TURUN YLIOPISTO
Kauppakorkeakoulu

Väärän tiedon leviäminen sosiaalisessa mediassa

Tietojärjestelmätieteen kandidaatintutkielma

Laatija:
Naomi Reisner

Ohjaaja:
KTM Tarja Matikka

27.4.2026
Turku

Opiskelijan lausunto tekoälyn käytöstä tähän tutkielmaan liittyen:

☐ **En ole käyttänyt tekoälyä hyödyntäviä työkaluja** tätä tutkielmaa kirjoittaessani.

☒ **Olen käyttänyt tekoälyä hyödyntäviä työkaluja** tätä tutkielmaa kirjoittaessani. Tämä käyttö on dokumentoitu tutkielman liitteessä. Vakuutan, että tekoälyä käytettiin yliopiston ohjeistuksen mukaisella tavalla.

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

Kandidaatintutkielma

Oppiaine: Tietojärjestelmätiede
Tekijä: Naomi Reisner
Otsikko: Väärän tiedon leviäminen sosiaalisessa mediassa
Ohjaaja: KTM Tarja Matikka
Sivumäärä: 28 sivua + liitteet 1 sivu
Päivämäärä: 27.4.2026

Tiivistelmä

Sosiaalista mediaa käytetään yhä useammin tiedonhaun ensisijaisena kanavana, mutta sen rakenteelliset ja teknologiset piirteet voivat kiihdyttää väärän tiedon leviämistä. Algoritmit synnyttävät somekuplia ja kaikukammioita, joissa käyttäjä altistuu toistuvasti virheelliselle tiedolle. Toistuva altistuminen lisää tiedon uskomisen todennäköisyyttä.

Tämä kirjallisuuskatsaus tarkastelee, miten väärä tieto leviää sosiaalisessa mediassa ja millaisin keinoin leviämistä voidaan ehkäistä. Tutkimuskysymyksiä lähestytään käyttäjän ja teknologian ominaisuuksien sekä alustojen käytössä olevien torjuntakeinojen kautta.

Tutkielma osoittaa, että kognitiiviset vinoumat ja heuristiikat, uutuusarvoa ja negatiivisia tunteita korostavat sisällöt sekä alustojen sitoutumista maksimoivat algoritmit muodostavat yhteisvaikutuksen, joka lisää väärän tiedon uskottavuutta ja jakamishalukkuutta.

Ehkäisyssä voidaan käyttää alusta- ja käyttäjätason toimia, kuten varoituksia, näkyvyyden hallintaa, faktantarkistusta ja tekoälypohjaista tunnistamista sekä käyttäjiä tarkkuuteen ohjaavia keinoja. Yksiselitteistä tehokkainta keinoa ei voida nimetä, sillä vaikutukset ovat kontekstisidonnaisia ja riippuvat ajoituksesta ja kohderyhmästä.

Jatkotutkimuksessa tulisi kartoittaa alustakohtaisia käytäntöjä sekä arvioida laajemmin, miten käyttäjien demografiset tekijät ja persoonallisuuden piirteet vaikuttavat ehkäisykeinojen tehoon.

Avainsanat: misinformaatio, disinformaatio, algoritminen vinouma, sosiaalinen media, viraalius, suosittelujärjestelmät

SISÄLLYS

TAULUKOT

1 Johdanto

Sosiaalinen media on nykyään erottamaton osa ihmisten arkea. Sillä on merkittävä rooli tiedon muodostumisessa ja leviämisessä (Hassani ym., 2024). Samalla sosiaalisessa mediassa esiintyvän mis- ja disinformaation sekä valeuutisten määrä on herättänyt laajaa huolta. World Economic Forum on nostanut väärän tiedon leviämisen yhdeksi merkittävimmistä uhista ihmiskunnalle (Törnberg, 2018). Sosiaalisessa mediassa leviävän väärän tiedon sekä kaikukammioiden ja somekuplien on pelätty heikentävän demokraattisia arvoja, kuten sananvapautta (van der Puil ym., 2023). Mis- ja disinformaatio sekä valeuutiset vaikuttavat yhteiskunnan eri osa-alueisiin aina politiikasta ja terveydenhuollosta talouteen ja osakekursseihin (Vosoughi ym., 2018; Hassani ym., 2024).

Koronapandemian aikana sosiaalisessa mediassa levinnyt väärä tieto ja salaliittoteoriat olivat yhteydessä lisääntyneisiin koronatartuntoihin sekä leviämisen estämiseen suunnattujen toimenpiteiden heikompaan vaikuttavuuteen (Wang ym., 2026). Laajalle levinnyt väite metanolin COVID-19:ää parantavasta vaikutuksesta johti lukuisten ihmisten sokeutumiseen, ja 5G-verkkoihin liitetyt viruksen leviämistä koskevat salaliittoteoriat puolestaan aiheuttivat televiestintäinfrastruktuuriin kohdistuneita omaisuusvahinkoja (Shirish ym., 2021). Ukrainan ja Venäjän sekä Israelin ja Palestiinan sodassa visuaalinen mis- ja disinformaatio on puolestaan vaikuttanut ihmisten mielipiteisiin ja yhteiskunnalliseen jakautumiseen sekä toiminut väkivallan oikeuttamisen välineenä (Hameleers, 2025). Pahimmillaan väärä tieto voi myös yllyttää väkivaltaan (Hassani ym., 2024).

Ihmisten turvautuessa yhä useammin sosiaalisen median välittämään tietoon on tärkeää, että mis- ja disinformaation sekä valeuutisten leviämistä pyritään minimoimaan. Sisällön moderointi ja väärän tiedon leviämisen rajoittaminen ovat kuitenkin pitkälti sosiaalisen median alustoja omistavien yritysten käsissä (Keswani, 2023). Alustojen liiketoimintamalli perustuu usein mainontaan ja suosittelujärjestelmiin, ja keskeisiä tavoitteita ovat käyttäjien huomion ja sitoutumisen maksimointi (Keswani, 2023; Gsenger, 2025). Samanaikaisesti tutkimuskirjallisuus kuvaa moderoinnin ja faktantarkistuksen käytäntöjä vaihteleviksi ja osin rajallisiksi (Hameleers, 2025).

Tutkimuksen mukaan väärä tieto leviää sosiaalisessa mediassa usein laajemmalle kuin totuudenmukainen tieto (Vosoughi ym., 2018). Lisäksi oikaisut eivät monesti saavuta samaa yleisöä riittävän nopeasti, mikä korostaa ennaltaehkäisevien keinojen merkitystä (Hassani ym., 2024). Tämän tutkielman tarkoituksena onkin tarkastella väärän tiedon leviämistä sosiaalisessa mediassa sekä keinoja, joilla sen leviämistä voidaan ehkäistä.

Tutkielma toteutetaan kirjallisuuskatsauksena, jossa aiemmasta tutkimuksesta tehdään synteesi ja arvioidaan tulosten soveltuvuutta nykyiseen sosiaalisen median ympäristöön.

Tutkielman tutkimuskysymykset ovat seuraavat:

1. Mitkä yksilölliset tekijät vaikuttavat siihen, että ihmiset uskovat ja jakavat väärää tietoa?
2. Miten sosiaalisen median suosittelumekanismit vaikuttavat väärän tiedon leviämiseen?
3. Mitkä ovat tehokkaimpia keinoja ehkäistä väärän tiedon leviämistä sosiaalisessa mediassa?

Tutkielmassa tarkastellaan väärän tiedon leviämistä ensisijaisesti käyttäjän näkökulmasta: millä tavoin sisältö tavoittaa käyttäjiä ja miksi sitä jaetaan? Torjuntakeinoja käsitellään alustayritysten näkökulmasta. Tutkielma painottuu alusta- ja käyttäjätason keinoihin eikä esimerkiksi sääntelyyn tai mediakasvatukseen.

Tässä tutkielmassa ”väärä tieto” käsittää mis- ja disinformaation sekä valeuutiset. Käsitteet eroavat toisistaan tarkoitusperän osalta: misinformaatio viittaa tahattomasti jaettuun virheelliseen tietoon, kun taas disinformaatio on tarkoituksellisesti tuotettua tai jaettua virheellistä tietoa. Myös valeuutiset sisältävät harhaanjohtavia tai tahallisesti vääristeltyjä elementtejä. (Bakir & McStay, 2018; Olenberger ym., 2026.) Nämä kaikki sisällytetään tässä tutkielmassa käsitteeseen ”väärä tieto”, sillä tarkoitusperä ei ole keskiössä tarkasteltaessa tiedon mahdollisesti aiheuttamaa haittaa.

