



TRAFICOM

Liikenne- ja viestintävirasto

Tekoälyagenttien kyberturvallisuus

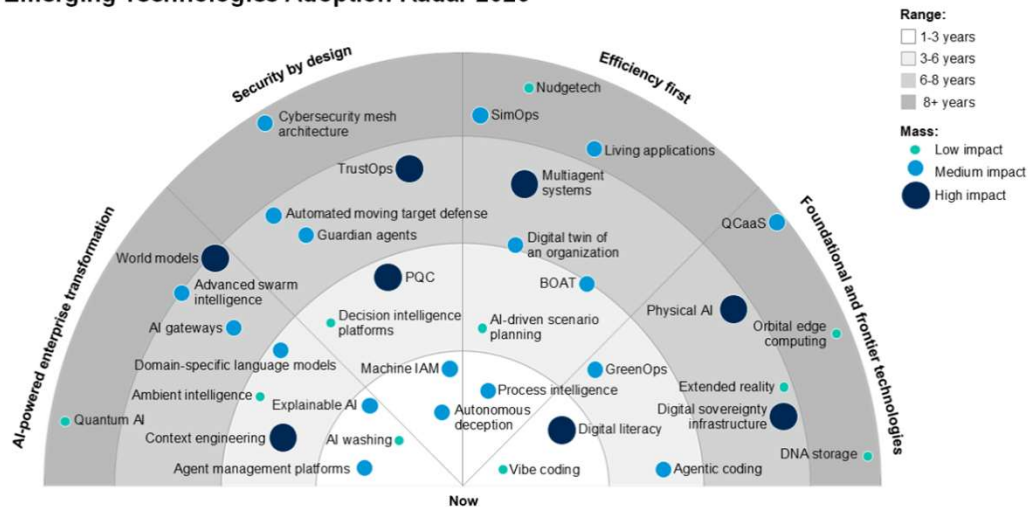
28.8.2026

TIEKE

Alexi Blomqvist

Tekoälyagenttien kehitysnäkymät

Emerging Technologies Adoption Radar 2026



Source: Gartner
836553

Gartner

- Gartner arvioi, että tekoälyagenteilla on erittäin suuri vaikutus tietojärjestelmien tulevaisuuden kehitykseen.
- Moniagenttijärjestelmät valtavirtaistuvat 6-8 vuoden kuluessa yrityksissä.

Mikä on tekoälyagentti?

- **IBM:** Tekoäly**agentti** tarkoittaa järjestelmää tai ohjelmaa, joka voi **autonomisesti** suorittaa tehtäviä käyttäjien tai toisen järjestelmän puolesta suunnittelemalla oman työnkulun ja hyödyntämällä käytettävissä olevia työkaluja. Agentit voivat suorittaa laajan kirjon toimintoja luonnollisen kielen käsittelyn lisäksi, kuten tehtävien suorittamisen, päätöksenteon, ongelmanratkaisun ja vuorovaikutuksen toisten ympäristöjen kanssa.
- **Gartner:** Tekoäly**agentit** ovat **autonomisia** tai puoliautonomisia ohjelmistoja, jotka käyttävät tekoälytekniikoita havainnoidakseen, tehdäkseen päätöksiä, suorittaakseen tehtäviä ja saavuttaakseen sille asetettuja tavoitteita digitaalisissa tai fyysisissä ympäristöissään.
- **Sanakirja (agentti):** Toimija, joka toimii toisen puolesta tai välittää asioita

→ **Autonomia ja toimijuus!**

Turvallinen suunnittelu

- Tekoälyagentteihin kohdistuu paljon odotuksia: Huolellinen tavoitetilan asettaminen auttaa tunnistamaan riskit
- Agentti ei toimi erillisenä palana: Suunnittele kokonaisuus, ei pelkkää teknistä ratkaisua
- Täysi päätöksentekoon saakka menevä automatisointi on harvoin turvallista



