvin erilaiseen peliin. Persuaded on Aladawy ym. (2018) kehittämä tietokoneella pelattava korttipeli, jonka tavoitteena on olla helposti opittava, tehokas ja uudelleenpelattavuus sattumanvaraisuuden kautta. Hafner ym. (2023) kehittivät tutkimuksessaan suosittuun Dungeons & Dragons -pöytäroolipeliin perustuvan opetuspelin, jossa pelaajat pelaavat tiimissä ja pelin vetäjä ohjaa tarinan kulkua. Pelissä omaksutaan hyökkääjän rooli ja yritetään saavuttaa peliskenaariossa määritellyt tavoitteet. Pelin immersion lisäämiseksi pelissä käytetään myös Lego-palikoista rakennettuja rakennuksia, joita pelaajien on tarkoitus päästä tutkimaan pelin edetessä.

Myös robottien käyttöä manipuloinnin torjunnassa on tutkittu. Ihmisillä on taipumus luottaa robotteihin epävarmoissa tilanteissa. Loogisin perusteluin käyttäjäänsä ohjaava robottikumppani vaikutti tutkimuksen mukaan varsin tehokkaalta tavalta suojella käyttäjää mahdollisilta uhilta (Pasquali ym., 2023).

Lähdeluettelo

Abroshan, H., Devos, J., Poels, G., & Laermans, E. (2021a). A phishing Mitigation Solution using Human Behaviour and Emotions that Influence the Success of Phishing Attacks. *UMAP 2021 - Adjunct Publication of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 345–350. https://doi.org/10.1145/3450614.3464472

Abroshan, H., Devos, J., Poels, G., & Laermans, E. (2021b). COVID-19 and Phishing: Effects of Human Emotions, Behavior, and Demographics on the Success of Phishing Attempts during the Pandemic. *IEEE Access*, *9*, 121916–121929. https://doi.org/10.1109/ACCESS.2021.3109091

Abroshan, H., Devos, J., Poels, G., & Laermans, E. (2021c). Phishing Happens beyond Technology: The Effects of Human Behaviors and Demographics on Each Step of a Phishing Process. *IEEE Access*, *9*, 44928–44949. https://doi.org/10.1109/ACCESS.2021.3066383

Aladawy, D., Beckers, K., & Pape, S. (2018). PERSUADED: Fighting social engineering attacks with a serious game. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11033 LNCS*, 103–118. https://doi.org/10.1007/978-3-319-98385-1_8/FIGURES/6

Aldawood, H., & Skinner, G. (2019). Challenges of implementing training and awareness programs targeting cyber security social engineering. *Proceedings - 2019 Cybersecurity and Cyberforensics Conference, CCC 2019*, 111–117. https://doi.org/10.1109/CCC.2019.00004

Carella, A., Kotsoev, M., & Truta, T. M. (2017). Impact of security awareness training on phishing click-through rates. *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, *2018-January*, 4458–4466. https://doi.org/10.1109/BIGDATA.2017.8258485

Derakhshan, A., Harris, I. G., & Behzadi, M. (2021). Detecting telephone-based social engineering attacks using scam signatures. *IWSPA 2021 - Proceedings of the 2021 ACM Workshop on Security and Privacy Analytics*, 67–73. https://doi.org/10.1145/3445970.3451152

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 581–590. https://doi.org/10.1145/1124772.1124861

Hafner, L., Wutz, F., Pöhn, D., & Hommel, W. (2023). TASEP: A Collaborative Social Engineering Tabletop Role-Playing Game to Prevent Successful Social Engineering Attacks. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3600160.3605005

Harris, I. G., Derakhshan, A., & Carlsson, M. (2021). A Study of Targeted Telephone Scams Involving Live Attackers. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12812 LNCS*, 63–82. https://doi.org/10.1007/978-3-030-79318-0_4/FIGURES/4

Jones, K. S., Armstrong, M. E., Tornblad, M. K., & Siami Namin, A. (2020). How social engineers use persuasion principles during vishing attacks. *Information and Computer Security*, *29*(2), 314–331. https://doi.org/10.1108/ICS-07-2020-0113

Mattera, M., & Chowdhury, M. (2021). Social Engineering: The Looming Threat. *2021 IEEE International Conference on Electro Information Technology (EIT)*. https://doi.org/10.1109/EIT51626.2021.9491884

Mouton, F., Leenen, L., & Venter, H. S. (2016). Social Engineering Attack Detection Model: SEADMv2. *Proceedings - 2015 International Conference on Cyberworlds, CW 2015*, 216–223. https://doi.org/10.1109/CW.2015.52

Mouton, F., Malan, M. M., Leenen, L., & Venter, H. S. (2014). Social engineering attack framework. *Information Security for South Africa - Proceedings of the ISSA 2014 Conference*. https://doi.org/10.1109/ISSA.2014.6950510

Mouton, F., Teixeira, M., & Meyer, T. (2017). Benchmarking a mobile implementation of the social engineering prevention training tool. *2017 Information Security for South Africa - Proceedings of the 2017 ISSA Conference*, *2018-January*, 106–116. https://doi.org/10.1109/ISSA.2017.8251782

Pasquali, D., Kothig, A., Aroyo, A. M., Muñoz Cadorna, J. E., Dautenhahn, K., Bencetti, S., Francesco, R., & Sciutti, A. (2023). That's not a Good Idea: A Robot Changes Your Behavior Against Social Engineering. *ACM International Conference Proceeding Series*, 63–71. https://doi.org/10.1145/3623809.3623879

Rao, R. S., Vaishnavi, T., & Pais, A. R. (2020). CatchPhish: detection of phishing websites by inspecting URLs. *Journal of Ambient Intelligence and Humanized Computing*, *11*(2), 813–825. https://doi.org/10.1007/S12652-019-01311-4/FIGURES/4

Sharevski, F., Devine, A., Pieroni, E., & Jachim, P. (2022). Phishing with Malicious QR Codes. *ACM International Conference Proceeding Series*, *2022*, 160–171. https://doi.org/10.1145/3549015.3554172

Uusitalo, H. (2021, syyskuuta 2). *Oulun yliopistolle tehtiin mittava tietojen kalastelu: Opiskelijoiden ja henkilöstön salasanoja joutui vääriin käsiin – "Poikkeustilanne" | Kaleva*. https://www.kaleva.fi/oulun-yliopistolle-tehtiin-mittava-tietojen-kalast/3925250

Wang, Z., Zhu, H., & Sun, L. (2021). Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods. *IEEE Access*, *9*, 11895–11910. https://doi.org/10.1109/ACCESS.2021.3051633