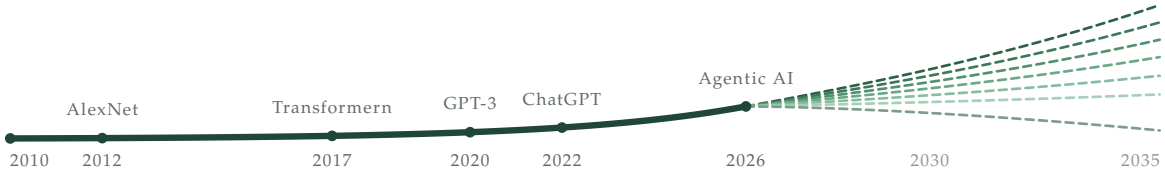

Två artiklar om artificiell intelligens

Obekväma tankar om jobb, bubblor och drönare

Jani S.

18 MINUTERS LÄSNING

20.9.2026

Del 1: Economic Scenarios for Transformative AI

Ekonomi · Arbetsmarknad · Agentic AI · EU-reglering · Bubblan

1.1. Varför ser vi inte vad som händer?

AI-utvecklingen har gjort enorma framsteg snabbare än någon annan teknisk utveckling genom historien. Varför är vi så blinda för det? Är det för att vi förknippar AI med negativa nyheter om fusk, deepfakes och propaganda, eller är det för att en framtid med AI helt enkelt är för skrämmande att tänka på?

Den första artikeln jag har valt är skriven av Santi Ruiz tillsammans med författarna till den tekniska rapporten *Economic Scenarios for Transformative AI*: Anton Korinek, Charles I. Jones, Szymon Sacher, Tess Cotter och Peter McCrory (Ruiz m.fl., 2026)[1]. Artikeln är skriven i bloggform och tar upp ämnen ur rapporten (Korinek, Jones, Sacher, Cotter & McCrory, 2026)[2]. Bloggformatet gör forskningen betydligt mera lättläst för vanliga människor.

1.2. Tre scenarier

Artikeln förutspår hur AI kommer att förändra ekonomin och arbetsmarknaden i USA. På sidan finns en interaktiv modell där läsaren själv kan ställa in hur kapabel hen tror att AI kommer att bli och hur mycket teknologin kommer att användas. Modellen ger sedan en

[1] Blogginlägget på The Anthropic Institute som del 1 handlar om, skrivet av Santi Ruiz tillsammans med rapportens fem författare. Se källförteckningen.

[2] Den tekniska rapporten (working paper) som blogginlägget bygger på. Se källförteckningen.

prognos för hur år 2030 kan se ut med de valda värdena. Det är en prediktiv modell och ingen av oss vet hur framtiden kommer att se ut.

Artikeln definierar tre scenarier för framtiden som grundar sig på hur snabbt AI-utvecklingen sker och hur snabbt industrin och arbetstagare tar i bruk de förmågor AI ger: ett måttligt, ett betydande och ett extremt scenario (Ruiz m.fl., 2026). I det måttliga scenariot får AI ungefär samma effekt på ekonomin som internet en gång fick: verkliga vinster, men inom det historiskt normala för ny teknik. I det betydande scenariot har AI en större effekt på ekonomin än vad både järnvägen och internet hade. Det extrema scenariot förutsätter rekursiv självförbättring (Recursive self-improvement, 2026)[3], det vill säga att AI själv utvecklar följande generation av AI, vilket helt och hållet kan ändra på ekonomin så som vi känner den.

1.3. Vi som borde veta bäst

Det verkar inte som om människor inom datorteknik (studerande, lärare, yrkesverksamma) fullt ut har insett hur revolutionerande och samtidigt skrämmande AI-utvecklingen är. Det finns knappast en enda programmerare som inte har chattat med en språkmodell för att fråga om, felsöka eller generera kod, men för de flesta stannar användningen där. Den allmänna uppfattningen är att AI är ett verktyg bland andra i samma kategori som de tusentals JavaScript-bibliotek programmerare använder för att skapa en så stor abstraktionsklyfta som möjligt mellan sig själva och datorns hårdvara.

