ELEKTRONIKTILLVERKNING · AI-INSPEKTION × ROBOTIK

AI på SMT-linjen:

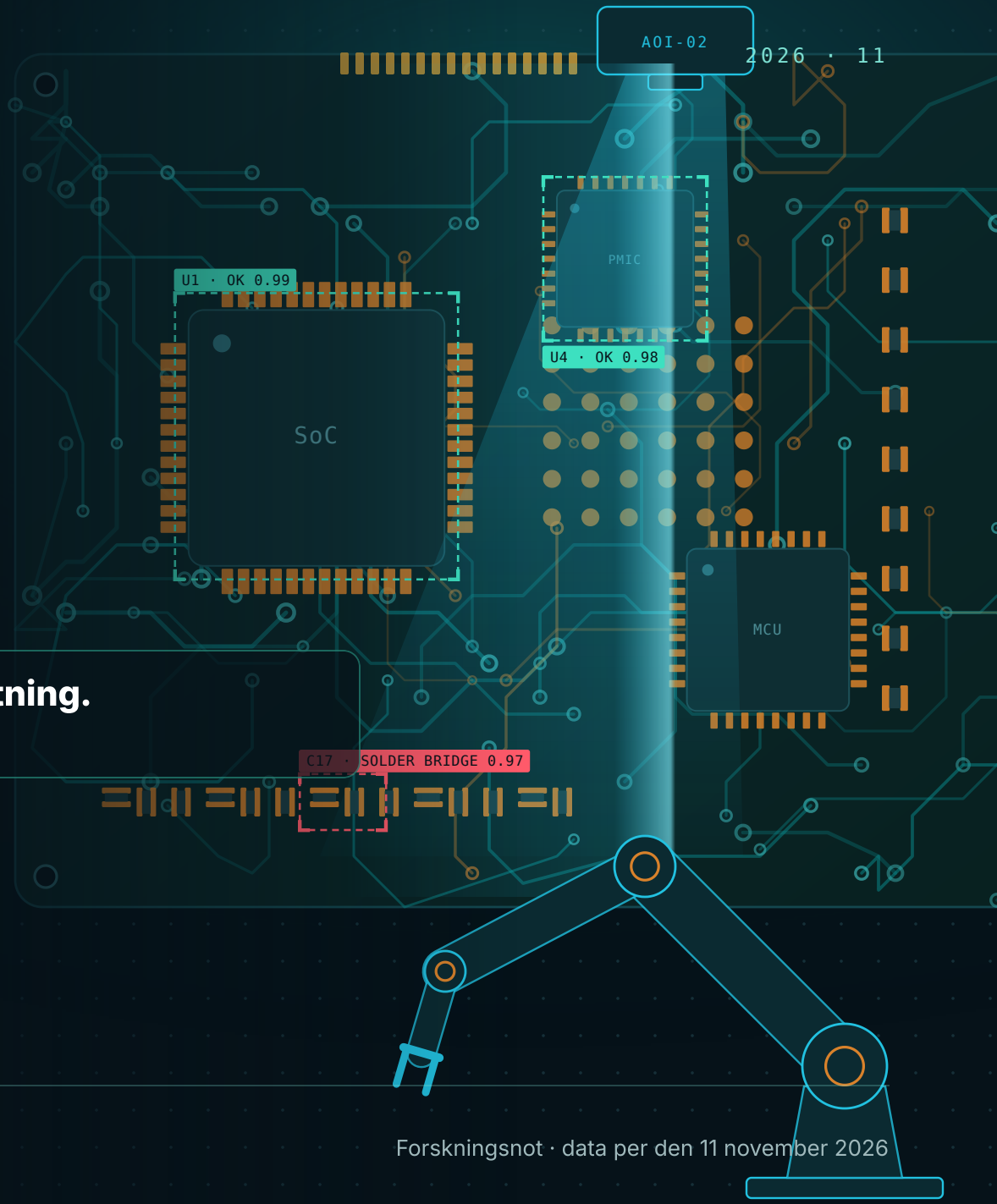
vad som betalar sig 2026

AI-visuell inspektion och robotarmar i elektronikfabriker

Fastlandskina · Taiwan · Sydostasien · USA

HUVUDFYND

AI-granskning är värd att göra, i begränsad omfattning. Förkroppsligade armar kan vänta.



Tan Shuai
tanshuai.comi@tanshuai.com

Forskningsnot · data per den 11 november 2026

Värt att göra, men smalt, nu och med partners

HUVUDFYND

Det som är moget 2026, och det fabriker betalar för, är AI-granskning och nischinspektion på AOI-bilder plus programmerad robot-maskinbetjäning, inte »en stor modell som tittar på kort plus förkroppsligade robotar som plockar dem«.

0.46–0.65 yr

Återbetalningstid för AI-granskning i USA, Taiwan och Shenzhen, vid RMB 50k per linje

≥95% · <20 PPM

Ett ledande EMS-företags interna AI-granskning: ribban för att sälja till tier-1-EMS

100–700 RMB M

Adresserbar tredjeparts-eftermonteringsmarknad i Fastlandskina, engångs, ≈5-årigt fönster

0 scores

Benchmarkpoäng för industridefekter för november 2026 års flaggskepp

1.1–1.5 yr

Återbetalningstid för en cobot-cell för maskinbetjäning, två skift

0.003–0.14 FTE

Arbete en utplockningsrobotarm sparar per skift: den betalar sig aldrig tillbaka

Källa: rapportens sammanfattning och kapitel 4–5; återbetalnings- och marknadssiffror är författarens uppskattningar.

Sex frågor, sex svar

① Vad fabriker kör

3D SPI, AOI före och efter reflow, stickprovsbaserad röntgen, ICT/FCT. En person granskar bakom varje AOI; takt ≈13–22 s per kort.

② Hur långt tekniken når

Inspektion är industrikvalitet som detektor + AI-granskning + några få personer; inte noll personer till något pris. Fastlandet: inhemska modeller; Taiwan: kunden avgör; USA: tullar och certifiering.

③ Lokalt och luftgapsisolerat

Praktiskt; begränsningarna är processuella. En fläktlös box ≤100 W vid linjen, servrar i serverrummet, 8-bit som standard.

④ Lärare och elev

Inspektion: frontiermodeller förmärker, människor avgör, ett guldset vaktar. Armar: lärarna är fortfarande mänskliga demos och simulering.