Ensimmäinen käsittelyluku tarkastelee väärän tiedon leviämistä käyttäjän ominaisuuksien näkökulmasta. Luvussa analysoidaan kognitiivisia tekijöitä, jotka altistavat käyttäjiä pitämään virheellistä tietoa uskottavana, sekä käyttäjätyyppeihin ja persoonallisuuden piirteisiin liittyviä eroja väärän tiedon uskomisessa ja jakamisessa. Lisäksi luvussa käsitellään sisällöllisiä piirteitä, jotka lisäävät väärän tiedon vakuuttavuutta.

Toinen luku tarkastelee väärän tiedon leviämistä sosiaalisen median rakenteellisten tekijöiden näkökulmasta. Luvussa kuvataan sosiaalisen median suosittelumekanismeja, algoritmeja ja suunnitteluratkaisuja sekä analysoidaan, miten nämä voivat edistää väärän tiedon leviämistä.

Kolmas luku keskittyy väärän tiedon leviämisen ehkäisemiseen. Luvussa tarkastellaan sosiaalisen median alustojen käyttämiä ehkäisykeinoja ja käyttäjätason keinoja, joilla väärän tiedon jakamista voidaan vähentää. Johtopäätöksissä kootaan keskeiset havainnot ja esitetään jatkotutkimuksen tarpeita.

2 Yksilölliset tekijät

Tässä luvussa tarkastellaan yksilöllisiä tekijöitä, jotka vaikuttavat väärän tiedon uskomiseen ja jakamiseen. Luvussa käsitellään erityisesti kognitiivisia tekijöitä, sisältöön liittyviä piirteitä sekä käyttäjien persoonallisuuteen ja käyttäjätyyppeihin liittyviä eroja.

Ihmiset etsivät tietoa useista eri lähteistä, ja sosiaalinen media on muodostunut keskeiseksi alustaksi tiedon jakamiselle ja omaksumiselle (Olenberger ym., 2026). Väärä tieto leviää sosiaalisessa mediassa totuudenmukaista tietoa nopeammin, laajemmalle ja syvemmälle (Vosoughi ym., 2018). Ihmisten kyky tunnistaa väärää tietoa on usein puutteellinen, mikä altistaa heidät omaksumaan tietoa ilman kriittistä arviointia. Riittämätön laadunvalvonta mahdollistaa väärän tiedon nopean leviämisen, ja ajan myötä tällaisesta tiedosta voi muodostua kulttuuriin juurtuvia myyttejä. (Olenberger ym., 2026.)

Väärän tiedon leviämiseen vaikuttavat kuitenkin myös laajemmat yhteiskunnalliset olosuhteet. Poliittisen vapauden on todettu olevan kytköksissä väärän tiedon leviämisen kasvuun, sillä se mahdollistaa poliittisten toimijoiden pyrkimykset vaikuttaa mielipiteisiin harhaanjohtavan tiedon avulla. Laaja mobiiliyhteyksien saatavuus puolestaan nopeuttaa sisällön kulutusta ja jakamista sekä siirtää käyttäjien huomion tiedon laadusta sen määrään. Sen sijaan taloudellinen vapaus ja median vapaus vähentävät väärän tiedon leviämistä. Vakaat ja ennakoitavat taloudelliset instituutiot lisäävät kansalaisten luottamusta yhteiskunnallisiin toimijoihin ja vähentävät tarvetta etsiä tai jakaa epävarmaa tietoa. Median vapaus mahdollistaa riippumattoman tiedonvälityksen sekä viranomaisten ja poliittisten toimijoiden toiminnan kriittisen tarkastelun, mikä heikentää virheellisen tiedon uskottavuutta ja leviämistä. (Shirish ym., 2021.) Nämä rakenteelliset tekijät osaltaan selittävät, miksi väärä tieto leviää eri tavoin eri yhteiskunnissa.

Kansalaiset altistuvat tunnepitoiselle ja provosoivalle väärälle tiedolle, jota on vaikea oikaista erityisesti kaikukammioiden vuoksi (Bakir & McStay, 2018). Kaikukammio on ympäristö, jossa samanmielisten käyttäjien vuorovaikutus vahvistaa heidän aiempia uskomuksiaan (Villa ym., 2021; Impicciché & Viviani, 2025). Somekupla viittaa puolestaan ympäristöön, jossa algoritmien personoima sisältö altistaa käyttäjän hänen omia uskomuksiaan ja mieltymyksiään vahvistavalle tiedolle. Toisin kuin kaikukammion tapauksessa, somekuplan synty ei edellytä aktiivista vuorovaikutusta muiden käyttäjien kanssa. (Impicciché & Viviani, 2025.)

2.1 Kognitiiviset tekijät

Useat tekijät heikentävät ihmisten kykyä tunnistaa väärää tietoa: aikaisempi altistuminen väärälle tiedolle, analyyttisen ajattelun puute, tunneperäinen sitoutuminen ja mieliala. Sisältö arvioidaan luotettavammaksi silloin, kun se on yhdenmukaista omien uskomusten, mielipiteiden ja ajatusmallien kanssa. (Olenberger ym., 2026.)

Väärän tiedon uskottavuuteen vaikuttavat myös kognitiiviset vinoumat, kuten vahvistusharha, ankkurointivaikutus, luottamusharha, kehystysvaikutus ja laumamieliala. Kognitiiviset vinoumat ovat epärationaalisia ajattelutapoja, jotka heikentävät kykyä tunnistaa harhaanjohtamista. Vahvistusharha viittaa omia uskomuksia tukevan tiedon suosimiseen. Ankkurointivaikutus tarkoittaa yksittäiseen tietoon takertumista, luottamusharha liiallista uskoa omaan tietämykseen, kehystysvaikutus tiedon esitystavan vaikutusta tulkintaan ja laumamieliala muiden toiminnan kopioimista. Useimmat käyttäjät eivät tarkista tiedon paikkansapitävyyttä, vaan luottavat sen näennäiseen johdonmukaisuuteen, mikä mahdollistaa kognitiivisten vinoumien hyödyntämisen väärän tiedon leviämisessä. (French ym., 2025.)

Vahvistusharhan on havaittu lisäävän myös väärän tiedon näkyvyyttä ja jakamista. Julkisuudenhenkilön esittämä, vastaanottajan uskomuksia tukeva väärä tieto kasvattaa katselu- ja jakamishalukkuutta 21 %. (Seol ym., 2024.) Viraalius vaikuttaa siihen, kuinka luotettavaksi tieto koetaan: mitä laajemmalle vaikuttajan julkaisu leviää, sitä luotettavammaksi se arvioidaan. Kriittinen kommentointi voi heikentää koettua luotettavuutta, kun taas kannustavat kommentit lisäävät sitä erityisesti vähemmän viraaleissa julkaisuissa. (Mulcahy ym., 2025.)

Sisältöön sitoutuminen käynnistyy sekä kognitiivisen että tunnetason prosessoinnin kautta, jossa sisältöä arvioidaan muun muassa luettavuuden ja esteettisyyden perusteella. Viime kädessä tunne vaikuttaa kuitenkin kognitiota enemmän siihen, jakaako käyttäjä sisältöä eteenpäin. (Maasberg ym., 2018.)

2.2 Käyttäjätyypit

Kognitiivisten tekijöiden lisäksi väärän tiedon uskomiseen ja leviämiseen vaikuttavat yksilön persoonalliset ominaisuudet, jotka ohjaavat käyttäytymistä sosiaalisessa mediassa. Luonteenpiirteitä voidaan arvioida käyttäjän julkaisemien sisältöjen perusteella, ja erityisesti tunnollisuuden ja moraalisen suvaitsevaisuuden on todettu olevan yhteydessä vastuulliseen verkkokäyttäytymiseen. Tutkimus erottaa neljä arkkityyppiä: narsistin, dogmaatikon, nihilistin ja altruistin. Narsisti hakee

näkyvyyttä ja käyttäytyy vastuuttomasti esimerkiksi ylijakamalla ja trollaamalla. Dogmaatikko hakeutuu kaikukammiohin, mikä voi lisätä polarisaatiota. Nihilisti suhtautuu välinpitämättömästi vallitseviin arvoihin ja välttelee osallistumista, mikä voi johtaa haitallisen sisällön hiljaiseen hyväksymiseen. Altruisti puolestaan suosii alustoja, jotka tukevat tiedon jakamista ja merkityksellisiä keskusteluja. (Ciriello ym., 2025.)

