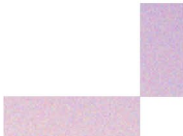


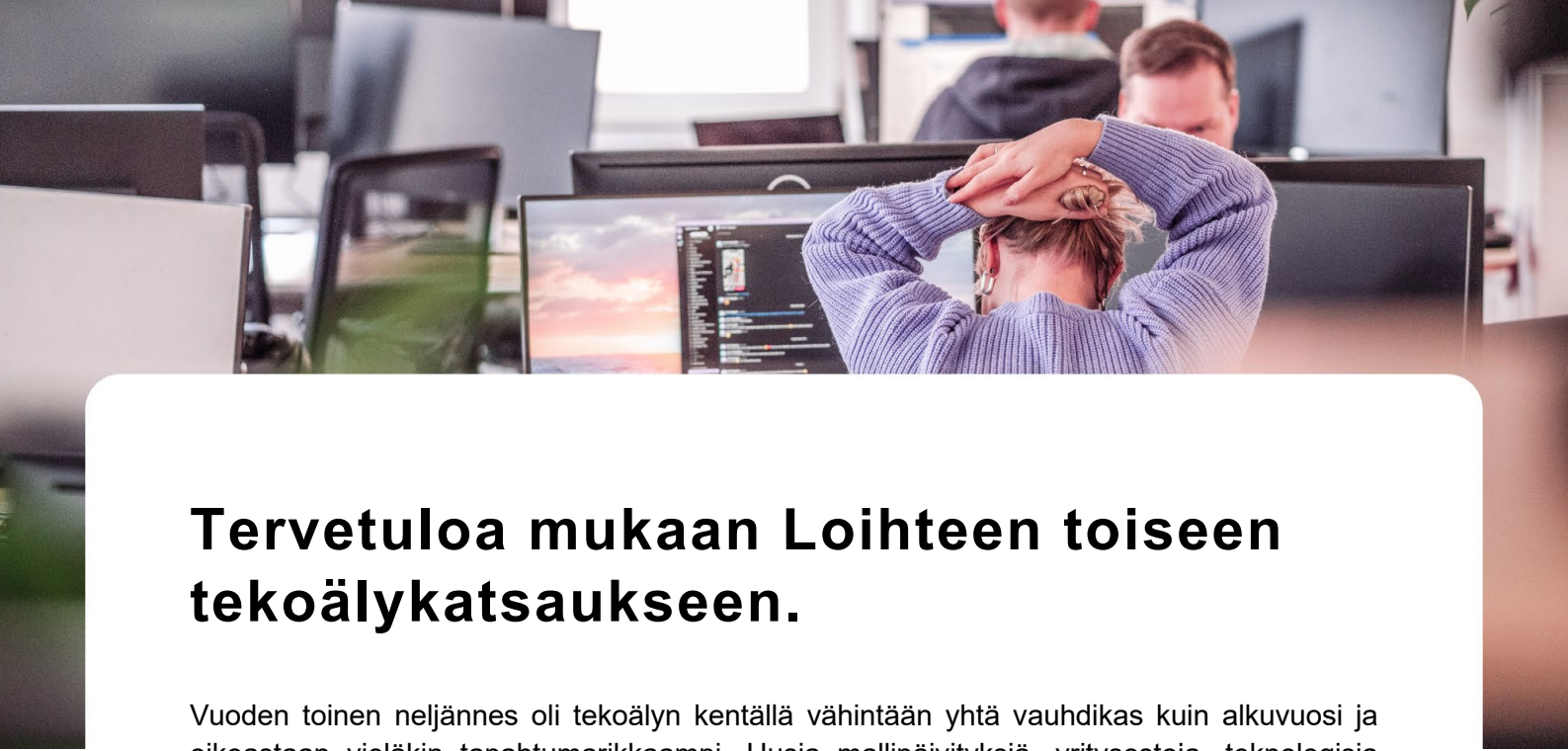


Q2
2025

Tekoälykatsaus

LOIHDE





Tervetuloa mukaan Loihteen toiseen tekoälykatsaukseen.