⑤ Kostnad och återbetalningstid

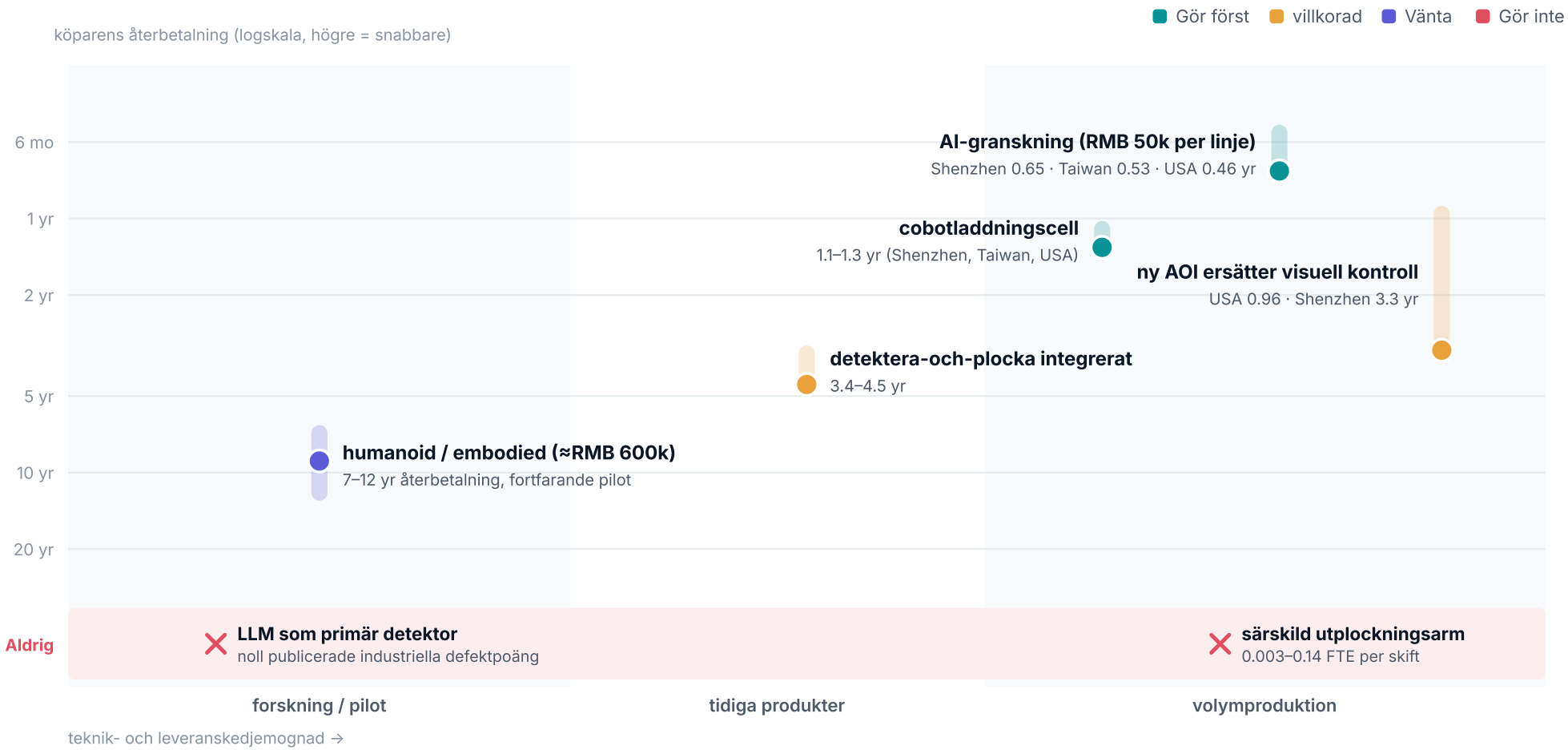
AI-granskning inom ett år i USA, Taiwan och Shenzhen; cobot-betjäning 1.1–1.5 år; Sydostasien 2–6 år; en utplockningsarm aldrig.

⑥ Nu eller vänta

Reell efterfrågan, trångt utbud; OEM-AI lämnar ett ≈5-årigt eftermonteringsfönster. Inspektion nu, smalt; förkroppsligad AI efter data från 2027–28.

Källa: rapportens sammanfattning.

Var AI passar på fabriksgolvet idag

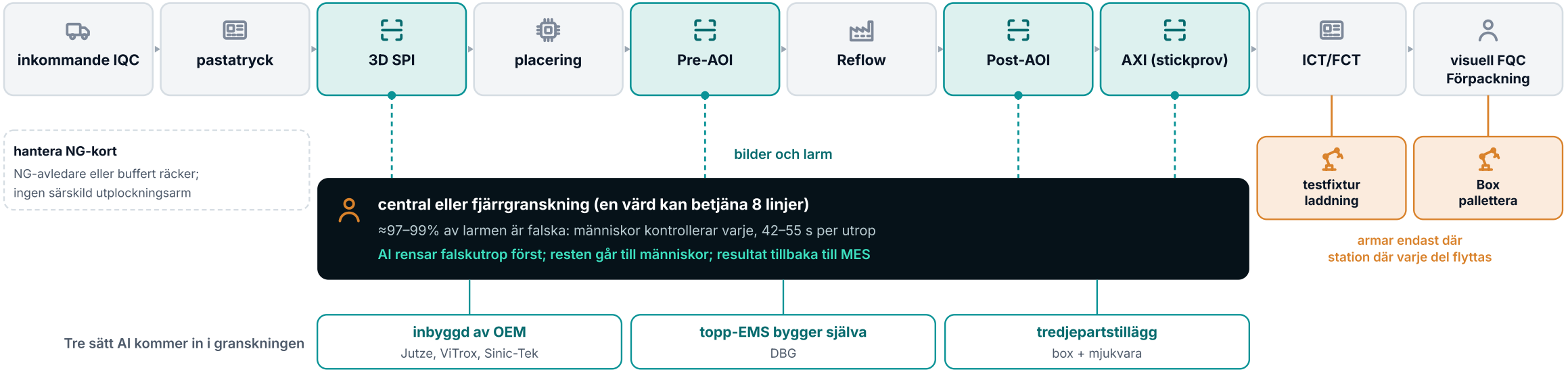


Källa: Återbetalning från författarens modell (kapitel 4); mognadsaxeln är författarens bedömning.

Människorna sitter bakom AOI:n. Det är där AI kommer in

En volym-SMT-linje

takt ≈13–22 s per kort · 3–5 AOI per linje



Källa: Offentliga årsredovisningar och leverantörsmaterial (Huaqin, DBG, Sinic-Tek med flera); sammanställt av författaren.

Omkring 97–99% av AOI-larmen är falskutrop

per 100 AOI-larm



Idag: människor kontrollerar varje utrop

42–55 s varje

≈448 falskutrop per skift på en hektisk linje

→ 5.2–6.8 h per skift ≈ 1 heltidsgranskare

fallbemanning: 2 per skift

AI-granskning: rensa falskutrop, ge resten till människor

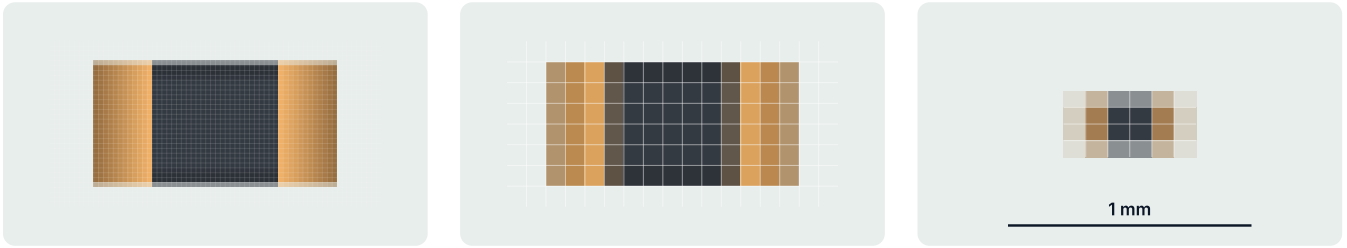
falskutrop eller granskningstimmar minskade 60–90% (leverantörsuppgifter)

topp-EMS eget riktmärke: täcker ≥95% av flaggade komponenter, escapes <20 PPM, mänsklig granskning ned ≥85%

vid fel till människor, linjen rullar vidare

Källa: Siemens fallstudie, iFactory (leverantörsuppgift), DBG:s årsredovisning 2025; granskningstimmar uppskattade av författaren.

