

Tekoälyagentit alkuvaiheen ohjelmistostartupeissa

Tekoälyagentit eli autonomiset ohjelmistorobotit, jotka hyödyntävät tekoälyä saavuttaakseen niille annetut tavoitteet. Toisin kuin tavalliset chatbotit tai co-pilot-työkalut, agenttimaiset tekoälyjärjestelmät pystyvät itsenäisesti suunnittelemaan tehtäviä, tekemään päätöksiä ja oppimaan aikaisemmasta tavoitteidensa saavuttamiseksi. Deloitte määrittelee autonomiset generatiiviset tekoälyagentit ratkaisuksi, jotka pystyvät suorittamaan monivaiheisia tehtäviä vähäisellä tai olemattomalla ohjauksella. Tällaisilla agenteilla on niin sanottu *agency*-ominaisuus, eli ne voivat toimia oma-aloitteisesti, pilkkoa isompia tehtäviä pienempiin ja saattaa prosessit loppuun asti itsenäisesti. Tyypillisiä kyvykkyyksiä ovat päättely, pitkäkestoinen muisti, oman toiminnan suunnittelu ja nykyään myös apuvälineiden, kuten ohjelmistojen tai toisten agenttien, käyttö (Stryker, C s.a.)

Tässä työssä sovellan teknologian kypsyysasteen eli Technology Readiness Level (TRL) viitekehystä tekoälyagentti teknologian analysoimista alkuvaiheen ohjelmistostartup kontekstissa. Alkuvaiheen startupit ovat pieniä kehitystiimejä, jotka ovat ketteriä julkaisemaan, mutta niitä rajoittaa erittäin rajalliset resurssit. Tällaisissa tiimeissä yksittäiset kehittäjät joutuvat usein venymään moniin rooleihin koodaamisen lisäksi, kuten asiakastukeen, liiketoiminnan analytiikkaan ja myyntiin. Oman ohjelmistostartupin näkökulmasta toistuvat ja rutiininomaiset tehtävät vievät suhteettoman paljon aikaa. Ihannetilanteessa kehittäjät voisivat käyttää enemmän aikaa uusien ominaisuuksien suunnitteluun, kun tekoälyagentit hoitaisivat toistuvan ja puuduttavan työn. Tekoälyagenttien hyödyntäminen varhaisen vaiheen ohjelmistostartupeissa ei kuitenkaan rajoitu pelkästään ohjelmistokehitykseen. Tekoälyagentit voisivat olla potentiaalisia ”digitaalisia kollegoita”, jotka voisivat sekä tukea ohjelmistokehitystä, että keventää asiakastuen, analytiikan, markkinoinnin ja hallinnon rutiineja.

Valitsin TRL-viitekehysten, koska se tarjoaa strukturoidun tavan arvioida uuden teknologian kypsyysta ideasta tuotantokäyttöön. TRL-asteikko 1-9 on vakiintunut mm. NASA:lla ja EU:ssa kuvaamaan, kuinka pitkällä kehityksessä teknologia on. TRL 1 tarkoittaa perustutkimusvaihetta, kun taas TRL 9 merkitsee täysin kypsää teknologiaa, joka on osoitettu toimivaksi operatiivisessa ympäristössä (National Aeronautics and Space Administration 2023). Näen tämän olevan erittäin hyvä viitekehys tekoälyagenttien valmiusasteen arviointiin, koska voimme yrittää sijoittaa tekoälyagenttien kehityksen tietyille tasolle tai tasoille ja perustella miksi ne eivät ole laajasti käytössä alkuvaiheen startupeissa tai missä rajatuissa tapauksissa ne jo ovat. TRL-viitekehys auttaa siivilöimään löydettyjä havaintoja ja uutisia yhtenäisen linssin läpi. Muita mahdollisia arviointikehikkoja olisivat olleet esimerkiksi Gartnerin hype-käyrä tai teknologian diffuusiomalli, mutta TRL sopii paremmin syväluotaavaan analyysiin, koska se keskittyy konkreettiseen tekniseen valmiusasteeseen hehkutuksen sijaan.