De flesta tänker fortfarande på ett chattfönster när det talas om AI. Det är inte särskilt länge sedan agentic AI slog igenom. Agentic AI går ut på att ge en AI-modell en miljö, till exempel en dator, där den självständigt kan arbeta och utnyttja deterministiska verktyg. Att använda en språkmodell i ett chattfönster kan liknas vid att baka en kaka med ögonbindel. Jag tror att en orsak till att så många studerande ännu inte har upptäckt agentic AI är att det i de flesta fall kostar omkring 20 euro i månaden att få tillgång till en stark modell och de verktyg för agentic AI som AI-företagen bygger runt modellerna.

Det är också svårare för äldre datoringenjörer att ta till sig AI. De har ett sätt att göra sitt arbete som passar dem. Jag tror inte heller att de är särskilt nöjda när det kommer nyutexaminerade som med hjälp av AI kan göra det som tagit dem tiotals år av hårt arbete att lära sig.

[3]Wikipedia-artikeln om rekursiv självförbättring, senast redigerad 4.8.2026. Se källförteckningen.

Under sommaren skapade jag datorprogram som utan AI (med de kunskaper jag har) skulle ha tagit flera år att skapa. Jag använde också AI för andra datortekniska uppgifter utanför programmering. Hittills har det inte funnits en enda uppgift som jag gett åt en ny AI-modell som den inte skulle ha klarat av. Vissa uppgifter har förstås tagit längre tid än andra och i vissa fall har modellen behövt mig för att manuellt prova olika lösningar. Datorteknik är ett så brett område att ingenjörer specialiserar sig inom två-tre delområden och rör aldrig de andra. För en AI-modell finns ingen sådan gräns. Jag tror att de flesta som har använt AI för något tekniskt arbete har fått flera aha-upplevelser.

Jag tror att en orsak till att många har en negativ inställning till AI är hur Microsoft har integrerat Copilot i nästan alla sina produkter. Tanken är rätt men målgruppen (vanliga datoranvändare) fel. De flesta bryr sig inte om AI och att med våld mata teknologin till vanliga användare skapar bara irritation. Hur länge kommer det att ta för allmänheten att inse vad teknologin är kapabel till? Finlands utrikesminister Elina Valtonen är inne på samma spår i en intervju med Helsingin Sanomat (Paananen, 2026)[4].

> *”Nyt kun kehitys on näin nopeaa, on tärkeää lisätä tietämystä koko kansakunnan parissa. Tämä ei ole mikään poliittisten päättäjien kabinettijuttu.”*
>
> Elina Valtonen (Paananen, 2026)

1.4. En nagel i ögat

Sommaren 2025 arbetade jag fem månader som automationstekniker inom processindustrin. Jag började just när aktien för företaget stod som lägst och stora uppsägningar hade nyligen ägt rum. Jag lyfter på hatten för att företaget över huvud taget tog in sommararbetare. I arbetet fick jag uppleva ett företag i amerikansk stil vars främsta uppgift är att generera pengar. Jag fick se allt från extrema automationsplaner till verkligt dåliga arbetsförhållanden och en högt uppsatt chef som dök upp berusad på ett distansmöte inför ett fyrtiotal människor. För ett sådant företag är människoarbete inte en styrka utan en nagel i ögat. Jag behöver inte undra över hur ett sådant företag ser på AI-utvecklingen, för jag vet redan svaret och har tyvärr en aning om att äpplet inte faller långt från trädet hos andra publika aktiebolag.

1.5. Elektriker eller sjukskötare?

Jag tycker inte alls att det är överdrivet att förutspå att arbetslösheten kan skjuta i höjden till historiska nivåer om rekursiv självförbättring uppnås, som i det extrema scenariot artikeln hänvisar till (Ruiz m.fl., 2026). En orsak till att jag valde den här artikeln är att den skapade diskussion när den kom ut eftersom den nämner programmerare specifikt: just vi kan i

[4]Intervju med utrikesminister Elina Valtonen i Helsingin Sanomat, 4.9.2026. Se källförteckningen.

framtiden bli tvungna att byta bana och bli elektriker eller sjukskötare, yrken som är mindre utsatta för AI.

Dario Amodei ser en rimlig möjlighet att AI orsakar betydande och bestående förluster av arbeten. I essän *Policy on the AI Exponential* (Amodei, 2026b)[5] tar han upp förslag på hur samhället kan sätta plåster på såren när arbeten ersätts av AI. Han föreslår bland annat en basinkomst till alla, finansierad genom att beskatta de företag som tjänar mest på AI.