En generell kamera är ingen AOI



AOI: 8–20 µm per pixel
0402 ≈ 100×50 px

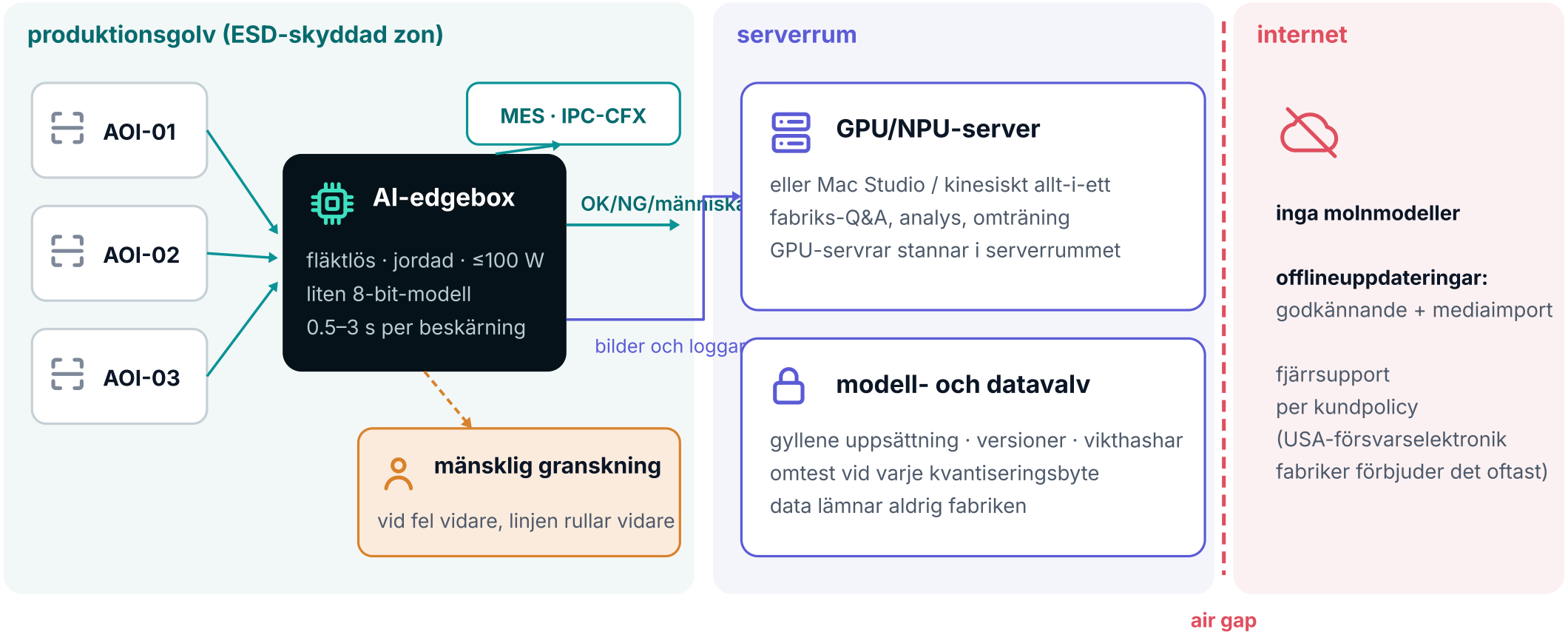
4K-kamera över en 300 mm-panel: ≈78 µm/px
0402 krymper till ≈13×6 px: räcker för närvaro och grov förskjutning

samma kamera på en 01005
≈5×3 px: i praktiken osynlig

- **AOI:** 8–20 µm per pixel, flervinklig flerfärgad belysning och strukturerat ljus
- **4K över en 300 mm-panel:** ≈78 µm per pixel; en 0402 är ≈13×6 pixlar, tillräckligt för närvaro, inte för lödfogar
- **01005:** ≈5×3 pixlar, i praktiken osynlig
- **Så:** AI:n arbetar på AOI:ns egna bilder; den ersätter inte optiken

Källa: Circuits Assembly; författarens beräkning för en 300 mm-panel och en 4K-sensor. Pixelrutnäten är illustrativa.

En luftgapsisolerad installation: boxen avgör, serverrummet svarar



Källa: Författarens syntes (kapitel 2).

Renrums- och ESD-zoner avgör var beräkningskraften står; hastighetsgränserna beror på användningen