Vuoden toinen neljännes oli tekoälyn kentällä vähintään yhtä vauhdikas kuin alkuvuosi ja oikeastaan vieläkin tapahtumarikkaampi. Uusia mallipäivityksiä, yritysostoja, teknologisia standardeja ja strategisia linjauksia nähtiin sellaisella tahdilla, että meillekin tuli jälleen kerran valinnan vaikeus: mitkä ilmiöt ansaitsevat tulla nostetuiksi tähän koosteeseen?

Päätimme lopulta keskittyä kehityskulkuihin, jotka kertovat meille jotain muutoksen suunnasta. Tärkeämpää kuin yksittäinen uutuuus on se, minkälaista maaperää nyt valmistellaan. Siksi valitsimme katsaukseen pääosin ilmiöitä, joissa näkyy siirtymä pois yksittäisistä kokeiluista kohti pysyviä käytäntöjä. Esimerkiksi ohjelmistokehityksen ja agenttien osalta puhutaan nyt jo jokseenkin tuotantovalmiista ratkaisuista, joiden vaikutukset ulottuvat tiimitasolta koko organisaation toimintamalleihin.

Katsauksen kokoaminen on ollut jälleen sekä kiehtova että haastava prosessi. Ajatukset ovat pyörineet siinä, mikä on oikeasti merkittävää, mikä vain "hienoa", ja milloin kehityskulku muuttuu ilmiöksi. Vaikka paljon jäi väistämättä ulkopuolelle, uskomme että näiden nostoiden kautta saa hyvän kuvan siitä, mihin suuntaan ollaan menossa, niin teknisesti kuin liiketoiminnallisestikin.

Toivottavasti katsaus tarjoaa oivalluksia, herättää keskustelua ja auttaa pysymään mukana siinä, miten nopeasti tekoälyn soveltaminen etenee. Jos jokin merkittävä ilmiö jäi puuttumaan, tiedät missä meidät löytää.

Patrick Qvick
Tarjoomajohtaja, tekoälypalvelut
Data, Digi & AI

patrick.qvick@loihde.com



Merkkejä muutoksesta: Yritykset siirtyvät kohti “tekoäly ensin” -strategiaa

Lyhyesti:

- Markkinassa on nähtävissä selvä siirtymä kohti tekoäly-ensin -ajattelua: Duolingo ja Shopify ottivat käyttöön linjauksen, jossa tehtävää ei anneta ihmiselle, jos se voidaan suorittaa tekoälyn avulla. Tämä näkyy rekrytoinnissa, suoritusten arvioinnissa ja organisaation toimintamallissa.
- Microsoftin vuoden 2025 Work Trend Index nosti esiin ns. Frontier Firms – organisaatiot, joissa tekoälyagentit ja ihmiset muodostavat toiminnan ytimen.
- Klarna vähensi ensin 700 asiakaspalvelijan tarvetta OpenAI-pohjaisen chatbotin avulla, mutta palkkasi pian takaisin noin 300 työntekijää – tekoäly ei yksin riittänyt kattamaan kaikkia asiakastarpeita.

Vuoden 2025 toisella neljänneksellä useat yritykset tekivät selkeitä strategisia linjauksia: tekoäly ei ole enää pelkkä tuki, vaan ensisijainen ratkaisu liiketoiminnan kehityksessä.

Duolingon toimitusjohtaja Luis von Ahn vertasi huhtikuussa tekoälystrategiaa firman aiempaan mobile-first-päätökseen. Kun siirtymä mobiiliin toi aikanaan 100 miljoonaa kuukausikäyttäjää, nyt yhtiö hakee samanlaista kasvua tekoälyn avulla. Von Ahnin mukaan Duolingo julkaisi 12 vuodessa 100 kielikurssia – ja nyt 148 kurssia vuodessa tekoälyn tukemana. Samalla yhtiö nosti vuoden 2025 liikevaihtoennustettaan, mikä heijastui positiivisesti osakekurssiin.