Persoonallisuudenpiirteet selittävät alttiutta uskoa ja jakaa väärää tietoa. Tunnollisuus ja sovinnollisuus vähentävät todennäköisyyttä omaksua tai levittää väärää tietoa. Sen sijaan ulospäinsuuntautuneisuus voi lisätä taipumusta uskoa ja jakaa sitä. (Chen, 2024.) Lisäksi pohdiskelu ja avoin ajattelu vähentävät väärän tiedon omaksumista, kun taas taipumus uskoa salaliittoteorioihin ja tulkita merkityksettömiä asioita merkityksellisiksi lisäävät riskiä uskoa virheellisiin väitteisiin. (Hubeny ym., 2025.)

Taulukko 1 kokoaa yhteen keskeiset yksilölliset tekijät ja niiden vaikutusmekanismit väärän tiedon uskomiseen ja jakamiseen.

Taulukko 1. Yksilöllisten tekijöiden vaikutus väärän tiedon uskomiseen ja jakamiseen

Tekijä	Kuvaus	Vaikutus väärän tiedon uskomiseen ja jakamiseen	Lähde
Aiempi altistuminen väärälle tiedolle	Toistuva kohtaaminen väärän tiedon kanssa	Lisää tiedon koettua tuttuutta ja luotettavuutta	Olenberger ym., 2026
Analyyttisen ajattelun puute	Heikko taipumus kriittiseen arviointiin	Altistaa tiedon hyväksymiselle ilman tarkistusta	Olenberger ym., 2026
Tunneperäinen sitoutuminen	Tunteet ohjaavat päätöksentekoa	Lisää jakamishalukkuutta riippumatta tiedon paikkansapitävyydestä	Maasberg ym., 2018; Olenberger ym., 2026
Kognitiiviset vinoumat	Epärationaaliset ajattelutavat heikentävät harhaanjohtavan tiedon tunnistamista	Lisää väärän tiedon näkyvyyttä ja leviämistä	French ym., 2025
Viraalius	Laaja näkyvyys lisää koettua luotettavuutta	Kasvattaa jakamishalukkuutta	Mulcahy ym., 2025
Käyttäjätyypit	Persoonallisuuteen perustuvat käyttäytymistyylit sosiaalisessa mediassa	Vaikuttavat siihen, kuinka todennäköisesti käyttäjät uskovat ja jakavat väärää tietoa	Ciriello ym., 2025
Persoonallisuudenpiirteet	Esimerkiksi tunnollisuus tai ulospäinsuuntautuneisuus	Lisäävät tai vähentävät alttiutta uskoa ja jakaa väärää tietoa	Chen, 2024

2.3 Sisältöön liittyvät tekijät

Väärä tieto sisältää usein uutuusarvoa, mikä lisää sen jakamisen todennäköisyyttä. Uutuus herättää huomiota, helpottaa päätöksentekoa ja kannustaa tiedon jakamiseen. (Vosoughi ym., 2018; Kumari ym., 2021.) Uusi tieto sisältää usein yllätyksellisyyttä, mikä on tyypillistä viraaliksi leviävälle väärälle tiedolle (Kumari ym., 2021).

Tiedon uutuus ja ajoitus vaikuttavat ajattelumallien muovaantumiseen. Ihmiset käyttävät ajattelumalleja monimutkaisten ilmiöiden ja tapahtumien jäsentämiseen. Sosiaalisen median suuret tietomäärät johtavat nopeaan prosessointiin, minkä vuoksi käyttäjät arvioivat sisältöjä usein pintapuolisesti. Tiedon määrän vuoksi ihmiset tukeutuvat arjessa ja erityisesti sosiaalista mediaa selatessaan nopeisiin tiedonkäsittelyn strategioihin. Tieto uskotaan helposti todeksi, jos tarjolla ei ole ristiriitaista näkökulmaa. Huhupuhe omaksutaan usein nopeammin kuin oikaiseva tieto, sillä nopea prosessointi on kognitiivisesti vähemmän kuormittavaa. Uuden tiedon omaksuminen muokkaa ajattelumalleja, ja virheellisen tiedon korjaaminen jälkikäteen vaatii huomattavaa kognitiivista ponnistelua. (Olenberger ym., 2026.) Lisäksi uusi tieto voi viestiä käyttäjän sosiaalisesta asemasta, kuten sisäpiiritiedon hallinnasta, mikä lisää jakamishalukkuutta (Vosoughi ym., 2018).

Ihmiset tunnistavat multimodaalista väärää tietoa, esimerkiksi tekstiä, kuvia ja ääntä yhdistävää sisältöä, helpommin kuin pelkkään tekstiin perustuvaa väärää tietoa (Qiu & Goh, 2025). Skeema on kognitiivinen rakenne, joka ohjaa informaation tulkintaa ja auttaa jäsentämään havaintoja aikaisempien kokemusten pohjalta (Vannucci ym., 2025). Multimodaalisuus tarjoaa useampia audiovisuaalisia vihjeitä, joiden perusteella käyttäjät voivat arvioida tiedon todenperäisyyttä useiden skeemojen kautta. Lisäksi multimodaalisuus voi vähentää kommenttien ja viraaliuden vaikutusta väärän tiedon tunnistamiseen. (Qiu & Goh, 2025.)

Väärän tiedon herättämät tunteet vaikuttavat puolestaan siihen, kuinka todennäköisesti ihmiset jakavat tietoa eteenpäin. Väärä tieto herättää usein pelkoa, inhotusta ja hämmästystä. Siinä missä robotit jakavat väärää ja totuudenmukaista tietoa yhtä paljon, ihmiset välittävät väärää tietoa eteenpäin huomattavasti todennäköisemmin: väärää tietoa jaetaan Twitterissä 70 % todennäköisemmin kuin totuudenmukaista tietoa. (Vosoughi ym., 2018.) Tutkimuksissa on osoitettu, että väärä tieto sisältää usein negatiivisia tunteita ja, että moraalista tunnearvoa sisältävä tieto leviää erityisen nopeasti (Kumari ym., 2021).

Tutkimusten mukaan jopa 75 % Facebookissa jaetusta sisällöstä jaetaan ilman, että käyttäjä avaa tai lukee sitä. Tämä koskee etenkin poliittisista ääripäistä peräisin olevaa tietoa. Jakaminen perustuu

pinnallisiin vihjeisiin, kuten otsikoihin, kuviin ja tykkäysten määrään. Käyttäjät, jotka muodostavat käsityksensä uutisotsikoiden ja tiivistelmien perusteella, kokevat tietävänsä enemmän kuin todella tietävät. Ilmiötä selittävät informaatiotulva ja heuristiikkoihin turvautuminen. (Sundar ym., 2025.)

Heuristiikat ovat päätöksentekoa ja informaation prosessointia helpottavia yleistyksiä. Käyttäjät voivat turvautua esimerkiksi maineeseen ja sosiaaliseen hyväksyntään: tunnettuun ja luotettavana pidettyyn lähteeseen luotetaan herkemmin, samoin tietoon, jonka muut ovat hyväksyneet. (Metzger ym., 2010.) Tykkäysten, kommenttien ja uudelleentwiittausten määrä vaikuttavat luotettavuusarvioihin sekä halukkuuteen jakaa tietoa. Jakaminen toimii myös sosiaalisena signaalina: käyttäjä voi viestiä olevansa varma tiedosta ja halukas osallistumaan keskusteluun. Käyttäjät tuntevat olevansa sitoutuneempia tietoon, kun he jakavat sitä ja liittävät jakamisen yhteydessä ystävilleen kohdistuvan kysymyksen. (Sundar ym., 2025.) Toistuva tieto mielletään luotettavammaksi, ja väärä tieto muotoillaan usein vetoavaksi, jotta se toistuisi eri lähteissä. Altistuminen lisää luottamusta ja kiihdyttää väärän tiedon leviämistä. (Vellani ym., 2023.)

Myös harhaanjohtava argumentointi heikentää käyttäjän kykyä tunnistaa virheellistä tietoa (Olenberger ym., 2026). Argumentaatiovirheillä pyritään vaikuttamaan lukijan käsityksiin tavalla, joka ei perustu itse väitteen sisältöön vaan sen esitystapaan. Tällaisia virheitä voivat olla esimerkiksi henkilöön kohdistuva argumentointi, tilanteen kaventaminen vääräksi vastakkainasetteluksi silloin, kun todellisuudessa vaihtoehtoja on useampia, tai vetoaminen epäolennaiseen tai epäluotettavaan lähteeseen. Lisäksi tieto voi sisältää tarkoituksellisia loogisia päättelyvirheitä, joiden tavoitteena on ohjata lukijan tulkintaa harhaan. (Beisecker ym., 2024.)