renhetskrav →

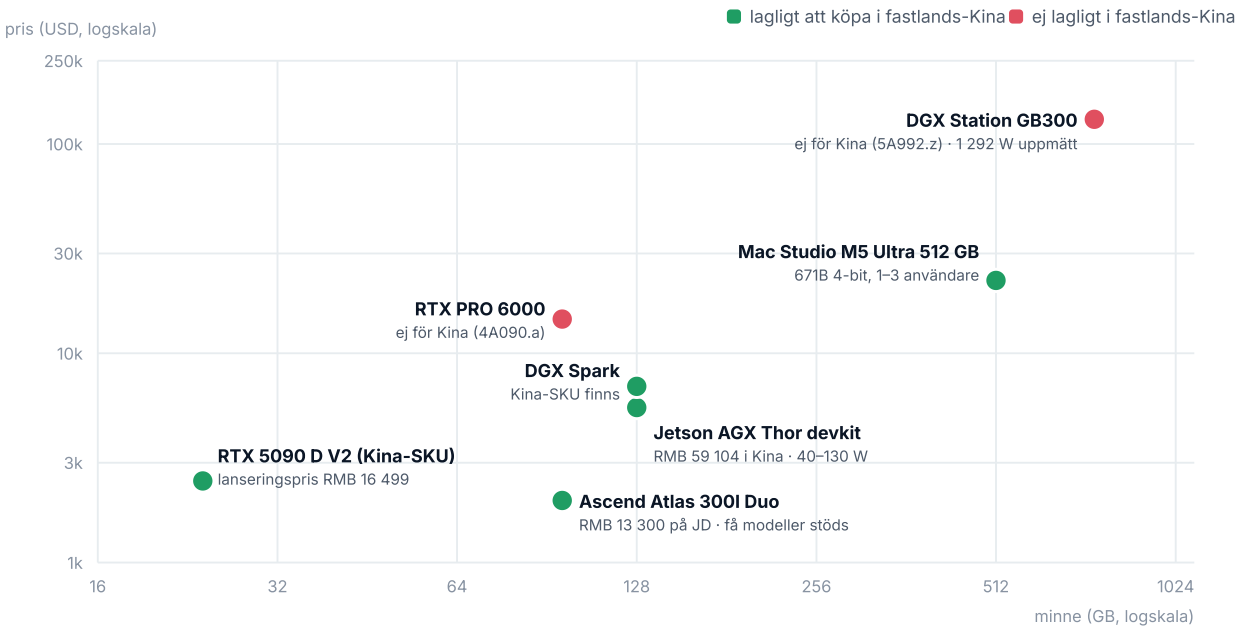
vanlig SMT/PCBA-hall	renrum ISO 7–8	mini-miljö ISO 5–6	fab, ISO 5 eller renare
ESD-skyddad zon (≈ISO 8) RH 40–60%	medicin, kameramoduler, EV renrum + ESD	bar die, wire bond, optik mini-miljö	SEMI E78 ESD · E187/E188 cyber fabspecifika planer
line-side beräkning fläktlös, jordad metallbox bredvid AOI:n eller granskningsplatsen	**line-side beräkning** fläktlös, förseglad IP65/66, lätt att torka av; be om ISO 14644-14-data	**line-side beräkning** bara kameror och ljus; beräkningsbox utanför	**line-side beräkning** endast verktygsintegrerade, fabcertifierade styrenheter
GPU/NPU-server serverrum eller klimatiserat IT-rack	**GPU/NPU-server** utanför renrummet, nätverk in	**GPU/NPU-server** serverrum	**GPU/NPU-server** vätskekylda rack i subfab eller datacenter



Nyckelpunkt En vanlig SMT-hall är ett ESD-skyddat område, inte ett klassat renrum: en fläktlös box i ett jordat metallhölje står vid AOI:n eller granskningsplatsen, och GPU-servrar hamnar aldrig på linjen.

Källa: ANSI/ESD S20.20, IEC 61340-5-1, ISO 14644, SEMI-standarder och leverantörsmaterial; sammanställt av författaren. Författarens syntes (kapitel 1–2).

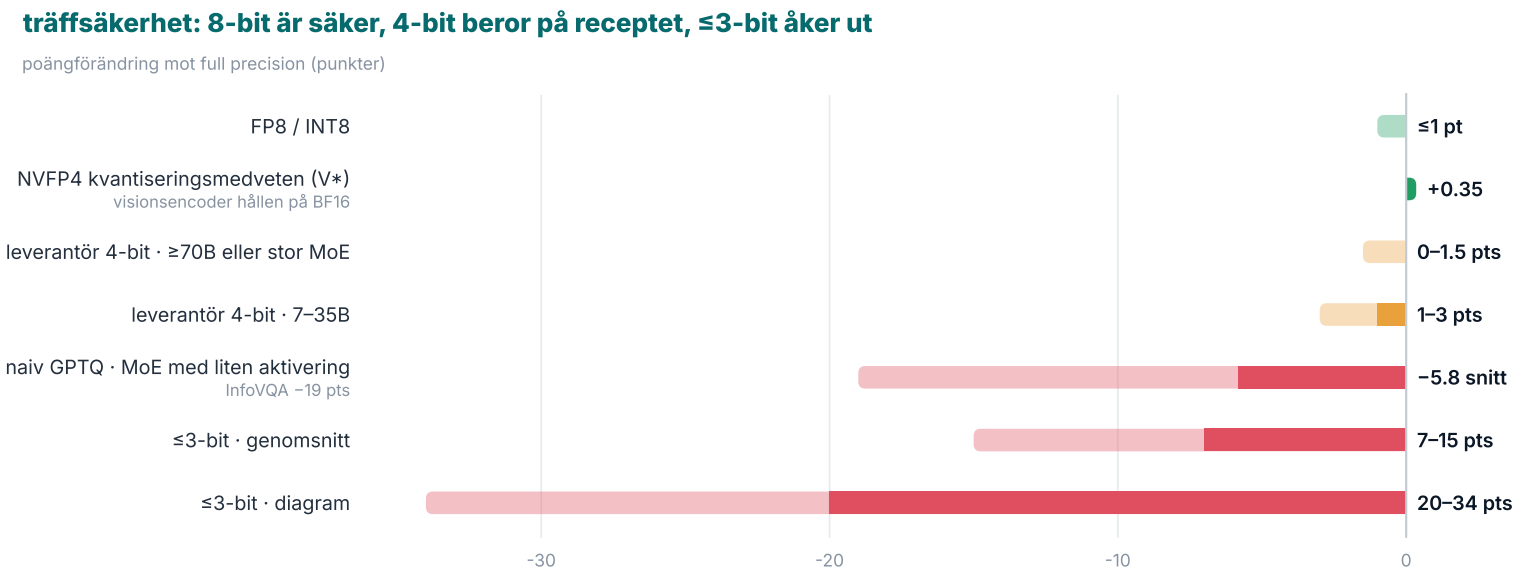
Exportreglerna väljer hårdvarans kortlista innan benchmark gör det

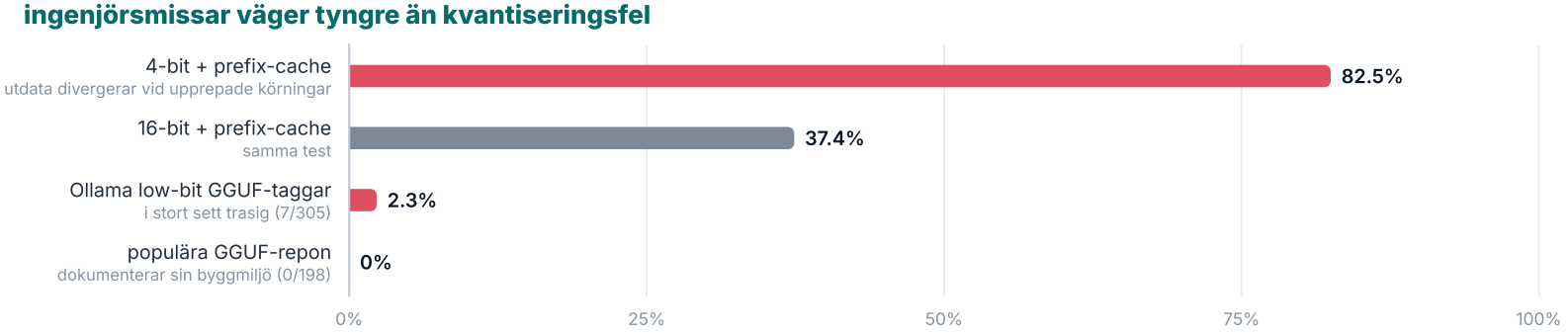


- **Jetson Thor:** 128 GB, 40–130 W; 2–8B-seendemodeller på AOI-utklipp, VLA:er i 3B-klass i realtid; RMB 59,104 i Kina
- **Mac Studio 512 GB:** DeepSeek 671B i 4-bit för 1–3 användare
- **DGX Station:** 5A992.z, inte för fastlandet; en 743B-modell ger 32 användare ≈4 tokens/s var
- **Inhemska acceleratorer:** två Ascend-servrar med åtta kort för fullstor Q&A för flera användare

Källa: NVIDIA:s exportklassificeringar, Apple, RS China, JD.com och recensionspriser (november 2026). StorageReview (aug och okt 2026); ServeTheHome. Oberoende tester från en källa.