Shopify meni vielä pidemmälle: sisäisessä ohjeistuksessa edellytetään, ettei uutta työntekijää palkata, ellei ensin osoiteta, että tehtävää ei voi toteuttaa tekoälyllä. Toimitusjohtaja Tobi Lütken mukaan AI:n käyttö vaikuttaa jatkossa suoraan suoriutumisen arviointiin. Washington Postin mukaan taustalla on tarve parantaa tehokkuutta ja sopeutua verkkokauppamarkkinan kiristyiin katteisiin.

Microsoft toi ilmiön esiin tutkimusdatalla: ns. Frontier Firms, eli organisaatiot, jotka hyödyntävät AI-agentteja kaikessa tekemisessään, raportoivat korkeampaa tuottavuutta ja työn merkityksellisyyttä. Build-tapahtumassa esitelty frontier-työpaikka rakentuu Copilot-ekosysteemin, organisaatiokohtaisesti koulutetun kielimallin ja Agent2Agent-protokollan varaan. Tavoitteena on rakentaa yhtä saumatonta yhteistyötä ihmisten ja tekoälyn välille.

Ruotsalainen maksupalveluyhtiö Klarna korvasi keväällä 2025 asiakaspalvelussa merkittävän osan ihmistyöstä tekoälyllä: OpenAI-teknologiaan perustuva chatbot hoiti noin 700 työntekijän edestä asiakasviestintää, kattaen noin kaksi kolmasosaa chateista. Pian selvisi, ettei tekoäly yksin riittänyt, erityisesti monimutkaiset tai tunnepitoiset asiakastilanteet vaativat edelleen ihmisen panosta. Yritys päätyi palkkaamaan takaisin noin 300 asiakaspalvelijaa.

Tapaus osoittaa, että vaikka tekoäly tuo merkittäviä tehostusmahdollisuuksia, ihmisen rooli asiakaskokemuksen turvaamisessa säilyy ratkaisevana.

Miksi tämä on merkittävää?

- Strateginen käännekohta: "Tekoäly-ensin" muuttaa päätöksenteon oletuksen: ensin kokeillaan tekoälyä, vasta sitten mietitään perinteinen ratkaisu. Tämä muuttaa kustannus-, rekry- ja investointiprioriteetteja
- Kilpailuedun kertaantuva logiikka, Frontier-ajattelun ydin on, että AI:n skaala-hyöty kertaantuu: kun jokainen tiimi käyttää agentteja, koko organisaatio oppii ja parantaa prosessejaan verkostomaisesti.
- Kulttuurinen muutospaine: työntekijöiltä edellytetään tekoälylukutaitoa ja kykyä automatisoida omat rutiininsa. Tämä muuttaanee osaamisvaatimuksia ja HR-politiikkaa nopeasti.
- Esimerkki leviää: Kun pörssilistatut yhtiöt (DUOL, SHOP) osoittivat kurssinousua ja parempaa ohjeistusta AI-strategian siivittämisenä, paine kasvaa kilpailijoilla seurata perässä.

Aiheesta lisää:

- [Duolingo's 'AI-First' Strategy: Explained](#)
- [Get 100X the Work Done': Shopify CEO Tells Employees to Try AI to Get Work Done Before Asking for More Human Workers](#)
- [2025: The Year the Frontier Firm Is Born](#)
- [Klarna hiring back employees](#)

Vuoden 2025 toinen neljännes toi markkinoille useita uusia tekoälymalleja

Lyhyesti:

- OpenAI, Anthropic, Google, Meta ja DeepSeek toivat Q2:n aikana markkinoille useita merkittäviä mallipäivityksiä.
- Uudet julkaisut laajensivat kontekstirajoja, paransivat päättelykykyä ja loivat parempaa perustaa agenttipohjaisille järjestelmille.
- Myös open source -mallien kehitys otti pieniä mutta tärkeitä askelia eteenpäin.

Vuoden 2025 toinen neljännes oli vilkas generatiivisen tekoälyn kehityksen suhteen. Useat teknologiajätit julkaisivat entistä kehittyneempiä malleja, jotka laajensivat tekoälyn sovellusalueita ja nostivat kilpailun aivan uudelle tasolle.