Här i Finland har det, precis som i USA, förts diskussion om att införa en automationsskatt för företag som med hjälp av AI driver automationen till en extrem nivå. Tyvärr tror jag inte att diskussionen ännu tas på allvar eftersom vi inte helt har uppfattat vad som händer. Vi får väl vänta några år och se hur saker och ting utvecklar sig, men då kan det redan vara för sent.

1.6. Europa och reglering

Det finns ingen chans för Europa att komma ikapp utvecklingen och det skapar frustration hos mig att se EU:s försök att vara en storspelare inom AI i form av uttalanden och regleringar som jag tycker gör mera skada än nytta. Reglerna biter bara på de amerikanska företagen eftersom de kinesiska inte har några planer på att följa dem. Det har förts en politisk debatt både i USA och i Europa om att förbjuda kinesiska AI-modeller. Att förbjuda modeller som körs i kinesiska datacenter är lätt, men att förbjuda kinesiska modeller med öppna vikter (som jag återkommer till i del 2) skulle vara det största masscensurförsöket någonsin.

AI-regleringar låter bra på papper när man kryddar dem med tillräckligt många positiva ord. De tal som EU-politiker har hållit har fått mig att tappa en väldigt stor del av mitt förtroende för unionen. Frågar de faktiskt inte experter inom området hur saker ligger till innan de öppnar munnen? Valtonen varnar i samma intervju (Paananen, 2026) för att reglera AI för hårt på EU-nivå.

Något som inte tas upp i artikeln, men som jag tror mycket väl kan hända, är att Europa blir beroende av en teknologi vi aldrig själva kommer att kunna utveckla. Pengarna flyter allt kraftigare till USA och Kina, och Europa blir till sist en u-kontinent.

> *”Teknologiaa koskevat säännöt ovat vaikeita tai mahdottomia ilman, että samalla rajataan meidän omaa kilpailuasetelmaamme. Ja samalla emme kuitenkaan välttämättä saisi sitä turvallisuushyötyä, mitä tavoittelemme.”*
>
> Elina Valtonen (Paananen, 2026)

[5] Dario Amodeis essä om AI-politik, juni 2026. Kapitlet om arbetsmarknaden innehåller förslagen som nämns här. Se källförteckningen.

1.7. Bubblan

Många talar om en ”AI-bubbla” som de tror kan spricka vilken sekund som helst. Bubbeltänkandet kring AI har sina rötter i att ett pris inte bara baserar sig på vad något är värt i dag, utan också på vad man tror att det kommer att vara värt i morgon. AI-företag tar stora lån och köper upp datorkapacitet innan den ens byggts. Om utvecklingen en vacker dag kör in i väggen är det möjligt att AI-företag går i konkurs.

Jag tror inte på bubblan. När en ännu icke-utgiven OpenAI-modell löste ett av de sju millennieproblemen, Navier-Stokes-ekvationerna (Louniala, 2026)[6], skapade det stor irritation bland både matematiker och vanliga människor. Förtjänar en AI-modell att få en miljon dollar för att ha löst ett problem som stått olöst i 90 år? Någon på X (före detta Twitter) tog upp vad det kostat OpenAI att köra modellerna konstant i 88 timmar. Har det faktiskt någon betydelse hur många miljoner det kostat att lösa något som kan gynna hela mänskligheten? OpenAI:s vd Sam Altman svarade själv på kritiken:

> *”ugh AI is such a bubble, i heard they are selling tokens at a loss, did they know this was only worth $1 million?”*
>
> Sam Altman (Altman, 8.9.2026)[7]

1.8. Börsraset och vädjan att bromsa

Den 14 september 2026 svängde börsen kraftigt och AI-relaterade företag som Nvidia, Nokia och Micron fick en rejäl smäll (Palosaari, 2026)[8]. Den största orsaken var Dario Amodeis (Anthropic) vädjan om att sakta ner utvecklingen och skapa hårdare regleringar för AI, något som både Sam Altman (OpenAI) och Elon Musk (xAI) höll med om (Varho, 2026)[9]. Alla tre är chefer för de företag som just nu har de mest avancerade modellerna, de så kallade frontier-modellerna. Deras motivering är att det finns en risk att mänskligheten utplånas om AI-utvecklingen fortsätter i samma takt som hittills.