Kvantisering: 8-bit är säkert, 4-bit beror på receptet, 3-bit och lägre åker ut

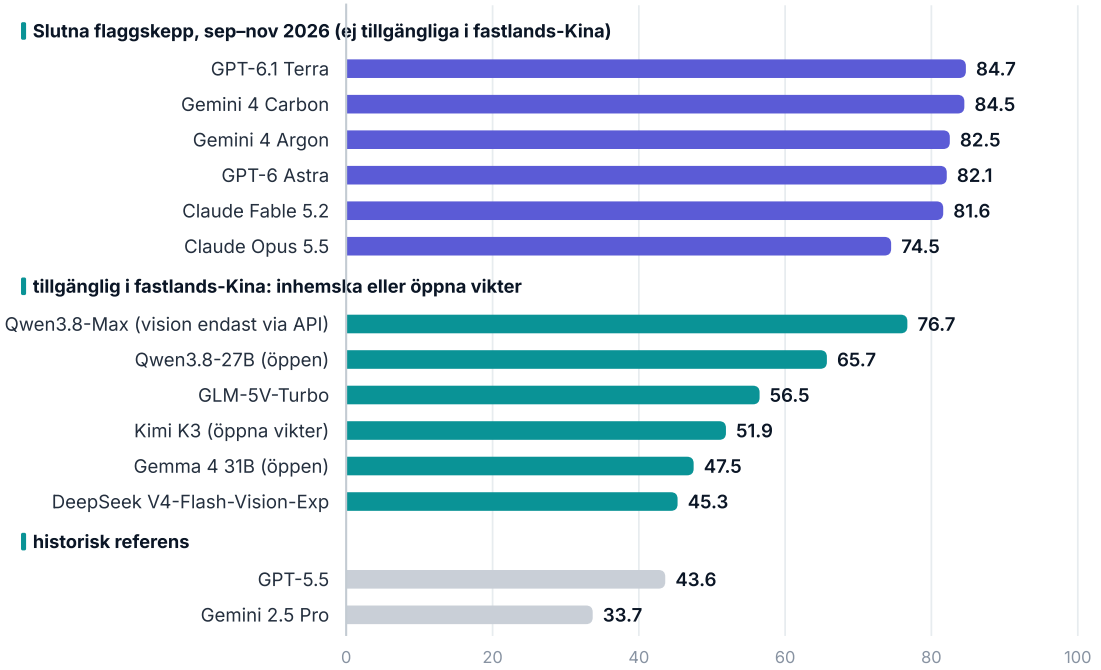




Källa: NVIDIA Nemotron-modellkort, VLMQ, VEQ, arXiv 2609.05881, 2610.19498, 2609.04748 med flera.

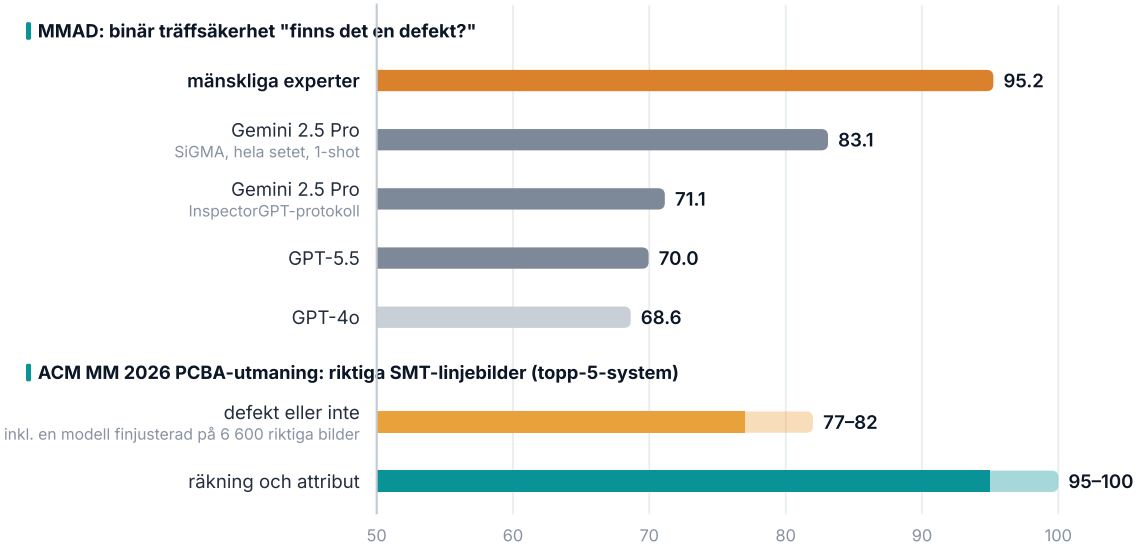
Generiskt seende fördubblades på en generation; industridefektpoäng saknas

Roboflow generisk detektering mAP@50



industriella defektbenchmarks (MMAD, MVTec AD 2, VisA, Real-IAD, UniPCB): noll publicerade poäng för dessa 2026-flaggske
generiska detektionspoäng fördubblades på en generation, men det är ingen defektbedömning; testa på din egen gyllene uppsättning

träffsäkerhet eller poäng (%)



Källa: Roboflow Vision Evals; modellrelease-noter; MMAD, MVTec AD 2 och andra benchmarksidor (per 11 nov 2026). MMAD; SiGMA; InspectorGPT; ACM MM 2026 PCBA-utmaning.