4.12.2025

Tekoälyagenttien kehitysvaiheen analyysi TRL:n näkökulmasta

Tekoälyagentit ovat jo tehneet liikkeitä ideasta kohti tuotantokypsyyttä, joita voidaan liittää TRL tasoihin. Vuonna 2022 on ensimmäiset maininnat generatiivisen tekoälyn agenttimaisesta toiminnasta ja sen tutkimuksesta nostivat päätään. Suurten kielimallien tuleminen suuren kansan tietoisuuteen loi pohjan agenttimaisen toiminnan idealle. Tutkimusyhteisö alkoi hahmotella, miten AI voisi suorittaa monivaiheisia tietokoneen käyttötehtäviä ihmisen puolesta. Esimerkkinä DeepMindin kokeilu, jossa AI tarkkaili ihmisen näppäimistö- ja hiiritoimintoja oppiakseen imitoimaan niitä, esimerkiksi varatakseen lennon tai hotellin. Tällöin TRL tasossa tekoälyagentit asettuivat joko toiselle tai kolmannelle tasolle, koska tämä edustaa vielä "proof-of-concept" tasoa laboratoriossa. Peruseriaatteet tekoälyagenttien toiminnasta olivat havaittuina, mutta käytännön toteutus oli vasta alustava.

Vuonna 2023 OpenAI:n GPT-4 mallin julkaisu mahdollisti ensimmäiset suositut autonomiset agentit kehittäjäyhteisössä. Ensimmäinen suurta suosiota saanut tekoälyagentti ohjelmistokehitykseen oli keväällä 2023 julkaistu AutoGPT-niminen avoimen lähdekoodin projekti. Se yritti ketjuttaa GPT-4:n päätöksiä ja luoda siitä koherentti itsenäinen toimija. AutoGPT saavutti nopeasti kymmeniä tuhansia tähtiä Githubissa, mutta kuitenkin kokemus osoitti, että nämä agentit olivat enimmäkseen vain hauskoja kokeiluja pienelle kehittäjäryhmälle. Kuten Jim Fan, NVIDIA:n robotiikkajohtaja, totesi viestintäpalvelu X:ssä: *"Omien kokeilujeni perusteella AutoGPT pystyy hoitamaan joitakin yksinkertaisia ja hyvin rajattuja tietotehtäviä kohtuullisen hyvin, mutta on epäluotettava suurimman osan ajasta vaikeammissa tehtävissä, jotka olisivat oikeasti hyödyllisiä. Minua huolestuttaa myös paljon aina, kun sille annetaan Pythonin suoritusoikeus ja pääsy levyille. On todella huono idea laittaa se täysin autopilotille (kuten tekijät itsekkin varoittivat). Kaikkiin tuotteisiin, jotka väittävät käyttävänsä AutoGPT:tä koodin suorituksella, pitäisi suhtautua erittäin varauksella."* AutoGPT ja vastaavat toteutukset kuitenkin osoittivat konseptin toimivuuden, agentti pystyi esimerkiksi debugaamaan koodia annetun tavoitteen pohjalta, mutta rajallisuudet tulivat nopeasti vastaan. Rajoitteita olivat mm. muistamattomuus, jossa agentti ei muistanut enää aiempia ratkaisujaan pitkissä tehtävissä, virhealttius, sekä tarve käyttäjän jatkuvalle valvonnalle. TRL-mielessä vuoden 2023 agentit olivat laboratoriotyyppejä, eli TRL tasolla 4. Useita komponentteja oli integroitu ja niitä testattiin suljetuissa ympäristöissä, mutta realistisissa olosuhteissa menestyminen oli vielä epävarmaa.

Vuonna 2024 sekä startupit että isot ohjelmistotalot sijoittivat voimakkaasti agenttitekнологiaan. Deloitte raportoi, että vuosina 2022–2024 tekoälyagentti startup yrityksiin sijoitettiin yli 2 miljardin dollarin edestä riskipääomaa (Stryker C s.a.). Samalla useat toimijat julkistivat omia agenttialustojaan. Esimerkiksi Cognition Labs julkaisi maaliskuussa 2024 Devin-nimisen autonomisen "ohjelmistoinsinööri"-agentin, jonka tavoite oli hoitaa itsenäisesti kokonaisten sovellusten suunnittelu, koodaus, testaus ja korjaus pelkästään ihmisen luonnollisella kielellä annettujen ohjeiden perusteella. Ensimmäisiä pilottiprojekteja tuotantoympäristössä nähtiin, kun esimerkiksi startup Adept AI ilmoitti saaneensa agenttinsa beta-asiakkaille käyttöön tuotantokriittisissä ympäristöissä ja automatisoineensa useiden kymmenien työvaiheiden prosesseja alusta loppuun. (Jackson 4.7.2024). Tämä kertoo, että tietyissä kapeissa käyttötapauksissa on saavutettu TRL 7, eli lähes lopullinen järjestelmäprototyyppi on osoitettu toimivaksi, mutta sekin vain yrityksen itsensä mukaan. Samalla on hyvä huomioida, että moni 2024