Taulukko 2 kokoaa yhteen keskeiset sisältöön liittyvät tekijät ja niiden vaikutusmekanismit väärän tiedon uskomiseen ja jakamiseen.

Taulukko 2. Sisältöön liittyvien tekijöiden vaikutus väärän tiedon uskomiseen ja jakamiseen

Tekijä	Mekanismi	Vaikutus väärän tiedon uskomiseen ja jakamiseen	Lähde
Uutuusarvo	Herättää huomiota ja helpottaa päätöksentekoa	Lisää sisällön jakamisen todennäköisyyttä	Vosoughi ym., 2018; Kumari ym., 2021
Sosiaalisen median suuret tietomäärät	Pintapuolinen ja nopea prosessointi	Heikentää kriittistä arviointia ja altistaa virheille	Olenberger ym., 2026

Taulukko 2. Jatkuu

Tekijä	Mekanismi	Vaikutus väärän tiedon uskomiseen ja jakamiseen	Lähde
Multimodaalisuus	Useat audiovisuaaliset vihjeet ja skeemojen aktivointi	Helpottaa väärän tiedon tunnistamista	Qiu & Goh, 2025
Tunteisiin vetoava sisältö	Herättää pelkoa, inhotusta ja hämmästystä	Lisää jakamishalukkuutta	Vosoughi ym., 2018; Kumari ym., 2021
Viraalius (tykkäykset, jaot)	Turvautuminen sosiaaliseen hyväksyntään	Lisää koettua luotettavuutta	Metzger ym., 2010
Otsikot ja kuvat	Pinnalliset vihjeet korvaavat sisällön arvioinnin	Sisältö jaetaan ilman lukemista	Sundar ym., 2025
Toistuvuus	Tuttuuden ja luottamuksen lisääntyminen	Kiihdyttää väärän tiedon leviämistä	Vellani ym., 2023
Harhaanjohtava argumentointi	Huomion kiinnittäminen esitystapaan sisällön sijaan	Heikentää virheellisen tiedon tunnistamista	Beisecker ym., 2024

3 Suosittelumekanismit

Tässä luvussa tarkastellaan sosiaalisen median suosittelumekanismeja ja niiden vaikutusta väärän tiedon leviämiseen. Huomio kohdistuu erityisesti algoritmeihin, kaikukammioihin ja somekupliin sekä alustojen suunnitteluratkaisuihin ja käyttömahdollisuuksiin.

Viraalius lisää väärän tiedon leviämistä ja kasvattaa sen näkyvyyttä sosiaalisessa mediassa (Vosoughi ym., 2018). Viraaliuden mekanismeja tukevat useat sosiaalisen median rakenteelliset ja teknologiset piirteet, jotka vaikuttavat siihen, kuinka nopeasti ja laajasti sisältö saavuttaa yleisönsä.

3.1 Algoritmit, kaikukammiot ja somekuplat

Sosiaalisen median painopiste on siirtynyt verkostoitumisesta algoritmiseen sisällönjakoon, jonka keskeisenä tavoitteena on käyttäjien sitouttamisen maksimoiminen ja kaupallisen hyödyn tuottaminen. Algoritmit vaikuttavat sananvapauteen säätelemällä viestien näkyvyyttä ilman läpinäkyviä toimintaperiaatteita. Alustat näyttävät käyttäjille sisältöä, joka pitää heidät mahdollisimman sitoutuneina, ja vahingollisen sisällön, kuten väärän tiedon, on havaittu olevan usein koukuttavampaa kuin vähemmän äärimmäisen sisällön. (Riemer & Peter, 2021.)

Luotettavaa tietoa jaetaan laajalle käyttäjäkunnalle, kun taas epäluotettava tieto kohdentuu usein samanmielisten ryhmille. Tämän seurauksena tieto leviää kaikukammiossa, joissa käyttäjät kohtaavat harvemmin kritiikkiä tai mainehaittaa väärän tiedon levittämisestä. (Acemoglu ym., 2024.)

Kaikukammiot perustuvat tiedon toistuvuuteen. Mainostuottojen lisäämiseksi selaamisesta on suunniteltu nopeaa ja säännöllistä, mikä altistaa käyttäjän toistuvasti samalle sisällölle. Toistuva tieto alkaa ajan myötä vaikuttamaan merkityksellisemmältä ja todenmukaisemmalta. Kaikukammiot kehittyvät herkästi alustoilla, joissa algoritmit vaikuttavat olennaisesti käyttäjän näkemään sisältöön. Ne lisäävät yli-itsevarmuutta, vähentävät oppimishalua ja vahvistavat ajatusten polarisoitumista. (Liu & Fowler, 2025.)

Kaikukammiot lisäävät myös tuttuuden tunnetta ja sujuvuusharhaa, jolloin helposti ymmärrettävä tieto koetaan oikeammaksi. Alustojen algoritmit edistävät somekuplien ja kaikukammioiden syntyä, mikä tekee alustoista koukuttavia ja tuottoisia. Koska algoritmit ovat keskeisiä tuottoisuuden kannalta, on epätodennäköistä, että alustojen omistajat muuttaisivat niitä ilman ulkopuolista sääntelyä. (Rhodes, 2022.)

Toisaalta tutkimukset osoittavat, että somekuplien syntyminen ja väärän tiedon leviäminen on mahdollista myös ilman suosittelualgoritmeja. Epäluotettavat lähteet pyrkivät sitouttamaan käyttäjän ohjaamalla hänet linkkien kautta yhä uusille epäluotettaville sivuille, mikä voi johtaa käyttäjän ajautumiseen syvemmälle väärän tiedon lähteisiin. (Greene ym., 2025.)

3.2 Alustojen suunnittelu

Sosiaalisen median alustat on muotoiltu siten, että käyttäjien vuorovaikutus sisällön kanssa maksimoidaan. Sosiaalinen media on nopea ja helppokäyttöinen tiedonvälityskanava, minkä vuoksi monet käyttävät sitä ensisijaisena tiedonlähteenään. Alustojen ominaisuudet tekevät väärän tiedon leviämisestä erityisen vaivatonta. (Liu & Fowler, 2025.)

Sosiaalinen media on suunniteltu henkilökohtaisten ja tunteisiin vetoavien tarinoiden jakamiseen, ja käyttäjät ovat tottuneet kuluttamaan sisältöä hedonistisesti. Tämä vähentää tiedon kriittistä arviointia ja faktantarkistusta. (Liu & Fowler, 2025.) Passiivisesti vastaanotettu tieto hyväksytään usein helposti (Rhodes, 2022). Lisäksi rajoitettu sanamäärä johtaa tiedon yksinkertaistamiseen ja liioitteluun. Jakamistoiminnon helppous kannustaa tiedon nopeaan levittämiseen ilman faktantarkistusta. (Liu & Fowler, 2025.)

3.3 Alustojen affordanssit (käyttömahdollisuudet)

Sosiaalisen median käyttömahdollisuudet määrittävät, millaisia toimintatapoja teknologia ja käyttäjien vuorovaikutus yhdessä mahdollistavat. Vastuulliseen käyttöön vaikuttavat keskeisesti epävakaa levittyvyys, algoritminen kohdentaminen ja sisällön lopullinen muokkaamattomuus. (Ciriello ym., 2025.)

Epävakaa levittyvyys voi johtaa tilanteisiin, joissa sisältö leviää hallitsemattomasti käyttäjältä eteenpäin. Algoritminen kohdentaminen puolestaan mahdollistaa sisällön tarkan suuntaamisen käyttäjän mieltymysten ja aiemman toiminnan perusteella, mikä on hyödynnettävissä myös haitalliseen tarkoitukseen. Lopullisella muokkaamattomuudella tarkoitetaan sitä, että sisällön muokkaaminen tai poistaminen on alustalla usein vaikeaa. Lisäksi poistoyritykset voivat synnyttää niin kutsutun Streisand-ilmiön, jossa piilottaminen lisää sisällön näkyvyyttä ja herättää entistä enemmän huomiota. (Ciriello ym., 2025.)

Taulukko 3 kokoaa yhteen sosiaalisen median keskeisimmät rakenteelliset piirteet ja niiden vaikutusmekanismit väärän tiedon uskomiseen ja jakamiseen.