Frontiermodeller förmärker, människor avgör, ett guldset vaktar

inspektion: frontmodellen för-märker, människor avgör, en gyllene uppsättning portar

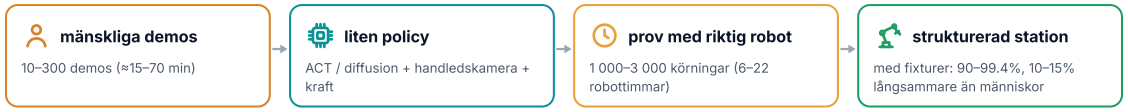


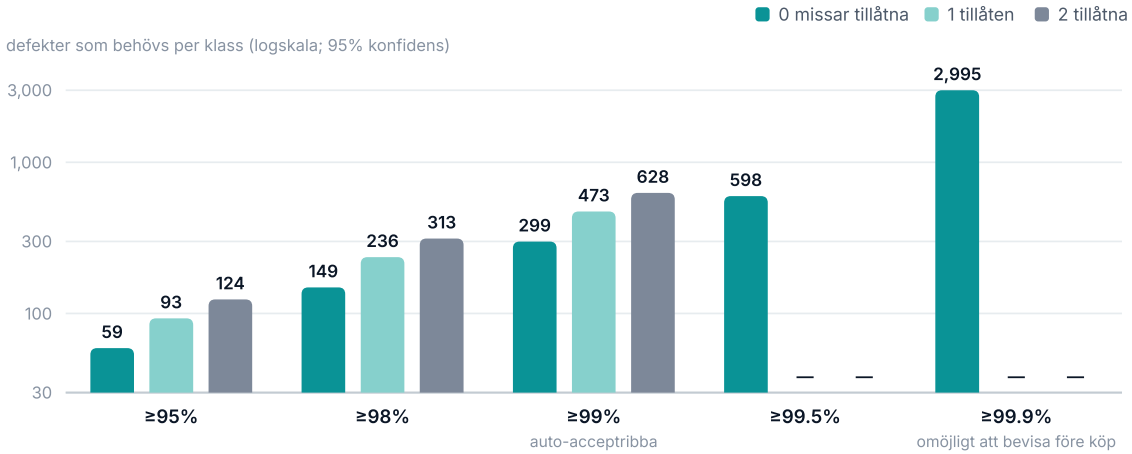
ribba för automatisk acceptans av lärar-etiketter

per klass: ≥299 defekter, noll missar (recall ≥99% vid 95% konfidens), överensstämmelse inte sämre än människa–människa; annars bara för-märkning

armar: lärarna är fortfarande mänskliga demos, RL på riktiga robotar och simulering

Simmodeller hjälper bara offline: programmering och omställningar

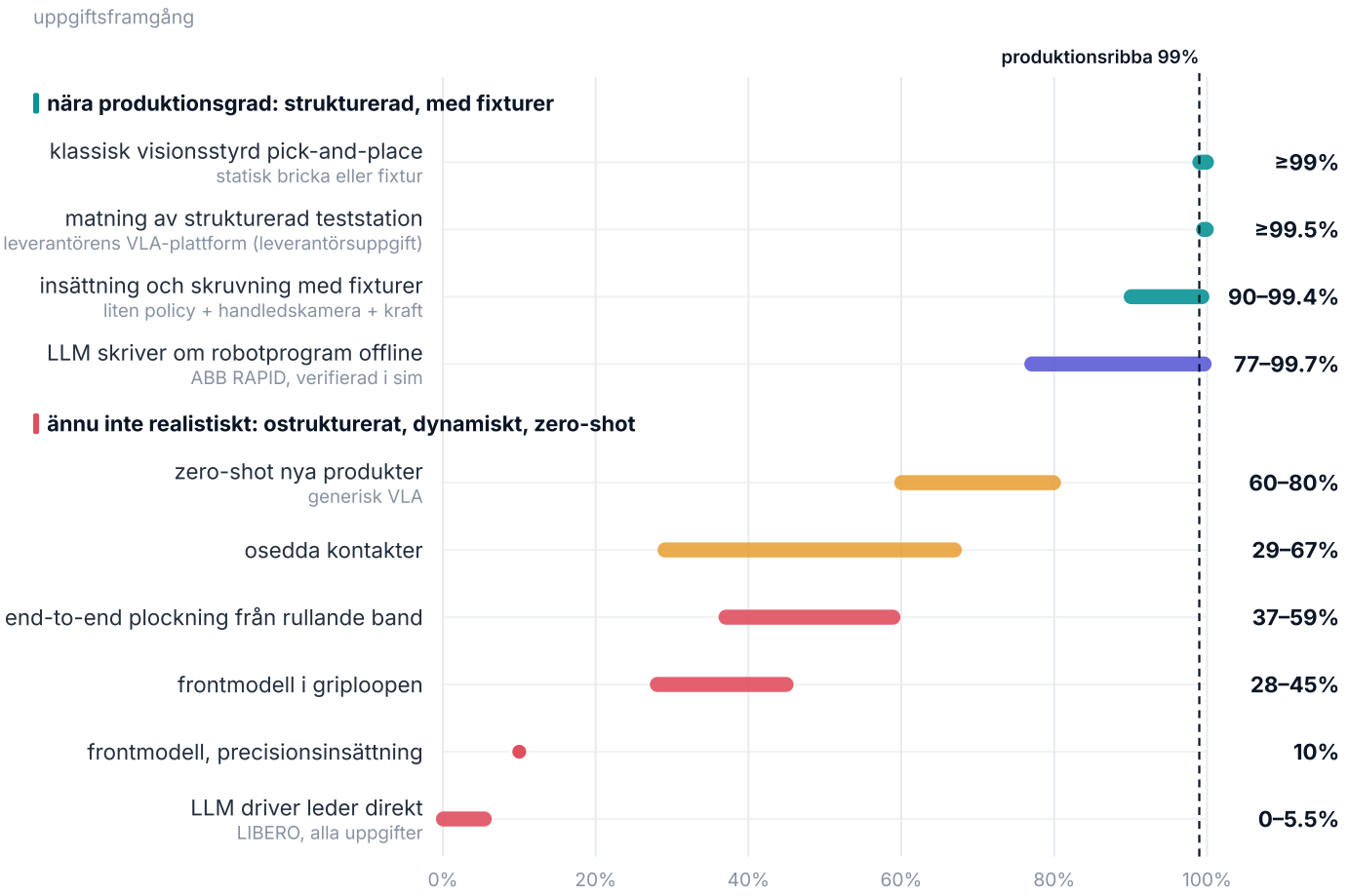


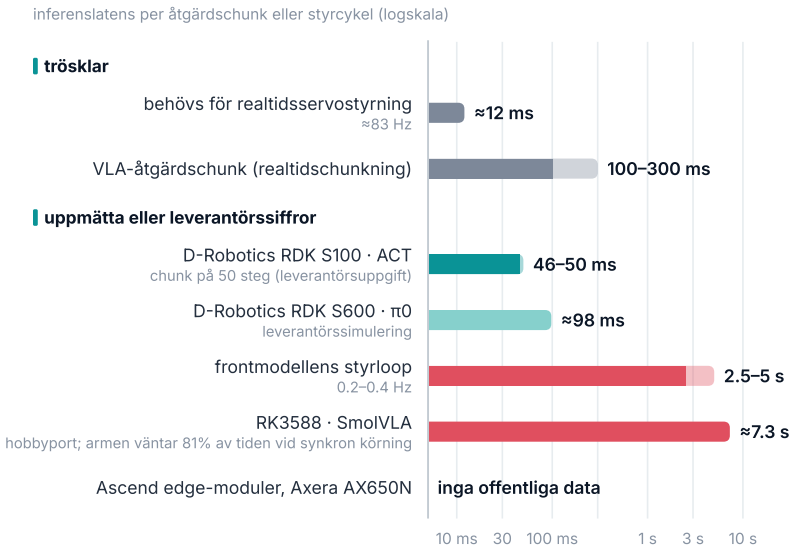


Nyckelpunkt Autoacceptera lärarens etiketter endast för en klass med noll missar på ≥299 guldset-defekter (99% recall vid 95% konfidens); kritiska klasser som kortslutningar, saknade delar och polaritet behöver mer.