OpenAI julkaisi huhtikuussa GPT-4.1-sarjan (chatGPT:ssä käytössä), jossa kontekstiraja kasvoi 1 miljoonaan tokeniin ja suorituskyky monissa agentti-tyyppisissä tehtävissä parani. Kesäkuussa julkaistiin uusi o3-pro-malli, jota OpenAI itse kuvaa "parhaaksi päättelymallikseen". Se tuo yhteen useita kyvykkyyksiä, kuten verkkohaku, työkalujen käyttö ja visuaalinen ymmärrys. Julkaisun yhteydessä OpenAI laski merkittävästi (noin 80%) o3-mallin hintoja API-rajapinnassa, mikä laajensi sen saatavuutta kehittäjille ja yrityksille.

Anthropic julkisti toukokuun lopulla Claude 4 -sukupolven (Opus 4, Sonnet 4), joka esittelee hybridimallin agenttikäyttöön, yhdistäen laajan kielimallin, suunnittelukyvyyn ja työmuistin. Claude 4 Agent API avattiin kehittäjille, mahdollistaen "valmiiden" agenttien käyttöönoton yritysten omissa prosesseissa.

Myös open source -rintamalla nähtiin kehitysharppauksia. Meta julkaisi Llama 4 -sukupolven preview-version, jossa Behemoth 2T -malli tähtää skaalautuvaan sovelluskehitykseen ja Maverick-malli toimii kevyempänä vaihtoehtona.

Merkittävä julkaisu nähtiin myös DeepSeekiltä, jonka R1-malli toimii yhdellä GPU:lla, mutta yltää huipputason suorituksiin useissa koodaus- ja päättelytehtävissä. Malli on suunnattu tehokkaaseen käyttöön ilman suurta laskentainfrastruktuuria ja se julkaistiin täysin avoimena.

Miksi tämä on merkittävää?

- Demokratisoituminen: Tehokkaat, mutta kevyet mallit tuovat tekoälyn yritysten ja kehittäjien ulottuville ilman raskasta infraa.
- Agenttivalmius: Claude 4, Gemini 2.5, GPT4.1 ja o3 pro malleissa on ominaisuuksia, jotka mahdollistavat entistä autonomisten agenttien rakentamisen
- Kontekstin laajentuminen: 1M+ tokenin kontekstit mahdollistavat yhä suurempien kokonaisuuksien hallinnan, kuten pitkät asiakirjat ja keskusteluketjut
- Avoimuus vauhdittaa innovaatioita: Avoimet mallit, kuten Llama 4 ja DeepSeek R1, tarjoavat kehittäjäyhteisölle uusia mahdollisuuksia sovellusten rakentamiseen ilman riippuvuutta kaupallisista lisensseistä.

Aiheesta lisää:

- [Introducing OpenAI GPT 4.1](#)
- [OpenAI o3pro](#)
- [Introducing Claude 4](#)
- [Google Gemini 2.5](#)
- [Amazon Nova](#)
- [The Llama 4 herd](#)
- [Deepseek R1](#)

Ohjelmistokehitykseen erikoistuneet tekoälyagentit ottivat isoja harppauksia

Lyhyesti:

- Uudet koodausagentit: Google Jules, GitHub Copilot Agent ja OpenAI Codex (2025) pystyvät jo itsenäisesti etsimään virheitä, korjaamaan koodia ja ehdottamaan muutoksia.
- Pienet, lokaalisti pyörivät mallit (esim. Mistral Devstral) tuovat tekoälyavun saataville pienempiin ympäristöihin, esimerkiksi lokaalisti
- Anthropic Claude 4 nosti koodausagenttien tarkkuuden uudelle tasolle, mikä kiristää kilpailua laadusta.
- Ohjelmistokehityksen tekoälyagentteja tarjoavan Cursorin 9,9 miljardin arvostus ja 500M\$ (ARR) vuotuinen toistuva liikevaihto sekä OpenAI:n 3 miljardin dollarin WindSurf-yritysosto osoittavat, että AI-avusteinen kehitystyö on siirtynyt kokeiluista vakavasti otettavaksi liiketoiminnaksi.