Det finns riktiga, väl förklarade argument i vädjan, men för mig låter det hela så extremt att jag har svårt att tro att det är den enda bakomliggande orsaken. Jag tror att det till en viss grad handlar om skrämselpropaganda för att skapa ekonomisk vinst. Det är ändå svårt att säga hur stor del som är det ena eller det andra. Geoffrey Hinton, ”AI:ns gudfader”, tycker inte att bedömningen är orimlig när Evan Hubinger vid Anthropic uppskattar sannolikheten för att AI utplånar mänskligheten under det kommande årtiondet till tio procent (Varho, 2026).

USA:s president Donald Trump har avfärdat vädjan om att bromsa utvecklingen. Det är uppenbart att USA kommer att hålla teknologin ekonomiskt vid liv även om den osannolika

[6] Yles nyhetsartikel om OpenAI:s påstådda lösning av Navier-Stokes-problemet, 8.9.2026. Se källförteckningen.
[7] Sam Altmans inlägg på X som svar på kritiken kring millenniepriset. Se källförteckningen.
[8] Yles artikel om börsraset och Nokias kursfall, 15.9.2026. Se källförteckningen.
[9] Yles analys av AI-chefernas varningar och Trumps reaktion, 15.9.2026. Se källförteckningen.

situationen att ”bubblan spricker” skulle inträffa. Teknologin har växit sig för stor både ekonomiskt och militärt (något jag tar upp i del 2) för att staten ska hålla sig utanför.

> *”AI, and Data Centers, will be the Greatest Economic Development Engine in History - Bigger than Oil, Gold, Diamonds, or even the Internet. It will not be stopped by brilliantly run Destructive Forces during the Term of President DONALD J. TRUMP!”*
>
> Donald Trump (Trump, 14.9.2026)[10]

1.9. Robotar

En sista och kanske mest skrämmande punkt i artikeln är en av den ekonomiska modellens uttalade begränsningar: scenarier där mänskligheten utvecklar hyperkapabla robotar finns inte med i beräkningarna (Ruiz m.fl., 2026). Det är precis det som Google (Google DeepMind, 2026)[11], OpenAI och xAI samt de hundratals företag som hållit på med robotik sedan länge (till exempel ABB) nu utvecklar i rekordfart. Inget arbete, vare sig på kontoret eller i fabriken, är säkert.

Mellandel: Språk och nivå

Båda artiklarna har ett väldigt bra språk (det vore underligt om de inte hade det). Svårighetsgraden är ändå låg, speciellt för en ämnesintresserad, vilket jag har märkt i all text som Anthropic ger ut. Jag tror att de flesta Anthropic-texter till en viss grad är AI-granskade eftersom skrivstilen är igenkännbar. Presentationen i Anthropics texter är också i en helt egen klass med interaktiva illustrationer och diagram. Förstås ska ett företag som säljer en produkt som kan generera snygga webbsidor ha en snygg webbsida.

Artiklarna väcker definitivt tankar hos mig. Jag tror att de är på en sådan nivå att de mycket väl skulle kunna användas som källor i ett examensarbete och jag är säker på att de redan har använts i det syftet.

[10] Trumps inlägg på Truth Social om AI och datacenter som ekonomisk motor, publicerat samma dag som AI-chefernas vädjan. Se källförteckningen.

[11] Google DeepMinds presentationsvideo för Gemini Robotics 2. Se källförteckningen.

Basstation Autonom svärm

Del 2: Measuring tactical intelligence targeting and conventional weapons capabilities of AI models

Målbearbetning · Geolokalisering · Drönare · Öppna vikter · Autonoma vapen

3.1. Fler problem än bara jobben

Undanträngande av arbetskraft är ett av många problem mänskligheten kommer att behöva tampas med. Andra storskaliga bekymmer som AI skapar (redan i dag) är kunskap om biologiska vapen och offensiv cyberförmåga, kunskap som förut varit ”säkrare” i händerna på statliga aktörer. Det finns olika åsikter om huruvida covid-19 skapades av människor eller uppkom naturligt. På cybersäkerhetssidan kan man peka på bakdörren i XZ Utils, som Veritasium berättar om i videon *The Internet Was Weeks Away From Disaster And No One Knew* (Veritasium, 2026, 47:53)[12]. I båda fallen har en statlig aktör misstänkts ligga bakom.

I den här delen kommer jag inte att ta upp biologi eller cybersäkerhet, utan i stället en tredje aspekt som jag tror kan bli ett lika stort problem för mänskligheten som de två tidigare nämnda.