Källa: Författarens syntes (kapitel 3); kostnader och stickprovsstorlekar är uppskattningar. Exakta binomiska konfidensintervall (95%).

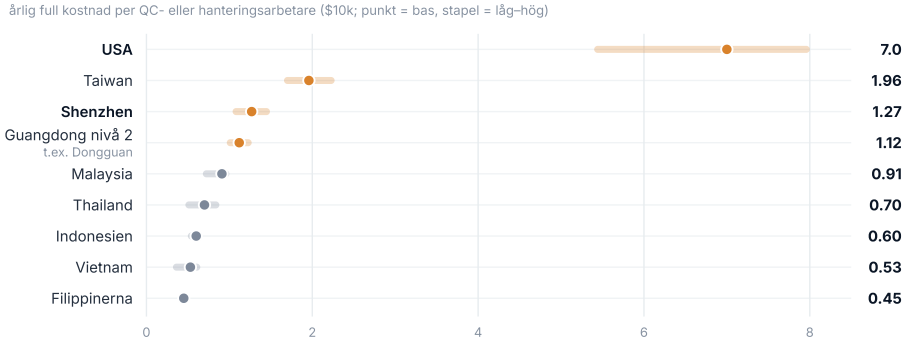
Robotelever: nära produktionsgrad bara när stationen är strukturerad

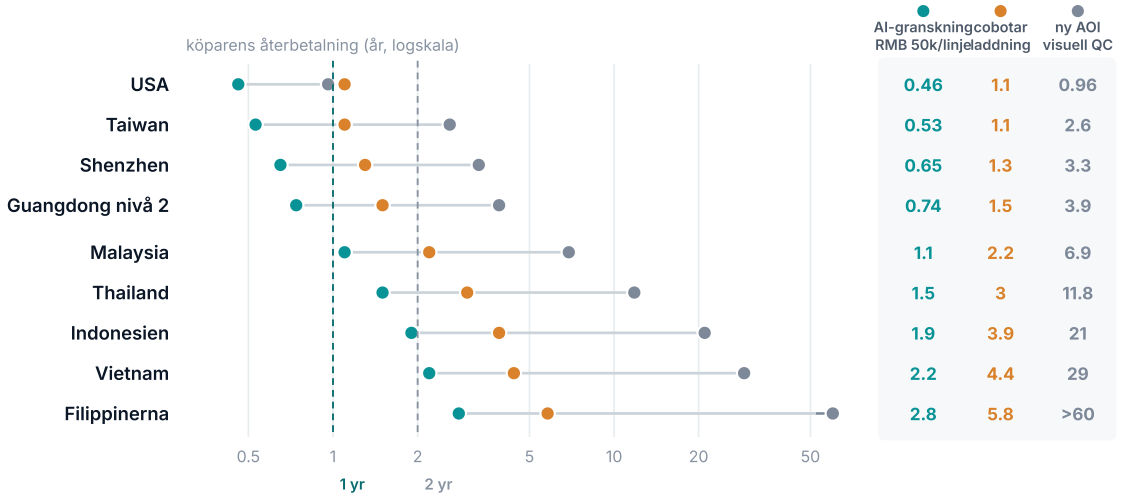




Källa: Artiklar och leverantörsmaterial citerade i kapitel 3; vissa siffror är leverantörsuppgifter. D-Robotics dokumentation och forum (leverantörssiffror); community-portar på GitHub; Anthropic-forskning.

AI-granskning betalar sig på månader; maskinbetjäning på ungefär ett år



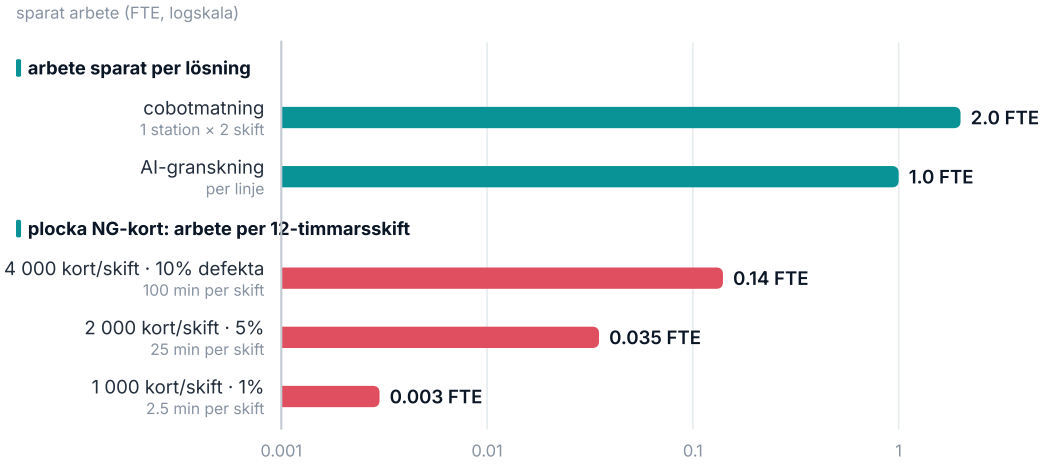


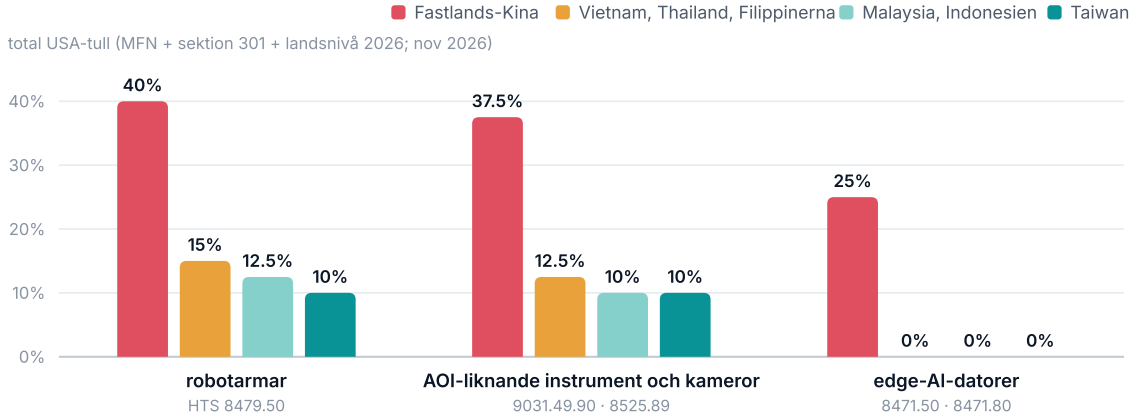
en robotarm enbart för att plocka ut NG-kort betalar sig aldrig, någonstans
vid first-pass yield över 90% tar utplockning bara 0.003–0.14 FTE per skift

Nyckelpunkt Vid RMB 50,000 per linje betalar AI-granskning sig på 0.46 år i USA, 0.53 i Taiwan, 0.65 i Shenzhen och 2.2 i Vietnam; en tvåskifts-cobotbetjäningscell på 1.1–1.5 år [estimates].