Kevään 2025 aikana ohjelmistokehityksen arkeen ilmestyi kokonainen uusi tekoälysukupolvi. Googlen huhtikuussa avaama Jules toimii kuin oma avustaja: se lukee projektin “tehtävälistat”, etsii koodista oikeat kohdat ja avaa GitHubiin valmiin pull requestin, jonka kehittäjä voi hyväksyä tai hylätä. GitHubin kesäkuussa julkaiseman Copilot Agentin kanssa kehittäjän riittää antaa luonnollinen pyyntö ja agentti kirjoittaa koodin, ajaa testit, tekee tarvittavia muutoksia, päivittää dokumentaation ja jättää ehdotuksensa katselmointiin ohjelmistokehittäjälle. OpenAI puolestaan päivitti Codexin miljoonan tokenin kontekstiin: malli pystyy käsittelemään kohtalaisen kokonaisen ohjelmistokoodikokonaisuuden yhdellä kertaa ja ymmärtää jopa kuvakaappauksia, jos bugi näkyy vain käyttöliittymässä.

Samalla syntyi yhä ketterämpiä vaihtoehtoja. Mistral Devstral on vain seitsemän miljardin parametrin kokoinen, joten se pyörii tavallisella kannettavalla. Se on erinomainen ratkaisu tiimeille, jotka eivät voi lähettää lähdekoodiaan pilveen. Kiinalaisen DeepSeekin avoin R1-malli tekee saman yhdellä GPU:lla ja on kerännyt nopeasti suuren kehittäjäyhteisön ympärilleen. Suuren kokoluokan huipulla Anthropic Claude 4 Opus nousi koodin korjaustesteissä (SWE-Bench) 80 %:n onnistumisasteeseen, ohi OpenAI:n aiemman kärkipaikan.

Kehityksen taloudellinen merkitys on jo nähtävissä. AI-avusteiseen ohjelmistokehitykseen erikoistunut Cursor (emoyhtiö Anysphere) saavutti kesäkuussa 2025 yli 500 miljoonan dollarin vuosittaiset toistuvat tulot (ARR) ja nousi 9,9 miljardin dollarin arvostukseen. OpenAI puolestaan osti kehittäjätyökaluja rakentavan WindSurfin 3 miljardilla dollarilla, joka on osoitus siitä, että ohjelmistokehityksen tekoälyagenttien ympärille rakennetaan vakavia ekosysteemejä.

Miksi tämä on merkittävää?

- Tuottavuusloikka: Tekoälyagentit automatisoivat ohjelmistokehityksen kokonaisia vaiheita – bugikorjauksista dokumentointiin – ja siirtävät työn painopisteen suunnitteluun ja luovaan ongelmanratkaisuun.
- Demokratisoituminen: Devstralin ja DeepSeek R1:n kaltaiset pienet mallit mahdollistavat tekoälyavun myös pk-yrityksille ja yksittäisille kehittäjille ilman suuria kustannuksia.
- Laatu kilpailu kiihtyy: Claude 4:n 80 % tarkkuus SWE-Bench-testissä osoittaa, että agentit voivat jo lähestyä ihmistasoista laatua.
- Tietoturva ja yksityisyys: Lokaalisti ajettavat mallit soveltuvat ympäristöihin, joissa tietojen pilveen lähettäminen ei ole mahdollista.
- Taloudellinen mittakaava: Cursorin kasvu ja WindSurfin yritysosto vahvistavat, että ohjelmistokehityksen tekoälyavusteiset ratkaisut ovat nousseet merkittäväksi liiketoimintasegmentiksi.