3.2. Artikeln

Den andra artikeln jag valt kommer också från Anthropic, publicerad som en forskningsartikel skriven av Anthropics Frontier Red Team som grupp (Anthropic Frontier Red Team, 10.9.2026)[13]. Inga namn på individerna bakom artikeln nämns, men man kan fråga sig om någon vill ha äran för en artikel med rubriken *Measuring tactical intelligence targeting and conventional weapons capabilities of AI models.*

Artikeln är väldigt tydlig med vad den handlar om och den animerade demonstrationen högst upp i inlägget gör det klart redan innan man börjar läsa. Frontier Red Team har skapat nya sätt att utvärdera AI-modellers förmågor inom militär målbearbetning och konventionella vapen.

[12] Veritasiums video om bakdörren i XZ Utils. Tidsangivelsen hänvisar till stället i videon. Se källförteckningen.
[13] Forskningsartikeln som del 2 handlar om, publicerad 10.9.2026. Se källförteckningen.

3.3. Anthropic, Pentagon och autonoma vapen

Dario Amodei förutspår att autonoma vapen förr eller senare kommer att skapas. Han och Anthropic råkade i början av året i gräl med den amerikanska staten och Pentagon, eftersom de inte gav Pentagon full kontroll över sina modeller. Synen på autonoma vapen och massövervakning gick för mycket isär (CBS News, 2026)[14]. Artikeln tar också upp övervakning av individer med hjälp av AI ur ett militärt perspektiv.

3.4. Bättre än ett GeoGuessr-proffs

Artikeln visar hur modeller utan hjälp av omvänd bildsökning, metadata, nätuppkoppling eller andra verktyg klarar av att lokalisera platser på jorden utifrån en bild. Spelet GeoGuessr nämns, vilket jag tycker är roligt. I GeoGuessr placeras spelaren någonstans på jorden i Googles Street View och ska sedan utifrån till exempel skyltar och arkitektur lista ut var i världen hen befinner sig. Artikeln visar att AI är bättre på att gissa platser utifrån en enda bild än vad en människa är med en hel Street View-vy som hjälp. Testet visar också hur mycket Anthropics modeller skiljer sig åt: i ett av exemplen (en pub i Kapstaden) var den smartaste modellen Mythos 5 bara 206 meter från målet, medan Opus 5 gissade på Melbourne och Sonnet 5 på Wellington, 10 314 respektive 11 309 kilometer fel (Anthropic Frontier Red Team, 10.9.2026).

3.5. Drönare i dag

I och med Ukrainakriget har drönare blivit den viktigaste delen av modern krigföring. Billiga drönare som bär en liten mängd sprängämne är den effektivaste metoden att döda, något som vi av goda skäl blundar för.

AI används redan i dag i drönare i form av mållåsning där operatören låser ett mål och drönaren sedan flyger de sista hundra metrarna på egen hand för att kringgå radiostörning. Artikeln har imponerande visuella representationer av hur det ser ut när en modell styr en drönare. Den testar också olika modellers förmåga att navigera längre sträckor utan GPS, en teknologi som visat sig vara nästintill oanvändbar i krig, eftersom den är så lätt att störa (Anthropic Frontier Red Team, 10.9.2026).

3.6. Öppna vikter

Artikeln lyfter tidigt fram oroande förmågor som teamet upptäckt i modeller med öppna vikter. En AI-modell är inte programkod som ett traditionellt, deterministiskt datorprogram, utan miljarder inlärda tal (vikter) som formats genom träning. Det en modell ger ut kommer inte från en stor databas utan skapas utifrån inlärda mönster. Samma inmatning ger inte nödvändigtvis samma utmatning. Modeller med öppna vikter kan köras lokalt på egen dator,

[14] CBS News intervju med Dario Amodei om konflikten med Trump-administrationen och Pentagon, 28.2.2026. Se källförteckningen.

men de bästa kräver i dag datorkraft för tiotusentals euro, och i full kvalitet hundratusentals. De flesta öppna modellerna kommer från Kina, men också Google, OpenAI och Meta har gett ut öppna modeller.

Eliezer Yudkowsky och Nate Soares lyfter i boken *If Anyone Builds It, Everyone Dies* (Yudkowsky & Soares, 2025)[15] fram att det i framtiden kan bli nödvändigt att reglera hur mycket datorkraft företag och privatpersoner får inneha, speciellt då öppna modeller sakta men säkert uppnår de förmågor som frontier-labbens nya modeller har.