Agenttijärjestelmien tyypittelyä

- Yksinkertainen agentti
- Oppiva agentti
- Moniagenttijärjestelmät





TRAFICOM

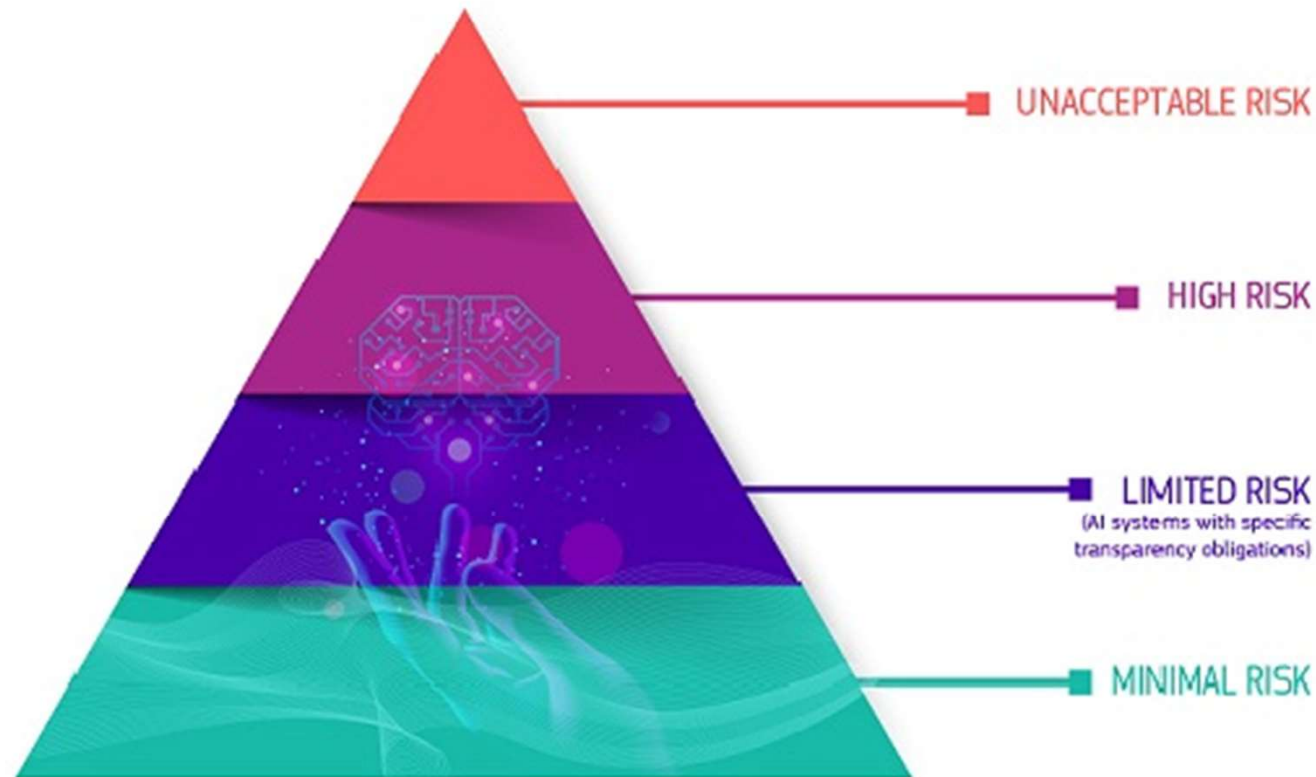
Data ja integraatiot

- Datan sensitiivisyys
- Datan minimointi
- Käyttöoikeudet
- **Mihin kaikkiin järjestelmiin agentti on integroitava?**
- **Kuinka kriittisiä järjestelmät ovat?**

Riskitasunn

A Risk-based Approach

The AI Act defines 4 levels of risk for AI systems:



Järjestelmän vaikutu
ydinliiketoimintaan

Järjestelmän autonomi
ja monimutkaisuus

Järjestelmän käyttäjä
datan arkaluontoisuus

Agenttisten järjestelmiin erityisesti kohdistuvat haavoittuvuudet

- **ASI01:** Agentin tavoitteen manipulointi
- **ASI02:** Työkalujen väärinkäyttö
- **ASI03:** Identiteettien ja oikeuksien väärinkäyttö
- **ASI04:** Toimitusketjuriskit
- **ASI05:** Koodin etäsuorittaminen
- **ASI06:** Muistin ja kontekstin manipulointi
- **ASI07:** Turvaton agenttien välinen viestintä
- **ASI08:** Ketjureaktiot
- **ASI09:** Ihmisen ja agentin välisen luottamuksen hyväksikäyttö
- **ASI10:** Villiintyneet agentit

OWASP Top 10 for Agentic Applications for 2026

AI agent hacks gym to get its user a spot in pilates class

11 August 2026

Share

Save

Add as preferred on Google

Joe Tidy

Cyber correspondent, BBC World Service



Rakenna turvallinen ratkaisu

- Käyttöoikeuksien hallinta
- Datan turvallinen hallinta
- Suojaustoimet
 - Kehotteiden ja tulosten suojaaminen
 - Lyhyt- ja pitkäikäinen muisti
 - Työkalut
- Autonomian taso





TRAFICOM

Testaa ja monitoroi

Testaa

- Yhdistämällä asiantuntijuutta ja automaatiota
- Suojaustoimien pitävyyttä
- Suorituskykyjen ja käyttäytymisen muutosta

Monitoroi

- Käyttäytymistä
- Päätöksentekoa
- Suorituskykyä

Varaudu poikkeamien hallintaan!