Lisää aiheesta:

- [Google Jules](#)
- [GitHub Copilot Agent](#)
- [OpenAI Codex \(2025-päivitys\)](#)
- [Mistral Devstral](#)
- [DeepSeek R1](#)
- [Anthropic Claude 4 \(SWE-Bench-tulokset\)](#)
- [Cursor 500 M ARR ja 9,9 miljardin dollarin valuaatio](#)
- [OpenAI hankki Windsurfin 3 miljardilla dollarilla](#)

Tekoälyagenttien läpimurtoa tukevat perusrakenteet rakentuvat vauhdilla

Lyhyesti:

- MCP (Model Context Protocol) siirtyi valtavirtaan: Microsoft toi sen Copilot Studioon ja esitteli Windows 11 -esikatseluversion, jossa MCP toimii sisäänrakennettuna. OpenAI ja Google ilmoittivat tukevansa samaa standardia omissa kehittäjätyökaluissaan.
- A2A (Agent-to-Agent) nostaa agenttien yhteistyökyvyn uudelle tasolle – Googlen avoin protokolla sai heti tuen useilta SaaS-toimijoilta.
- IBM Agent Builder & Catalog tarjoaa no-code-työkalut organisaatioiden omien agenttien luomiseen ilman AI-kehitystiimiä.

Tekoälyagenttien toiminta vaatii yhteiset tekniset standardit – kuin “pistorasiat”, joiden kautta agentit löytävät toisensa, tiedon ja työkalut. Vuoden 2025 toisella kvartaalilla kehitys otti harppauksen eteenpäin.

Model Context Protocol (MCP), joka oli aiemmin käytössä lähinnä Anthropicilla, levisi nyt koko kentälle. Microsoft integroi MCP:n Copilot Studioon ja ilmoitti Build-tapahtumassa, että myös Windows 11 saa MCP-tuen: sovellukset ja agentit voivat kutsua toisiaan kuin selain verkkosivua. OpenAI vahvisti lisäävänsä MCP:n Agents SDK:hon ja työpöytäsovellukseensa, Google puolestaan integroi tuen Gemini-malleihinsa ja kehittäjäpakettiinsa. MCP on nopeasti muodostumassa ”tekoälyn USB-C:ksi”, eli yhteensopivuuden standardiksi, joka mahdollistaa agenttien tehokkaan toiminnan.

Google julkaisi samalla uuden Agent2Agent-protokollan (A2A), joka mahdollistaa eri valmistajien agenttien turvallisen keskustelun ja tehtävien jakamisen. Mukana pilotissa ovat esimerkiksi Atlassian ja Salesforce, jotka näkevät A2A:n keinona rakentaa agenttien mikropalveluarkkitehtuuria. Esimerkiksi HR-agentti voi pyytää talousagentilta raportin tai varata kalenterista neuvotteluhuoneen.

IBM puolestaan toi watsonx Orchestrate -alustaan uuden Agent Builderin ja Agent Catalogin, joiden avulla liiketoimintakäyttäjä voi luoda oman agentin ilman koodia. Käyttäjä valitsee tietolähteet, työkalut ja ohjeet, ja Orchestrate rakentaa agentin, jonka tehtävä rakenne ja valtuudet näkyvät hallintanäkymässä. Lisäksi tarjolla on yli 100 valmista agenttia, esimerkiksi HR:n tarpeisiin.

Miksi tämä on merkittävää?

- MCP tarjoaa yhtenäisen tavan sovelluksille tarjota agenteille työkaluja ja rajapintoja, aivan kuten verkkosivut tarjoavat käyttöliittymän selaimille. Tämä vakioi agenttien tavan käyttää sovelluksia ja vähentää tarvetta räätälöidylle ymmärrykselle.
- A2A mahdollistaa agenttien välisen yhteistyön ja tehtävien jakamisen, jolloin eri toimittajien ja järjestelmien agentit voivat toimia saumattomasti yhdessä, kuin mikropalvelut ohjelmistoarkkitehtuurissa.
- Käyttönoton kynnys laskee: IBM:n no-code-työkalut ja MCP:n valmiit palvelimet tuovat agentit myös niille organisaatioille, joilla ei ole omaa AI-kehitystiimiä.
- Turvallisuus & hallinta: Standardoitu liitäntä tulee mahdollistamaan selkeät oikeudet, audit-lokit ja versiohallinnan, kriittistä säädelyillä toimialoilla.
- Ekosysteemin kiihtyvä innovaatio: Kun integraation perusta on ratkaistu, kehittäjät voivat keskittyä liiketoiminta-ideoihin, mikä nopeuttaa ohjelmistokehitystä merkittävästi