4.12.2025

aikana julkistettu agenttiratkaisu oli vasta demo vaiheessa, eikä valmis laajaan yrityskäyttöön. Agenttitekniologia kuitenkin ohjelmistokehityksen kontekstissa oli vielä vuoden 2024 aikana varhaisessa kehitys- ja omaksumisvaiheessa, vaikka demonstroinnit olivat vaikuttavia, agentit tekivät virheitä, jäivät jumiin silmukoihin tai jopa levittivät "hallusinaatioita" agenttien välillä monen agentin järjestelmissä. Tyypillisesti jokaisessa vaiheessa vaadittiin ihminen tarkistamaan agentin tekemiä päätöksiä ennen toimeenpanoa. Kokonaisuutena tekoälyagentit olivat siis vaiheessa, jossa teknologiaa testattiin relevantissa ympäristössä ja ensimmäisiä operatiivisia kokeiluja käynnistettiin, eli TRL tasolla 5 tai 6. Yksittäiset projektit saattoivat yltää demottavaan prototyyppiin oikeassa käytössä, mutta tämä ei edustanut alan kypsyyttä valitussa kontekstissa.

Kuluvana vuonna 2025 odotukset tekoälyagenttien suhteen ovat korkealla, mutta varovaisuus säilyy. Deloitte ennustaa, että vuoden 2025 loppupuoliskolla joissakin organisaatioissa tietyt agenttisovellukset voivat siirtyä osaksi yrityksen prosesseja. Moni yritys on myös ryhtynyt hankkimaan agenttitekniologiaa startupeilta lisenssien tai yritysostojen kautta, kuten Amazon, joka lisensoi Adeptin agenttialustan ja palkkasi sen ydintiimiä keväällä 2024 vauhdittaakseen omia AGI tavoitteitaan. Silti haasteena on luottamus ja luotettavuus. Jotta yrityksen varsinkin ohjelmistokehityksessä ottaisivat agentin kriittiseen käyttöön, sen on toimittava lähes virheettömästi. Kokonaiskuva on kuitenkin se, että tekoälyagentit eivät vielä vuonna 2025 ole yleisessä tuotantokäytössä ilman ihmisen valvontaa. Alan uusimmatkin autonomiset ohjelmistoinfinöörit tekevät yhä liikaa virheitä hoitaakseen edes tarkoin rajattuja kehitystehtäviä täysin itsenäisesti. Esimerkiksi Devin-agentin ohjelmistobugien ratkaisuaika oli yrityksen oman vertailun mukaan vain n. 14%, joten se ei vielä ole lähellekään täysin autonominen (Cognition Labs 2024). Tällä hetkellä aikaisen startuppien ja suurien tekoälyjättien huomio keskittyy agenttien autonomisuuden parantamiseen, jotta ihmiskehittäjät voivat uskoa osan työkuormastaan agentin harteille. Yleisellä tasolla tällä hetkellä tekoälyagentit ohjelmistokehityksessä ovat nyt TRL 6 ja lähellä TRL 7 tasoa Täysi tuotantovalmius on kuitenkin vielä kaukana. Agenttien pitää kehittyä vielä huomattavasti todellisissa olosuhteissa, ennen kuin niiden voidaan katsoa olleen täysin kypsiä kriittiseen käyttöön.

Tekoälyagenttien konsepti on ylittänyt alkuvaiheen tutkimuksen kynnyksen ja on edennyt prototyyppien ja pilotoitien kautta lupaaviin, joksikin rajattuihin tuotantokokeiluihin Tekniologia on siis noussut "labrasta lattialle" mutta ei vielä laajemmin yritysten toiminnan peruspalikaksi.