Taulukko 3. Sosiaalisen median rakenteen vaikutus väärän tiedon uskomiseen ja leviämiseen

Tekijä	Mekanismi	Vaikutus väärän tiedon uskomiseen ja leviämiseen	Lähde
Algoritminen sisällönjako	Käyttäjien sitouttaminen ja kaupallisen hyödyn maksimointi	Virheellinen ja äärimmäinen sisältö saa suhteellisesti enemmän näkyvyyttä	Riemer & Peter, 2021
Kaikukammiot	Samanmielisten käyttäjien keskuudessa toistuva sisältö	Lisäävät yli-itsevarmuutta ja polarisaatiota sekä vähentävät oppimishalua	Acemoglu ym., 2024; Liu & Fowler, 2025
Somekuplat	Algoritmien kohdentama sisältö	Rajaa käyttäjän kohtaamaa tietoa ja vahvistaa ennakkokäsityksiä	Impicciché & Viviani, 2025
Tiedon toistuvuus	Nopeaan ja säännölliseen selaamiseen perustuva käyttölogiikka	Lisää koettua todenmukaisuutta	Liu & Fowler, 2025
Hedonistinen käyttötapa	Sisältöä kulutetaan viihteellisesti ja passiivisesti	Vähentää kriittistä arviointia ja faktantarkistusta	Rhodes, 2022; Liu & Fowler, 2025
Alustojen rakenteelliset piirteet	Rajoitettu sanamäärä ja jakamistoiminnon helppous	Johtavat tiedon yksinkertaistamiseen, liioitteluun ja levittämiseen ilman faktantarkistusta	Liu & Fowler, 2025
Alustojen affordanssit	Sisällön leviämisen ennakoimattomuus ja poistamisen hankaluus	Kiihdyttävät väärän tiedon leviämistä ja lisäävät näkyvyyttä	Ciriello ym., 2025

4 Leviämisen ehkäiseminen

Tässä luvussa tarkastellaan keinoja, joilla väärän tiedon leviämistä voidaan ehkäistä sosiaalisessa mediassa. Ensimmäisessä alaluvussa käsitellään sosiaalisen median alustojen käytössä olevia keinoja leviämisen rajoittamiseksi, kuten sisällön moderointia ja teknologisia ratkaisuja. Toisessa alaluvussa keskitytään käyttäjätason ehkäisykeinoihin ja siihen, miten käyttäjien tietoisuutta ja kriittistä ajattelua lisäämällä voidaan vähentää väärän tiedon jakamista.

EU on todennut, että X:n (entisen Twitterin) toimenpiteet väärän tiedon ja vihapuheen leviämisen estämiseksi eivät ole riittäviä. Digitaalisia palveluja koskeva säädös velvoittaa sosiaalisen median alustoja omistavat yritykset poistamaan laittoman sisällön taloudellisten seuraamusten uhalla. Väärä tieto ei kuitenkaan ole useinkaan yksiselitteisesti laitonta, ja sääntely vaihtelee maittain, mikä tekee alustojen työstä haastavaa. Yritysten onkin tasapainoteltava väärän tiedon torjunnan ja sananvapauden turvaamisen välillä. (Olenberger ym., 2026.) Alustojen julkilausuttujen periaatteiden ja niiden käytännössä toteuttaman sisällön moderoinnin välillä on myös merkittäviä eroja (Flynn ym., 2025). Leviämistä ehkäisevien toimien tavoitteena on auttaa käyttäjiä arvioimaan sisältöä kriittisesti ja tarjota lisätietoa vääristetyistä väitteistä (Olenberger ym., 2026). Sosiaalisen median alustat puuttuvat erityisesti harhaanjohtavaan sisältöön, identiteettiväärennöksiin, näennäistieteeseen, salaliittoteorioihin ja manipulointiin. Moderointikeinot eivät juuri eroa eri alustojen välillä: sisällön poistaminen on tyypillinen toimenpide, ja alustalta poistamista käytetään, kun sisältö on laitonta tai rikkoo alustan käyttöehtoja. (Gsenger, 2025.)

Bakir ja McStay (2018) tuovat esiin, että tunteita hyödynnetään katselukertojen kasvattamiseksi, mikä lisää mainostuottoja. Mainostajat eivät kuitenkaan hyödy siitä, että heidän mainoksensa yhdistetään epäluotettavaan tietoon. Tämän vuoksi mainostajien kanssa tehtävä yhteistyö voi vähentää väärän tiedon leviämistä. (Bakir & McStay, 2018.)

4.1 Alustojen interventiot

Olenberger ym. (2026) jakavat ehkäisevät toimet kolmeen tasoon: yhteisötasoon, tilitasoon ja sisältötasoon. Yhteisötasolla käyttäjiä voidaan kannustaa raportoimaan väärää tietoa ja pysäyttämään sen leviäminen. Tilitasolla alustat voivat rajoittaa yksittäisten tilien toimintaa. Sisältötasolla voidaan korostaa oikaisevaa tietoa tai tarjota selventäviä lähteitä.

Varoitukset voivat vähentää sekä väärän tiedon uskomista että sen jakamista (Olenberger ym., 2026). Kevyet muistutukset tiedon paikkansapitävyydestä parantavat jaetun tiedon laatua

(Pennycook ym., 2021; Pennycook & Rand, 2022). Tiedon merkitseminen vääräksi voi myös katkaista kierteen, jossa toistuva virheellinen tieto alkaa tuntua uskottavalta (Vellani ym., 2023).

Porter ym. (2024) kuitenkin esittävät, että käyttäjälle kohdennettu suorempi viesti väärän tiedon haitallisuudesta ja käyttäjän oman toiminnan merkityksestä vähentää uskoa virheellisiin väitteisiin ja halua jakaa niitä – mutta voi samalla heikentää halukkuutta jakaa myös totuudenmukaista tietoa. Olenberger ym. (2026) huomauttavat lisäksi, että varoitusten vaikutus on usein lyhytaikainen: käyttäjän voi olla vaikeaa muuttaa ajattelumallejaan, minkä takia tiedon saamisjärjestys on oleellista ajattelumallien muodostumisen kannalta. Tieto varoituksesta tallentuu ajattelumalleihin, mutta unohtuu ajan myötä, koska varoitus on toissijaista tietoa ja alttiimpi kognitiivisille vinoumille. Väärän tiedon jälkeen esitetty varoitus tai lisätieto voi aktivoida perusteellisemman prosessoinnin ja näin vähentää uskoa väärään tietoon. (Olenberger ym., 2026.)

French ym. (2025) korostavat varoitusten vaihtelevaa tehokkuutta ja tuovat esille, että dynaamiset varoitukset, kuten vilkkuvat lipukkeet, käyttäjän ohjaaminen aineiston arviointiin ja osittain peitetty sisältö, voivat vähentää kognitiivisten vinoumien vaikutusta.

Varoitusten lisäksi alustoilla käytetään myös sisällön sensurointia ehkäisykeinona. Sisällön sensurointi, kuten epäluotettavan tiedon poistaminen, voi toisinaan tuottaa päinvastaisen vaikutuksen. Käyttäjät saattavat päätellä, että jäljelle jäänyt sisältö on automaattisesti luotettavaa. Myös asiayhteyden tai lisälähteiden esittäminen voi johtaa siihen, että käyttäjät olettavat muiden tehneen faktantarkistuksen puolestaan. (Acemoglu ym., 2024.) Vaikuttajan poistaminen alustalta vähentää hänen saamaansa huomiota ja häntä koskevaa keskustelua. Samalla se lieventää myös hänen seuraajayhteisönsä toksisuutta. (Gsenger, 2025.)

Sensuroinnin ohella väärän tiedon leviämistä pyritään hillitsemään myös muilla keinoilla. Olenberger ym. (2026) esittävät, että keskeisiä väärän tiedon leviämistä ehkäiseviä keinoja ovat faktantarkistuspalvelut, käyttäjien kannustaminen harkittuun käyttäytymiseen, tekoälyn hyödyntäminen sekä luotettavien lähteiden tarjoaminen käyttäjille. Ehkäisevät toimet voivat kuitenkin kääntyä itseään vastaan: mieltymysten vastainen tieto voi johtaa puolustusreaktioon, jossa käyttäjä vahvistaa uskomuksiaan vastaväitteiden avulla. Rhodes (2022) huomauttaa, että faktantarkistus voi vähentää väärän tiedon jakamista, mutta vaikutukset ovat usein lyhytaikaisia algoritmien synnyttämien kaikukammioiden vuoksi.