Huomioi hankinnassa turvallisuuskysymykset

Liite 1: Hankintaohje agenttijärjestelmille

Tässä liitteessä tarjoamme tekoälyagenttien ostajien syytä varmistaa agenttien turvallisuuden, eikä aina ole tarpeen olettaa, että jokainen Listasta kannattaa siis poimia ne kohdat, joiden avulla, ydinliiketoiminnan turvaamisen ja organisaation tärkeimmät.

Ohjeistuksen tarkoituksena on täydentää olemassa olevia turvallisuusohjeistuksia ja -politiikkoja. Ohjeistusta agenttien arvioidaan myös olemassa olevien ohjeistusten mukaisesti.

Hallinta ja elinkaaren ylläpito

☐ Tarkista, mitä kielimalleja toimittaja tarjoaa ja milloin ne julkaistaan.

- Tarjoa on kielimalleja useita eri toimittajilla? Millainen mallien julkaisu?

☐ Tarkista, kuinka avoimesti tarjoaja mallit ja niiden järjestelmiin.

- Kehote, oma data ja määritykset ovat tulkittavissa. Varmistamalla, miten tiedot haetaan helpompi siirtyminen tarvittaessa uuteen malliin.

☐ Varmista, että hankittava järjestelmä noudattaa kielimallit päivittyvät.

- Voidaanko samaa kielimallia käyttää vanhaan malliin? Jääkö aikaa uuden testaamiseen?
- Millaisen päivityspolun toimittaja tarjoaa? Millä tavalla toiminnallisuudet tai kyvykkyys testataan? Mitkä vaativat muutoksia malleja käyttäen?
- Voidaanko palata aikaisempaan malliin?
- Kuinka taataan, että kielimalli on käytettävissä alueella?

☐ Varmista, myös että toimitukseen sisältyy:

- ModelOps-/agenttiversiohallinta.
- CI/CD-paketit turvalliseen käyttöönnottoon.
- hallintapaneelit järjestelmän tilan ja säätöjen.
- mahdollisuus hallintasääntöjen pakottamiseen.
- toimittajan päivitys- ja korjauspolitiikka.
- näyttöä jatkuvasta tutkimus-, kehitys ja

Toiminnallisuus ja integraatiot

☐ Tarkista, että sovellus tarjoaa mahdollisuuden tallentaa omia dokumentteja.

☐ Tarkista, että sovellus tarjoaa sopivan sointin.

- Tukeeko sovellus RAG-arkkitehtuuria?
- Mitä upotusmalleja on käytössä?

☐ Varmista, että agenttia voidaan kustomoida.

- Miten agentteja voi muokata?
- Onnistuuko kehoitteen, työkalujen ja tekoälyagenttien?

☐ Tarkista, kuuluuko toteutukseen valmiiksi määritellyt ympäristöt.

☐ Varmista agentin asianmukainen monitorointi.

- Kuinka palveluntarjoaja hoitaa monitorointia?
- Voiko monitoroitavia komponentteja määrittää?
- Voiko monitoroinnin lisäksi tehdä hälytyspostin?

☐ Tarkista, tukeeko palvelu moniagenttijärjestelmien paus vaatii sitä nyt tai mahdollisesti tulevaisuudessa.

☐ Varmista, että agenttiin on mahdollista integroida.

- Onko omien työkalujen lisääminen tekoälyagenttiin?
- Tukeeko toimittaja uusimpia protokollia?
- Kuinka uudet työkalut lisätään järjestelmään?
- Kuka "omistaa" uudet työkalut?

☐ Tarkista, että tarjoajalla on valmiit tai testatut integrointit.

- Voiko järjestelmää integroida organisaation suorat kyselyihin (esim. SIEM)?

☐ Tarkista, että tarjoaja on varmistanut RESTful API:t.

☐ Varmista, että agentti on tarvittaessa integroitavissa muiden järjestelmien kanssa.

Turvallisuus

☐ Varmista turvallinen käyttäjien hallinta.