> *”So the safest bet would be to set the threshold low — say, at the level of eight of the most advanced GPUs from 2024 — and say that it is illegal to have nine GPUs that powerful in your garage, unmonitored by the international authority.”*
>
> Yudkowsky och Soares (2025)

3.7. Svärmen

I essän *The Adolescence of Technology* lyfter Amodei (2026a)[16] fram idén om en teoretisk svärm av fullt automatiserade drönare. En sådan AI-drönarsvärm skulle kunna kopplas samman i ett meshnätverk där drönarna pratar med varandra och en eller några av dem pratar med en basstation. Trådlösa förbindelser går dock alltid att störa, så en bättre lösning vore att köra AI-modellerna lokalt på drönarna. För några år sedan skulle det ännu inte ha varit möjligt med tanke på den datorkraft som krävs, men Googles familj av öppna Gemma-modeller har visat att det går att köra förvånansvärt kraftfulla modeller på väldigt begränsade resurser, till exempel lokalt på en telefon.

Arbetet med att effektivisera AI-modeller så att de kräver mindre datorkraft går framåt hela tiden. Anthropic slår i artikeln fast att AI-utvecklingen korrelerar starkt med militär utveckling (Anthropic Frontier Red Team, 10.9.2026).

[15] Bok om riskerna med superintelligent AI, utgiven 2025. Se källförteckningen.
[16] Dario Amodeis essä om riskerna med kraftfull AI, januari 2026. Se källförteckningen.

> *”A swarm of millions or billions of fully automated armed drones, locally controlled by powerful AI and strategically coordinated across the world by an even more powerful AI, could be an unbeatable army, capable of both defeating any military in the world and suppressing dissent within a country by following around every citizen. Developments in the Russia-Ukraine War should alert us to the fact that drone warfare is already with us (though not fully autonomous yet, and a tiny fraction of what might be possible with powerful AI).”*
>
> Dario Amodei (Amodei, 2026a)

3.8. Någonting fysiskt

Slutsatsen forskarna drar i artikeln är att modeller (både öppna och stängda) mycket väl kan användas redan i dag för att ge en stor fördel i krig (Anthropic Frontier Red Team, 10.9.2026). Krigsindustrin har historiskt sett fört utvecklingen framåt, något som fortfarande gäller.

I Svenska Yles artikel *Därför varnar tunga AI-namn för att mänskligheten kan utplånas* (Fagerström, 2026)[17] intervjuas Laura Ruotsalainen, professor i datavetenskap vid Helsingfors universitet. Hon ser inte varifrån ”någonting fysiskt” som kunde döda människor skulle komma. Med lite fantasi är det inte svårt att komma på ett och annat sådant, vilket det också finns många exempel på i artikelns kommentarsfält.

När OpenAI:s modell GPT-6 kom ut var det ett stort försäljningsargument att modellen tagit ett enormt språng när det gäller 3D. Det tog inte många dagar innan människor delade videor och reflektioner kring hur bra modellen självständigt kan spela videospel. Att se en modell lösa utmaningar i ett spel är roligt och jag tror inte många drar paralleller till det hemska man kan läsa mellan raderna. Finns det någon skillnad för en modell mellan att döda människor i ett videospel och i verkliga livet? Nej. Det går att lära en modell etik och att simulera känslor, men det går lika bra att låta bli.

Det är minst sagt ironiskt att lanseringsvideon för GPT-6 (OpenAI, 2026)[18] visar hur GPT-6 Astra skapar 3D-modeller av en raket.

> *”För att döda människor behövs det ändå alltid någonting fysiskt och jag ser inte varifrån det skulle komma.”*
>
> Laura Ruotsalainen (Fagerström, 2026)

[17] Svenska Yles artikel om AI-chefernas varningar, med finländska experters kommentarer, 13.9.2026. Se källförteckningen.
[18] OpenAI:s lanseringsvideo för GPT-6 Astra, 3.9.2026. Se källförteckningen.

Källförteckning

Samtliga källor hämtade 16.9.2026.