Perinteisten menetelmien rinnalle ovat nousseet tekoälypohjaiset tunnistusmenetelmät. Kumari ym. (2021) esittävät, että useita rinnakkaisia tehtäviä suorittava koneoppimismalli pystyy tunnistamaan

väärää tietoa tehokkaasti kouluttamalla mallin tunnistamaan tiedon uutuusarvon, tunnetilan ja sävyn. Koneoppimisella tarkoitetaan automatisoitua päätöksentekoa datasta opittujen rakenteiden avulla ilman, että mallin toimintalogiikka ohjelmoidaan käsin (Kufel ym., 2023). Ozcelik ym. (2025) puolestaan osoittavat, että käyttäjäyhteisöt jakautuvat usein totuudenmukaista ja väärää tietoa jakaviin ryhmiin. He kuvaavat SiMiD-järjestelmää, joka tunnistaa nämä ryhmät ja hyödyntää kielimalleja väärän tiedon havaitsemisessa.

Generatiivinen tekoäly tuottaa suurten kielimallien avulla ihmismäisiä vastauksia saamiinsa syötteisiin. Laajat kielimallit ovat syväoppimiseen perustuvia järjestelmiä, joita koulutetaan erittäin laajoilla tekstiaineistoilla luonnollisen kielen tuottamiseen ja ennakointiin. Laajat kielimallit kykenevät tunnistamaan virheellisiä väitteitä, mutta niiden tarkkuus vaihtelee muun muassa käytetyn kielen ja kehotteiden muotoilun mukaan. Lisäksi generatiivisen tekoälyn kyky tuottaa uskottavaa väärää tietoa heikentää luottamusta siihen, kuinka luotettavasti sama teknologia pystyy tunnistamaan väärää tietoa. Tutkimus korostaa, että tekoälyavusteinen väärän tiedon tunnistus edellyttää edelleen kontekstisidonnaista hienosäätöä sekä selkeää sääntelyä. (Park & Nan, 2026.)

4.2 Käyttäjätason ehkäisykeinot

French ym. (2025) katsovat, että ehkäiseviä toimia voidaan kohdistaa kognitiivisten vinoumien vaikutusten vähentämiseksi. Vastakkaisten näkemysten ja kilpailevien selitysten esiin tuominen voi hillitä ankkurointi- ja kehystysvaikutusta. Vilkkuvat varoituslipukkeet sekä suostumukseen perustuva sisällön näkyvyyden rajoittaminen voivat vähentää luottamus- ja vahvistusharhaa. Tekstin visualisointi ohjaa huomiota sisältöön esitystavan sijaan ja voi siten lieventää kehystysvaikutusta. Lisäksi viraalimittareiden (tykkäysten, kommenttien ja jakojen) piilottaminen voi heikentää laumamielialaa. Rhodes (2022) puolestaan korostaa erilaisten näkemysten esittämistä ja somekuplien sekä kaikukammioiden purkamista keinona vähentää väärään tietoon uskomista.

Ciriello ym. (2025) painottavat käyttäjäarkkityyppien ja luonteenpiirteiden tuntemisen merkitystä. Dogmaatikko, joka suosii uskomuksiaan vahvistavaa sisältöä, voi hyötyä monimuotoisille näkemyksille altistumisesta. Välinpitämätön nihilisti voi hyötyä matalan sitoutumisen aktivismista, kuten sivistyksellisen sisällön jakamisesta. Narsisti voi hyötyä molemmista. Altruistin kampanjoiden näkyvyyttä voidaan lisätä algoritmisella kohdentamisella. Ciriello ym. (2025) esittävät myös, että mindfulness ja monipuoliselle sisällölle altistuminen voivat lisätä harkitsevuutta. Lyhyt pysähdys ennen sisällön jakamista voi auttaa käyttäjää tiedostamaan oman toimintansa vaikutukset.

Beisecker ym. (2024) osoittavat, että käyttäjän tietoisuuden lisääminen estää väärän tiedon leviämistä. Harhaanjohtava argumentointi voi perustua totuudenmukaiseen tietoon, jota muokataan tai liioitellaan. Kun käyttäjä oppii tunnistamaan tällaiset argumentointikeinot, hän tunnistaa väärän tiedon paremmin ja ylittää omia ennakkoluulojaan. Olenberger ym. (2026) huomauttavat, että kehotukset syvällisempään kognitiiviseen sitoutumiseen tukevat luotettavuuden arviointia.

Taulukko 4 kokoaa yhteen keskeiset väärän tiedon leviämisen ehkäisykeinot.

Taulukko 4. Väärän tiedon leviämisen ehkäisykeinot

Ehkäisykeino	Mekanismi	Vaikutus väärän tiedon leviämiseen	Rajoitteet	Lähde
Sisällön sensurointi (poistaminen, tilien sulkeminen)	Haitallisen sisällön näkyvyyden rajoittaminen	Vähentää väärän tiedon näkyvyyttä ja siihen liittyvää keskustelua	Voi herättää sananvapauskritiikkiä ja luoda näennäistä luottamusta jäljelle jäävään sisältöön	Acemoglu ym., 2024; Gsenger, 2025
Varoitukset	Ohjaa kriittiseen arviointiin ja perusteellisempaan prosessointiin	Vähentää väärän tiedon uskomista ja jakamista	Vaikutus voi olla lyhytaikainen ja se voi vähentää myös totuudenmukaisen sisällön jakamista	Pennycook & Rand, 2022; Vellani ym., 2023; Porter ym., 2024; Olenberger ym., 2026
Faktantarkistus ja lisälähteet	Tarjoaa oikaisevaa tietoa	Vähentää väärän tiedon jakamista	Kaikukammioiden vuoksi vaikutus on usein lyhytaikainen	Rhodes 2022, Olenberger ym., 2026
Tekoälypohjainen tunnistus	Automaattinen virheellisen sisällön tunnistus	Mahdollistaa nopean ja laajamittaisen seulonnan	Edellyttää kontekstisidonnaista hienosäätöä ja selkeää sääntelyä	Kumari ym., 2021; Ozcelik ym., 2025; Park & Nan, 2026
Kognitiivisiin vinoumiin kohdistuvat interventiot	Vastakkaisten näkemysten tarjoaminen, tiedon visualisointi ja viraalimittareiden piilottaminen	Parantaa kykyä arvioida tiedon luotettavuutta	Vaatii aktiivista osallistumista	French ym., 2025
Käyttäjätyyppien huomioiminen	Räätälöidyt interventiot käyttäjille	Lisää ehkäisyn vaikuttavuutta	Edellyttää käyttäjäprofilointia	Ciriello ym., 2025
Käyttäjän tietoisuuden lisääminen	Harhaanjohtavan argumentoinnin tunnistaminen ja kannustaminen syvällisempään kognitiiviseen sitoutumiseen	Parantaa väärän tiedon tunnistamista	On riippuvaista käyttäjästä	Beisecker ym. 2024; Olenberger ym., 2026

5 Yhteenveto

Tässä tutkimuksessa tarkasteltiin väärän tiedon leviämistä sosiaalisessa mediassa sekä keinoja sen ehkäisemiseksi. Työn tavoitteena oli vastata kolmeen tutkimuskysymykseen:

1. Mitkä yksilölliset tekijät vaikuttavat siihen, että ihmiset uskovat ja jakavat väärää tietoa?
2. Miten sosiaalisen median suosittelumekanismit vaikuttavat väärän tiedon leviämiseen?
3. Mitkä ovat tehokkaimpia keinoja ehkäistä väärän tiedon leviämistä sosiaalisessa mediassa?

Tutkielman tulokset osoittavat, että väärän tiedon leviämistä selittävät sekä yksilötason kognitiiviset tekijät että sosiaalisen median rakenteelliset ominaisuudet. Niiden yhteisvaikutus tekee ehkäisystä monimutkaista. Väärä tieto voi uhata demokraattisia arvoja, mutta samalla sen rajoittamiseen liittyvät toimet voivat herättää huolta sananvapaudesta.

Yhä useampi käyttää sosiaalista mediaa ensisijaisena tiedonlähteenään (Liu & Fowler, 2025). Alustat sisältävät valtavan määrän tietoa, mikä voi johtaa informaatiotulvaan ja pintapuoliseen sisällön arviointiin (Olenberger ym., 2026). Sosiaalinen media on myös suunniteltu nopeaa ja hedonistista selailua varten, mikä vähentää käyttäjien faktantarkistusta ja kriittistä tarkastelua. Myös rajoitettu sanamäärä ja jakamistoiminnon helppous kannustavat tiedon nopeaan levittämiseen. (Liu & Fowler, 2025.) Väärän tiedon herättämät tunteet puolestaan kasvattavat tiedon eteenpäin jakamisen todennäköisyyttä (Vosoughi ym., 2018).