Tekoälyagentit varhaisen vaiheen ohjelmistostartupeissa

Varhaisen vaiheen ohjelmistostartupissa tekoälyagenttien keskeisin käyttöalue ohjelmistokehityksen näkökulmasta on koodin kirjoittamista, testaamista sekä teknisen dokumentaation tuottamista. Agentti voi toimia eräänlaisena digitaalisena juniorikehittäjänä, jolle delegoidaan rajattuja ja hyvin määriteltäviä tehtäviä, kuten yksinkertaisia toiminnallisuuksia tai olemassa olevan koodin testien generointia. Näissä tehtävissä tekoälyagentti kukee tuottamaan nopeasti päälipuolisin toimivia ratkaisuja, joita ihminen voi jatkojalostaa. Tämä vapauttaa kehittäjien aikaa korkeamman tason suunnitteluun.

Tekniologian kypsyydystason näkökulmasta tekoälyagentit ovat jo toimivan prototyypin tasolla ja niitä on jo demonstroitu relevantissa toimintaympäristössä. Perusteina tälle ovat laboratoriorio- ja pilottiesimerkit vuosilta 2023-2024 joissa agentit ovat suoriutuneet

4.12.2025

rajatuista kehitystehtävistä, kuten koodin generoinnista ja testien suorittamisesta oikeassa kehitysympäristössä. Yksittäisissä tapauksissa, kuten Adept AI:n asiakkailla, on jopa ylitetty TRL 7 -taso, kun agentin on raportoitu toimivan tuotantokriittisissä prosesseissa ja automatisoivan useiden kymmenien työvaiheiden kokonaisuuksia. Tämä osoittaa, että tietyissä kapeasti rajatuissa käyttötapauksissa tekoälyagentti on jo lähes tuotantovalmis teknologia. Tässä kuitenkin on muistettava, että tämä raportointi on Adept AI:n omaa viestintää, eikä sitä ole vahvistettu kolmannen osapuolen testaamana.

Varhaisen vaiheen ohjelmistostartupin näkökulmasta tämä ei kuitenkaan tarkoita vielä teknologista kypsyttää. Ohjelmistokehitys pienessä tiimissä on vahvasti sidoksissa liiketoimitalogiikkaan, asiakasymmärrykseen ja jatkuvasti muuttuviin vaatimuksiin. Tämän vuoksi tekoälyagentti ei vielä nykykypsytydellään pysty ottamaan vastuulleen kriittisiä kokonaisuuksia tai edes sen palasia, kuten hinnoittelulogiikka, tai ulkoisten järjestelmien integraatiota ilman ihmisen valvontaa. Nämä osa-alueet sisältävät erittäin paljon hiljaista tietoa, jota ei ole dokumentoitu ja jota tekoäly ei kykene luotettavasti päättämään pelkän koodin tai tekstiohjeiden perusteella.

Omassa startupissa Fiksatenissa hyödynnetään jo nyt joitakin tekoälyagentteja tuottavuuden parantamiseen. Tiimimme on ottanut käyttöön OpenAI:n Codexiin perustuvan komentorivityökalun, joka suorittaa tehtäviä agenttimaisesti suoraan koodikannassa kehittäjän luonnollisella kielellä antamien ohjaiden perusteella. Käytännössä tämä näkyy esimerkiksi siinä, että agentti voi luoda uusia tiedostoja, muokata olemassa olevaa koodia, suorittaa testejä ja ehdottaa korjauksia ilman, että kehittäjän tarvitsee kirjoittaa kaikkea käsin. Fiksaten hyödyntää myös asiakaspalvelussaan tekoälyagentteja. Asiakaspalvelua ei ole ulkoistettu kokonaan tekoälylle vaan agentti luo asiakaspalvelun viesteistä kokonaisuuksia, kategorisoi niitä sekä ehdottaa asiakaspalvelun henkilölle mahdollisia vastauksia edellisten kokemusten perusteella. Nämä ovat konkreettisia esimerkkejä siitä, miten tekoälyagentti voi kasvattaa pienen tiimin kapasiteettia ilman lisähenkilöstöä.