Altman, S. [@sama]. (8.9.2026). *ugh AI is such a bubble, i heard they are selling tokens at a loss, did they know this was only worth $1 million?* [Inlägg]. X. Hämtat 16.9.2026 från https://x.com/sama/status/2097408559583772986

Amodei, D. (2026a). *The Adolescence of Technology: Confronting and Overcoming the Risks of Powerful AI* [Essä]. Hämtat 16.9.2026 från https://darioamodei.com/essay/the-adolescence-of-technology

Amodei, D. (2026b). *Policy on the AI Exponential* [Essä]. Hämtat 16.9.2026 från https://darioamodei.com/post/policy-on-the-ai-exponential

Anthropic Frontier Red Team. (10.9.2026). *Measuring tactical intelligence targeting and conventional weapons capabilities of AI models* [Forskningsartikel]. Anthropic. Hämtat 16.9.2026 från https://www.anthropic.com/research/intelligence-targeting-conventional-weapons-capabilities

CBS News. (28.2.2026). *Full interview: Anthropic CEO responds to Trump order, Pentagon clash* [Videoklipp]. Hämtat 16.9.2026 från https://www.youtube.com/watch?v=MPTNHrq_4LU

Fagerström, N. (2026, 13.9). Därför varnar tunga AI-namn för att mänskligheten kan utplånas. *Svenska Yle.* Hämtat 16.9.2026 från https://yle.fi/a/7-10105189

Google DeepMind. (30.7.2026). *Gemini Robotics 2 brings whole body intelligence to robots* [Videoklipp]. Hämtat 16.9.2026 från https://www.youtube.com/watch?v=4lSQnrMC6nY

Korinek, A., Jones, C. I., Sacher, S., Cotter, T. & McCrory, P. (2026). *Economic Scenarios for Transformative AI* (Working Paper No. 2026-02). The Anthropic Institute. Hämtat 16.9.2026 från https://www-cdn.anthropic.com/files/4zrzovbb/website/cf58f84d46a4a76bf5a5b039ac695fba6b80041c.pdf

Louniala, O.-P. (2026, 8.9). Open AI väittää ratkaisseensa yhden vuosituhannen matemaattisista mysteereistä. *Yle.* Hämtat 16.9.2026 från https://yle.fi/a/74-20245250

OpenAI. (3.9.2026). *Introducing GPT-6 Astra: the most intelligent and aligned model in the world* [Videoklipp]. Hämtat 16.9.2026 från https://www.youtube.com/watch?v=1QNsdr-Qx_I

Paananen, V. (2026, 4.9). Suomessakin pitää herätä nyt tekoälyn riskeihin, sanoo ulkoministeri Elina Valtonen. *Helsingin Sanomat.* Hämtat 16.9.2026 från https://www.hs.fi/politiikka/art-2000012236890.html

Palosaari, A. (2026, 15.9). Pörssit heiluvat - Professori Vesa Puttonen kertoo, mitä sijoittajan kannattaa nyt muistaa. *Yle.* Hämtat 16.9.2026 från https://yle.fi/a/74-20246317

Recursive self-improvement. (4.8.2026). *Wikipedia.* Hämtat 16.9.2026 från https://en.wikipedia.org/wiki/Recursive_self-improvement

Ruiz, S., Korinek, A., Jones, C. I., Sacher, S., Cotter, T. & McCrory, P. (2026). *Economic Scenarios for Transformative AI* [Blogginlägg]. The Anthropic Institute. Hämtat 16.9.2026 från https://www.anthropic.com/institute/econ-scenarios

Trump, D. J. [@realDonaldTrump]. (14.9.2026). *Concerning AI, when, in the History of Business, did anyone see the Leaders of an Industry call for Regulation…* [Inlägg]. Truth Social. Hämtat 16.9.2026 från https://truthsocial.com/@realDonaldTrump/115270591511950591

Varho, E. (2026, 15.9). Analyysi: Tekoälypomot varoittavat nyt vaaroista, mutta sääntely törmää vuoreen nimeltä Donald Trump. *Yle.* Hämtat 16.9.2026 från https://yle.fi/a/74-20246438

Veritasium. (25.2.2026). *The Internet Was Weeks Away From Disaster And No One Knew* [Videoklipp]. Hämtat 16.9.2026 från https://www.youtube.com/watch?v=a0ag03mSuXQ

Yudkowsky, E. & Soares, N. (2025). *If Anyone Builds It, Everyone Dies: The Case Against Superintelligent AI.* London: The Bodley Head. Hämtat 16.9.2026 från https://en.wikipedia.org/wiki/If_Anyone_Builds_It,_Everyone_Dies

Anmärkning om AI-användning: Artificiell intelligens har använts vid dokumentets grafiska utformning, för att granska textens grammatik samt för att kontrollera att källhänvisningarna följer APA-mallen. Innehåll och slutsatser är skribentens egna.