Kognitiiviset vinoumat, kuten vahvistusharha, altistavat käyttäjät virheelliselle tiedolle (Olenberger ym., 2026). Lisäksi käyttäjät turvautuvat heuristiikkoihin arvioidessaan tiedon luotettavuutta (Metzger ym., 2010). Luottamus voi perustua esimerkiksi tykkäysten ja kommenttien määrään (Sundar ym., 2025). Uutuusarvo ja yllätyksellisyys lisäävät tiedon uskottavuutta ja leviämistä (Kumari ym., 2021). Viraalius puolestaan kasvattaa näkyvyyttä sosiaalisessa mediassa (Vosoughi ym., 2018). Sosiaalisen median rakenteelliset ja teknologiset piirteet vahvistavat tätä mekanismia: sisältö voi levitä hallitsemattomasti eikä sen poistaminen ole aina mahdollista (Ciriello ym., 2025). Algoritmit pyrkivät maksimoimaan käyttäjien sitoutumisen (Riemer & Peter, 2021; Acemoglu ym., 2024). Tämä edistää kaikukammioiden ja somekuplien syntymistä (Rhodes, 2022). Toistuva tieto koetaan luotettavammaksi, mikä kiihdyttää väärän tiedon leviämistä (Vellani ym., 2023).

Väärän tiedon leviämistä voidaan ehkäistä sekä alustatason että käyttäjätason toimilla. Alustat voivat kannustaa käyttäjäyhteisöä raportoimaan väärää tietoa, rajoittaa epäasiallisesti toimivien tilien toimintaa ja tarjota selventäviä lähteitä (Olenberger ym., 2026). Sisällön sensurointi voi toisaalta

johtaa siihen, että käyttäjät pitävät jäljelle jäävää sisältöä automaattisesti luotettavana (Acemoglu ym., 2024).

Varoitukset ovat yksi keskeinen interventiomuoto (Olenberger ym., 2026). Varoitusten tehokkuudesta ja toimeenpanosta on vaihtelevia näkemyksiä. Pennycook ym. (2021) osoittavat, että kevyet muistutukset tiedon paikkansapitävyydestä voivat parantaa jaettavan tiedon laatua, kun taas Porter ym. (2024) esittävät, että suoremmat viestit voivat vähentää virheellisen tiedon jakamista tehokkaammin. Samalla varoitukset voivat heikentää luottamusta myös totuudenmukaista sisältöä kohtaan (Porter ym., 2024; Olenberger ym., 2026). French ym. (2025) toteavat, että varoitusten tehokkuus vaihtelee ja että dynaamiset varoitukset voivat vähentää kognitiivisten vinoumien vaikutusta. Olenberger ym. (2026) puolestaan korostavat varoituksen lyhytaikaista vaikutusta, sillä varoitus on toissijaista tietoa ja alttiimpaa unohtamiselle. Väärän tiedon jälkeen esitetty varoitus voi silti vähentää uskoa virheelliseen tietoon.

Tekoälyä voidaan hyödyntää väärän tiedon (Kumari ym., 2021) ja sitä jakavien ryhmien tunnistamisessa (Ozcelik ym., 2025). Olenberger ym. (2026) nostavat faktantarkistuspalvelut keskeiseksi ehkäisykeinoksi, mutta Rhodesin (2022) mukaan niiden vaikutukset voivat jäädä lyhytaikaisiksi kaikukammioiden vuoksi. Vaihtoehtoisten näkemysten esille tuominen (French ym., 2025), somekuplien ja kaikukammioiden purkaminen (Rhodes, 2022), erilaisten toimien kohdistaminen käyttäjäarkkityyppeihin sekä käyttäjien harkintaan kannustaminen (Ciriello ym., 2025) tukevat kriittistä arviointia. Myös argumentointikeinojen tunnistaminen lisää käyttäjän kykyä havaita väärää tietoa (Beisecker ym., 2024).

Ehkäisevät toimet voivat kuitenkin johtaa käyttäjän puolustusreaktioon, jossa käyttäjä vahvistaa omia uskomuksiaan vastaväitteiden avulla (Olenberger ym., 2026).

6 Johtopäätökset ja jatkotutkimus

Tutkielma tarjoaa kokonaiskuvan siitä, millaiset yksilöiden ominaisuudet ja sosiaalisen median rakenteelliset piirteet vaikuttavat väärän tiedon leviämiseen. Se lisää käyttäjien ymmärrystä heidän omista kognitiivisista ja tunneperäisistä taipumuksistaan sekä keinoista, joilla heihin pyritään vaikuttamaan sosiaalisessa mediassa. Lisäksi tutkielma antaa sosiaalisen median alustayrityksille yleiskatsauksen käytössä olevista ehkäisykeinoista. Tutkielman perusteella väärän tiedon leviämistä voidaan hillitä useilla menetelmillä, mutta niiden vaikuttavuus vaihtelee. Ehkäisyä vaikeuttavat erityisesti sananvapauden turvaamiseen liittyvät haasteet, sääntelyn epäyhtenäisyys sekä alustojen eri toimintaperiaatteet. Tilannetta monimutkaistaa myös käyttäjien välinen heterogeenisyys ja se, miten tämä voidaan huomioida ehkäisytoimia suunniteltaessa.

Väärän tiedon vaikutus kasvaa, kun yhä useampi käyttää sosiaalista mediaa ensisijaisena tiedonhankinnan kanavana. Koska ilmiöllä on laaja yhteiskunnallinen ulottuvuus, olisi perusteltua harkita yritysten ja julkisen sektorin tiiviimpää yhteistyötä ongelman hillitsemiseksi.

Tutkimuskirjallisuudessa ei ole yksimielisyyttä siitä, mitkä toimenpiteet ovat tehokkaimpia, ja lisätutkimusta tarvittaisiin erityisesti eri ehkäisykeinojen vaikuttavuudesta. Vaikka tekoälyä on hyödynnetty jonkin verran väärän tiedon tunnistamiseen, teknologian nopea kehitys edellyttää lisätutkimusta siitä, mitkä tekoälyratkaisut ovat kaikkein luotettavimpia väärän tiedon havaitsemisessa ja miten niiden käyttö näkyy käytännössä sosiaalisen median sisältövirrassa. Kirjallisuudesta ei myöskään löytynyt kattavaa kuvaa siitä, mitä keinoja eri alustat tosiasiallisesti hyödyntävät. Koska alustojen käyttötarkoitukset ja käyttäjäkunnat voivat poiketa toisistaan, tulevissa tutkimuksissa olisi hyödyllistä kartoittaa, millaisia ehkäisykeinoja tietyt alustat soveltavat ja miten käyttäjien demografiset tekijät sekä persoonallisuuden piirteet vaikuttavat näiden keinojen tehokkuuteen.

Lähteet

Acemoglu, D., Ozdaglar, A., & Siderius, J. (2024). A Model of Online Misinformation. *Review of Economic Studies*, *91*(6), 3117–3150. https://doi.org/10.1093/restud/rdad111

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital journalism*, *6*(2), 154–175.

Beisecker, S., Schlereth, C., & Hein, S. (2024). Shades of fake news: How fallacies influence consumers' perception. *European Journal of Information Systems*, *33*(1), 41–60. https://doi.org/10.1080/0960085X.2022.2110000

Chen, X. (2024). Misinformation and personality: Big-5 trait related to misinformation believing and sharing. *Communications*, *47*(1), 25–29.