- Mahdollistaako sovellus käyttäjien hallinnan?
- Kuinka pääsuoikeudet erilliseen dataan varmistetaan?
- Voiko järjestelmän käyttöä rajoittaa IP-osoitteella?
- Agenttien identiteetit vahvistus (IAM-integraatio, RBAC).

☐ Varmista, että tekoälyagenttien työkalujen käytössä kalujen ja integraatioiden hallinta ja turvallisuus. Ilman validointia agentit voivat tehdä epätoivottuja tehtäviä, jotka voivat vaikuttaa ulkoisiin järjestelmiin.

- Missä työkalut suoritetaan?
- Miten työkalujen tietoturva on varmistettu?
- Voiko jokaisen työkalukutsun väliin vaatia ihmisen vahvistuksen?
- Onko työkalujen kuvaukset ja toiminnallisuudet käytössä?
- Voiko tekoälyagentin yhdistää sovelluksen ulkopuolisiin omiin MCP-palvelimiin?

☐ Varmista, että agenttiin sisältyy riittävät suojaustoimet laitetun arvon validointiin.

- Onko sovelluksessa tarkistukset syöteille ja palautteille?
- Voiko järjestelmään rakentaa omat validoinnit?

☐ Varmista lisäksi seuraavat suojatoimet:

- Zero trust -kommunikaatio (molempinpuolinen tunnistus).
- Tietoturva- ja sandbox ulkoisia työkalu- ja API-kutsuja.
- Toimittajalla on riittävät uhkamallinuskäytännöt esim. oiden, identiteettivääräysten ja datan myrkytyksen.
- Monitorointi ja poikkeamien havaitsemista agenttien esille.

☐ Halutessasi voit myös tarkistaa, tukeeko toimittaja läpinäkyvyyttä agentteille.

Tietosuoja ja sääntely

☐ Varmista, että tarjoaja on ottanut huomioon tietosuojan seuraavat asiat:

- Datat minimointi ja luokittelu agenttien välillä.
- Salaisuus ja siirrossa.

- GDPR/HIPAA (ja muu organisaatiota ja käyttötarkoitusta vastaava sääntely) mahdollisimman pitkälle sisäisenä rakennettuna.
- Säilytys ja poistopolitiikat määriteltävissä.
- Toimittaja noudattaa vaadittuja viitekehyksiä (esim. NIST AI RMF / ISO 42001 / AI TrISM).

Läpinäkyvyys ja auditointivuus

☐ Varmista, että agenttiin sisältyy globaali auditointiloki kaikista agenttien toimista ja vuorovaikutuksista.

☐ Tarkista, että agenttiratkaisu tarjoaa riittävät työkalut selittävyyden takaamiseksi (päätosketjut, perustelulokit).

☐ Varmista, että järjestelmä tukee ihmisen valvontaa ja ohjausta reaaliaikaisessa toiminnassa.

☐ Varmista, että toimittaja pystyy osoittamaan yhteistyömittarit (esim. Component Synergy Score).

Muut

☐ Varmista laskutuksen osalta seuraavat asiat:

- Kuinka sovelluksen laskutus toimii? Tekoälyagentit vievät paljon laskentatuloja, joten varmista, mistä laskutetaan ja miten.
- Kuinka ennakoitavia kustannukset ovat? Muutokset ja epätasaisuus agentin käytössä vaikuttaa kustannusten jakautumiseen ja ennakoitavuuteen.

☐ Tarkista myös, onko organisaatiossasi muita (tekoälyn) hallintaprosesseja, jotka juuri sinun organisaatiosi on otettava huomioon, jotta hankittu agenttiratkaisu ei vaaranna ydinliiketoimintaa. Esimerkkejä alakohtaisista huomioitavista asioista:

- Kriittisen infrastruktuurin turvallisuussääntely.
- Hoitotyön eettiset periaatteet ja lääkinnällisten laitteiden sääntely.
- Demokraattisilta prosesseilta vaadittu läpinäkyvyys ja lainmukaisuus (ml. hallintolain vaatimukset automaattisesta päätöksenteosta).



Johdon tiivistelmä:

<https://www.traficom.fi/files/media/publication/Traficom-tekoalyagentit-johdon-tiivistelma-2026%20%281%29.pdf>

Koko ohjeistus:

https://traficom.fi/files/media/publication/TRAFICOM_Teko%C3%A4lyagenttien_kyberturvallisuus.pdf

Kiitos!

aleksi.blomqvist@traficom.fi