Samalla oma kokemukseni kuitenkin osoittaa selvästi, että tekoälyagentit eivät vielä nosta teknologiaa TRL 9 -tasolle. Jokainen agentin tuottama koodirivi edellyttää edelleen ihmisen tekemää tarkistusta ja testausta ennen tuotantoon viemistä. Ohjelmistokehityksessä ja asiakaspalvelussa tekoäly toimii lähinnä teknisenä apurina, ei itsenäisenä päätöksentekijänä.

Yhteenvetona voidaan todeta, että tekoälyagentit ovat varhaisen vaiheen ohjelmistostartupeissa tällä hetkellä kypsyystasolla noin TRL 6, eli teknologia on osoittanut toimivuutensa relevantissa ympäristössä rajatuissa käyttötapauksissa. Yksittäisissä piloteissa on yletty jopa TRL 7 -tasolle, mutta laajamittaiseen ja kriittiseen tuotantokäyttöön ilman jatkuvaa ihmisen valvontaa agentit eivät vielä sovellu. Käytännössä agenttien rooli nykytilassa on toimia kehittäjien tuottavuuden vahvistajina ja rutiinityön automatisoijina, ei itsenäisinä ohjelmistokehityksen vastuunkantajina.

Lähteet:

Loucks J, Crossan G, Sarer B, Widener C, Bucaille A 9.11.2024. Autonomous generative AI agents: Under development. Deloitte Center for Technology Media & Telecommunications Luettavissa:

4.12.2025

<https://www.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2025/autonomous-generative-ai-agents-still-under-development.html> Luettu: 5.12.2025

Stryker C. s.a. What is tool calling. IBM Think. Luettavissa: <https://www.ibm.com/think/topics/tool-calling> Luettu: 8.12.2025

Kapuściński, M. 11.9.2025. Technology Readiness Levels (TRL) in Space Projects – Explanation and Significance. TTMS Blogi. Luettavissa: <https://tms.com/technology-readiness-levels-trl-in-space-projects-explanation-and-significance>. Luettu: 5.12.2025.

National Aeronautics and Space Administration. 2023. Technology Readiness Levels. Luettavissa: <https://www.nasa.gov/directorates/somd/space-communications-navigation-program/technology-readiness-levels/>. Luettu: 5.12.2025

Humphreys, P. Raposo, D. Pohlen, T. Thornton, G. Chhaparia, R. Muldal, A. Abramson, J. Georgiev, P. Santoro, A. Lillicrap, T 2022. A Data-Driven Approach for Learning to Control Computers. DeepMind. London. Luettavissa: <https://arxiv.org/pdf/2202.08137>. Luettu: 8.12.2025

Jackson, A. 4.7.2024. What the Adept Hires Will Mean for Amazon's AI Development. AI Magazine. Luettavissa: <https://aimagazine.com/ai-strategy/what-the-adept-hires-will-mean-for-amazons-ai-development>. Luettu: 8.12.2025.

Cognition Labs. 2024. SWE-bench technical report. Luettavissa: <https://cognition.ai/blog/swe-bench-technical-report>. Luettu: 8.12.2025

Jim Fanin viestintäpalvelu X:n postauksen teksti on suomennettu OpenAI:n GPT-5.1 mallilla käyttäen promptia: "Suomenna: AutoGPT just exceeded PyTorch itself in GitHub stars (74k vs 65k). I see AutoGPT as a fun experiment, as the authors point out too. But nothing more. Prototypes are not meant to be production-ready. Don't let media fool you - most of the "cool demos" are heavily cherry-picked: In my experiments, AutoGPT can solve certain simple & well-defined knowledge tasks well, but is unreliable *most of the time* for harder tasks that are truly useful. I also worry a lot whenever I give it python execution and disk access. It's a *terrible* idea to have it on autopilot (as the authors warned too). You should be very wary of any product that claims to use AutoGPT with code execution. Much of the unreliability can be attributed to GPT-4's inherent limitations. I don't think these can be fundamentally addressed by fancier prompting tricks, without having access to GPT-4's weights and finetuning more. I do want to applaud @SigGravitas for putting together many great ideas (e.g. ReAct, Reflexion, memory, self-cloning) into a very neat repo that people can easily experiment with. @SigGravitas has tried the best to overcome GPT-4's limitations. But just like no amount of prompting could turn GPT-3 into GPT-4's capability, I don't think AutoGPT + a frozen GPT-4 can magically and reliably solve complex decision making that matters. The current media hype is pushing the project to completely unrealistic expectations."