Ciriello, R., Gal, U., Hannon, O., & Thatcher, J. (2025). Responsible social media use: How user characteristics shape the actualisation of ambiguous affordances. *European Journal of Information Systems*, *34*(5), 799–821. https://doi.org/10.1080/0960085X.2024.2444249

Flynn, A., Vakhitova, Z., Wheildon, L., Harris, B., & Robards, B. (2025). Content Moderation and Community Standards: The Disconnect Between Policy and User Experiences Reporting Harmful and Offensive Content on Social Media. *Policy & Internet*, *17*(3), e70006. https://doi.org/10.1002/poi3.70006

French, A. M., Storey, V. C., & Wallace, L. (2025). The impact of cognitive biases on the believability of fake news. *European Journal of Information Systems*, *34*(1), 72–93. https://doi.org/10.1080/0960085X.2023.2272608

Greene, K. T., Pereira, M., Pisharody, N., Dodhia, R., Lavista-Ferres, J., & Shapiro, J. N. (2025). Using website referrals to identify unreliable content rabbit holes. *Behaviour and Information Technology*, *44*(7), 1340–1349. https://doi.org/10.1080/0144929X.2024.2352093

Gsenger, R. (2025). Platform governance under the Digital Services Act: A perspective on disinformation. *Information Communication and Society*. https://doi.org/10.1080/1369118X.2025.2590561

Hameleers, M. (2025). The visual nature of information warfare: The construction of partisan claims on truth and evidence in the context of wars in Ukraine and Israel/Palestine. *Journal of Communication*, *75*(2), 90–100. https://doi.org/10.1093/joc/jqae045

Hassani, H., Komendantova, N., Rovenskaya, E., & Yeganegi, M. R. (2024). Unveiling the waves of mis- and disinformation from social media. *International Journal of Modeling, Simulation, and Scientific Computing*, *15*(3). https://doi.org/10.1142/S1793962324500338

Hubeny, T., Nahon, L., Ng, N., & Gawronski, B. (2025). Who Falls for Misinformation and Why? Personality and social psychology bulletin. (WOS:001457480900001). https://doi.org/10.1177/01461672251328800

Impicciché, P., & Viviani, M. (2025). Comparing Echo Chamber Detection Metrics: A Cross-modeling and Cross-platform Analysis of Twitter and Reddit. *ACM Transactions on the Web*, *19*(4). https://doi.org/10.1145/3707701

Keswani, V. (2023). *Social Media Platform Structures and Their Implications*. *3456*, 126–131. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85171154129&partnerID=40&md5=792c5b2105ed2e6c2cb70037bc5a8a48

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023). What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. *Diagnostics (Basel)*, *13*(15), 2582. https://doi.org/10.3390/diagnostics13152582

Kumari, R., Ashok, N., Ghosal, T., & Ekbal, A. (2021). *A Multitask Learning Approach for Fake News Detection: Novelty, Emotion, and Sentiment Lend a Helping Hand*. *2021-July*. https://doi.org/10.1109/IJCNN52387.2021.9534218

Liu, M., & Fowler, S. (2025). Engineering the Discourse: The Role of Engineers in the Health Infodemic. *Asian Bioethics Review*, *17*(3), 463–476. https://doi.org/10.1007/s41649-024-00352-y

Maasberg, M., Ayaburi, E., Liu, C., & Au, Y. (2018). *Exploring the propagation of fake cyber news: An experimental approach*.

Metzger, M., Flanagin, A., & Medders, R. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of communication*, *60*(3), 413–439. (WOS:000281150600001). https://doi.org/10.1111/j.1460-2466.2010.01488.x

Mulcahy, R., Barnes, R., de Villiers Scheepers, R., Kay, S., & List, E. (2025). Going Viral: Sharing of Misinformation by Social Media Influencers. *Australasian Marketing Journal*, *33*(3), 296–309. https://doi.org/10.1177/14413582241273987

Olenberger, C., Schoch, M., & Utz, L. (2026). The information processing of fake news: How intervention order influences perception over time. *Information & Management*, *63*(1), 104256. https://doi.org/10.1016/j.im.2025.104256

Ozcelik, O., Toraman, C., & Can, F. (2025). Detecting Misinformation on Social Media using Community Insights and Contrastive Learning. *ACM Transactions on Intelligent Systems and Technology*, *16*(2). https://doi.org/10.1145/3709009

Park, S., & Nan, X. (2026). Generative AI and misinformation: A scoping review of the role of generative AI in the generation, detection, mitigation, and impact of misinformation. *AI and Society*, *41*(2), 1501–1515. https://doi.org/10.1007/s00146-025-02620-3

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A., Eckles, D., & Rand, D. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590-+. (WOS:000629906100002). https://doi.org/10.1038/s41586-021-03344-2

Pennycook, G., & Rand, D. (2022). Nudging Social Media toward Accuracy. Annals of the American Academy of Political and Social Science, *700*(1), 152–164. (WOS:000791952600011). https://doi.org/10.1177/00027162221092342

Porter, E., Wood, T., Broniatowski, D., & Hosseini, P. (2024). Shoves, nudges and combating misinformation: Evidence on a new approach. Behavioural public policy. (WOS:001346241000001). https://doi.org/10.1017/bpp.2024.42

Qiu, H., & Goh, D. H.-L. (2025). Disinformation Recognition in Social Media: Effects of Multimodality, Comments and Virality Metrics. *Proceedings of the Association for Information Science and Technology*, *62*(1), 1647–1649. https://doi.org/10.1002/pra2.1495

Rhodes, S. C. (2022). Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation. *Political Communication*, *39*(1), 1–22. https://doi.org/10.1080/10584609.2021.1910887

Riemer, K., & Peter, S. (2021). Algorithmic audiencing: Why we need to rethink free speech on social media. *Journal of Information Technology*, *36*(4), 409–418. https://doi.org/10.1177/02683962211013358

Seol, S., Mejia, J., & Dennis, A. R. (2024). Lying for Viewers: Commingled Partisan Falsehoods Increase Viewing and Sharing of News Media1. *Management Information Systems Quarterly*, *48*(2), 551–582. https://doi.org/10.25300/MISQ/2023/17928

Shirish, A., Srivastava, S. C., & Chandra, S. (2021). Impact of mobile connectivity and freedom on fake news propensity during the COVID-19 pandemic: A cross-country empirical examination. *European Journal of Information Systems*, *30*(3), 322–341. https://doi.org/10.1080/0960085X.2021.1886614

Sundar, S. S., Cho, E. C., Liao, M., Yin, J., Wang, J., & Chi, G. (2025). Sharing without clicking on news in social media. *Nature Human Behaviour*, *9*(1), 156–168. https://doi.org/10.1038/s41562-024-02067-4

Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. Plos one, *13*(9). (WOS:000445626400045). https://doi.org/10.1371/journal.pone.0203958

van der Puil, R., Spahn, A., & Royakkers, L. (2023). Which Democratic Way to Go? Using Democracy Theories in Social Media Design. *International Journal of Technoethics*, *14*(1). https://doi.org/10.4018/IJT.331800

Vannucci, A., Yu, W., Martin, N., Patel, S., & Tottenham, N. (2025). Affective Schemas: Acquisition, Updating, and Inference. *Emotion*. https://doi.org/10.1037/emo0001561

Vellani, V., Zheng, S., Ercelik, D., & Sharot, T. (2023). The illusory truth effect leads to the spread of misinformation. Cognition, *236*. (WOS:000952589600001). https://doi.org/10.1016/j.cognition.2023.105421

Villa, G., Pasi, G., & Viviani, M. (2021). Echo chamber detection and analysis A topology- and content-based approach in the COVID-19 scenario. Social network analysis and mining, *11*(1). (WOS:000687153200001). https://doi.org/10.1007/s13278-021-00779-3

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-+. (WOS:000426835900044). https://doi.org/10.1126/science.aap9559

Wang, Y., Basellini, U., & Camarda, C.-G. (2026). The infodemic–pandemic nexus: Cross-national analysis of online misinformation's impact on population health. *Social Science and Medicine*, *390*. https://doi.org/10.1016/j.socscimed.2025.118881

Liitteet

Liite 1 Selvitys tekoälyn käytöstä

Tutkielman laatimisen tukena olen käyttänyt OpenAI ChatGPT:tä ja Microsoft Copilotia (versiot 2/2026). Hyödynsin kumpaakin työkalua tutkimusaiheen ja siihen liittyvien hakutermien ideoimisessa sekä alustavien tutkimuskysymyksien muotoilussa. Käyttämäni kehotteet olivat esimerkiksi:

”I’m interested in researching mis/disinformation. Could you come up with a current topic related to data, processes or people?”

“What research terms could I use to find studies made on this topic?”

“What research questions are related to this topic?”

Aiheen, hakutermien ja tutkimuskysymysten lopullinen valinta perustui kuitenkin aiempaan tutkimuskirjallisuuteen, joka vahvisti rajauksen ja auttoi hienosäätämään niitä tarpeen mukaan.

Käytin Microsoft Copilot -työkalua myös tekstin kielelliseen viimeistelyyn varmistaakseni, että esitystapa oli selkeä ja akateemiseen tyyliin sopiva. Käyttämäni kehote oli:

”Voitko tarkistaa, että teksti on kielellisesti sujuva ja muotoilla sitä yksinkertaisemmaksi ja akateemiseen tyyliin sopivaksi siltä osin kuin on tarpeellista? Osoita kohdat, joihin ehdotat muutoksia äläkä tee muutoksia automaattisesti.”

Kävin kaikki tekoälyn antamat muokkaus- ja parannusehdotukset läpi itse varmistaakseni tekstin sisällöllisen ja kielellisen oikeellisuuden.