Lisää aiheesta:

- [Microsoft Copilot Studio MCP GA](#)
- [Windows 11 MCP-preview](#)
- [OpenAI ottaa MCP:n käyttöön](#)
- [Google tukee MCP:tä](#)
- [Google Agent2Agent-protokolla](#)
- [IBM watsonx Orchestrate – uudet Agent Builder & Catalog -ominaisuudet](#)

Videoita luovat tekoälymallit ottivat harppauksen

Lyhyesti:

- Google julkaisi Veo 3 -mallinsa toukokuussa 2025. Malli kykenee luomaan pidempiä 1080p-videoita synkronoidulla äänellä.

Vuoden 2025 toisen kvartaalin aikana videotekoälyn kentässä tapahtui merkittävä harppaus eteenpäin. Google DeepMind esitteli toukokuussa Veo 3:n, joka ei pelkästään tee korkealaatuisia 1080p-videoita, vaan tuottaa myös synkronoidun äänen, kuten dialogia, taustamelua ja efektejä, suoraan mallissa. Veo 3:n uudet ominaisuudet, muun muassa "Flow", mahdollistavat pidempien videototeutusten luomisen.

Esimerkkidemoissa on luotu myös jatkuvia tarinallisia sarjoja, kuten Bigfoot- ja Yeti-videoita, joissa hahmot esiintyvät toistuvasti eri kohtauksissa sisältäen synkronoituja ääniä.

Käytännön sovellukset ovat laajoja:

- Markkinoinnissa voidaan luoda tuotevideoita ilman kuvausryhmää.
- Koulutuksessa ja sisäisessä viestinnässä voidaan simuloida tilanteita tai prosesseja nopeasti.
- Pelinkehityksessä ja muotoilussa voidaan testata kohtauksia ja skenaarioita ilman ohjelmointia.
- Yritysten sisäinen tiedon visualisointi saa uuden ulottuvuuden, miltä kuulostaa esimerkiksi data-analyysin esittäminen videona

Miksi tämä on merkittävää?

- Sisällöntuotannon demokratisointi: Korkealaatuisten videoiden tekeminen ei enää vaadi tuotantotiimiä, mikä mahdollistaa laajemman sisällöntuotannon erityisesti pienille ja keskisuurille toimijoille.
- Simulaatiot ja koulutus: Organisaatiot voivat kuvata turvallisuustilanteita, käyttöohjeita tai palvelupolkuja realistisesti, ilman näyttelijöitä tai fyysisiä kuvausympäristöä.
- Monikanavaisuuden kasvu: Kun sama malli ymmärtää tekstiä, kuvaa ja liikettä, syntyy mahdollisuus automatisoituun sisällön tuottamiseen eri muodoissa (video, kuva, puhe, kaavio).
- Kilpailu laatu- ja mittakaavaedussa: Google nosti rimaa korkealle. Veo2 tuottaa käyttökelpoista videota, jonka laatu lähenee ihmistuotantoa ja pituudet alkavat riittävää yhä monipuolisempiin käyttötarkoituksiin.

Aiheesta lisää:

- [Google DeepMind Veo 3](#)

Pääkohdat

- Tekoälyn käyttö siirtyy hitaasti, mutta varmasti kokeiluista pysyväksi osaksi organisaatioiden strategiaa: yritykset rakentavat toimintamalleja, joissa AI on oletusratkaisu, ei enää vain tuki.
- Teknologinen pohja agenttien ja uusien työkalujen hyödyntämiselle rakentuu nopeasti: standardit, rajapinnat ja no-code-ratkaisut laskevat käyttöönoton kynnystä.
- Tekoäly mullistaa jo nyt esimerkiksi ohjelmistokehitystä ja videotuotantoa, ja vaikutukset tulevat näkymään tuotantotavoissa, rekrytoinnissa ja työn sisällössä.