Källa: Shenzhens officiella löneriktlinjer, Taiwans arbetsmarknadsdepartement, JETRO, BLS; författarens uppskattningar. Författarens modell: RMB 50k AI-granskningsbox per linje; cobotmatningscell på RMB 200k eller $140k, två skift.

En utplockningsarm betalar sig aldrig tillbaka; var en arm tillverkas ändrar dess USA-tull





Nyckelpunkt Över 90% förstapassutbyte tar utplockning av NG-kort 0.003–0.14 av en person per skift, så en NG-avledare eller buffert gör jobbet; en fastlandstillverkad arm betalar dessutom cirka 40% kombinerad USA-tull.

Källa: Författarens uppskattningar. USITC Harmonized Tariff Schedule kapitel 99; Federal Register; USTR (november 2026).

Reellt men litet: 16 000 adresserbara linjer, ≈RMB 320M engångs, ett femårsfönster

SMT-linjer i Kina
40 000 rader
35k–46k

× 80% × 50%
ingen AI ännu: 75–85%
nåbar: 40–60%

servicebar
16 000 rader
10.5k–23.5k

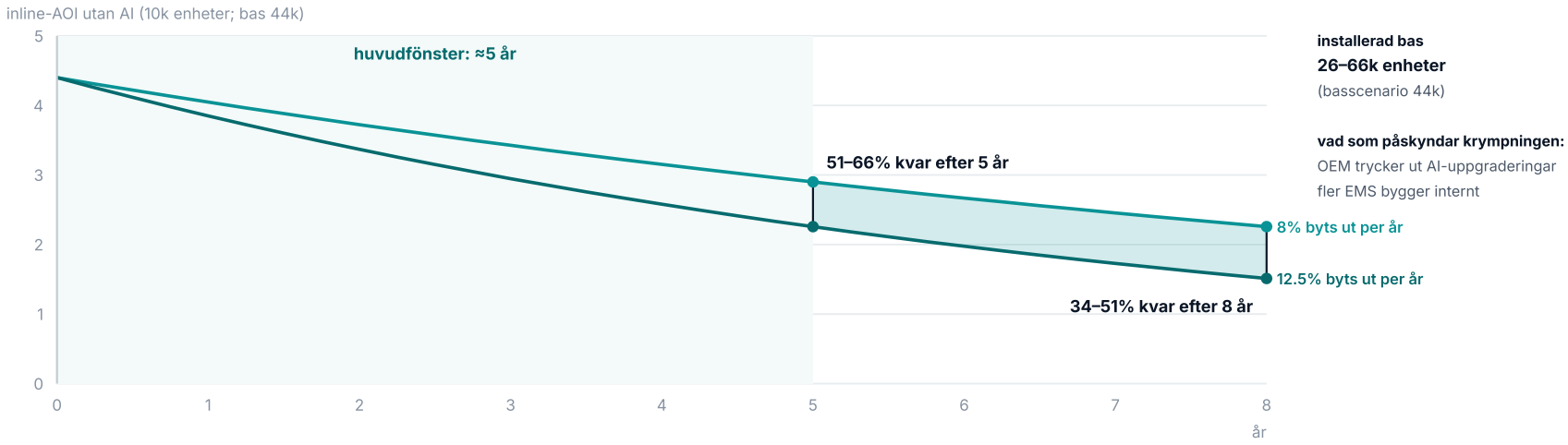
× RMB 20k
RMB 10–30k engångs

engångsmarknad
RMB 320M
RMB 105–705M

÷ ≈5 år
krymper 8–12%/yr

5-årsfönster
RMB 64M/yr
RMB 21–141M

en nykomling som tar 2–5% av servicebara linjer på 3 år: ≈320–800 linjer, ≈RMB 2–5M per år
som prenumeration (RMB 5k per linje och år): RMB 1.6–4M återkommande, bara 0.3–2.3% av IDC:s prognos för industriell AI-QC i Kina



Källa: NVR (bas 2020); författarens uppskattningar. Författarens uppskattningar; installerad bas och andelar inte oberoende verifierade.

Granskningskilen är omringad; marginalen ligger i produktifierad hårdvara

topp-EMS bygger själva

DBG (2025): täcker ≥95% av flaggade komponenter
escapes <20 PPM, mänsklig granskning −85%
detta är ribban för försäljning till tier-1 EMS

AI inbyggd i AOI av OEM

Jutze AI 3D-AOI i produktion, "tar bort mänsklig granskning vid inspektion"
ViTrox V510Ai; Sinic-Tek utvecklar med Tencent
nya maskiner kommer med AI; gamla behöver uppgraderingar

tredjeparts AI-granskning

box + mjukvara skruvad på befintlig AOI

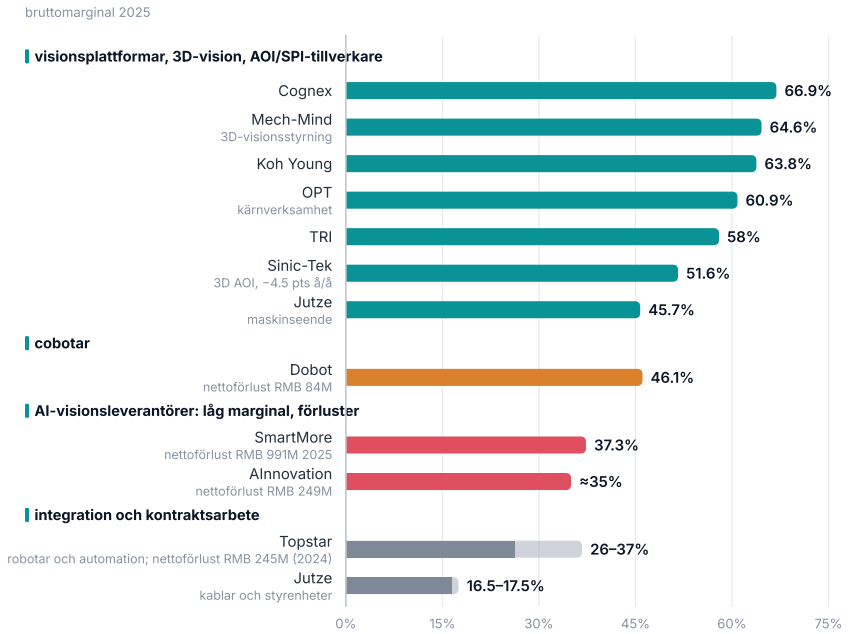
plattformsleverantörer

Hikvision Guanlan och SmartMore IndustryGPT säljs
ledande AI-leverantör: intäkter 2025 RMB 1,086B
nettoförlust RMB 991M: priskrig

vad som återstår för tredje part

mellanstora EMS (≈10–40 linjer) med äldre eller blandad AOI och ingen OEM-AI-uppgradering

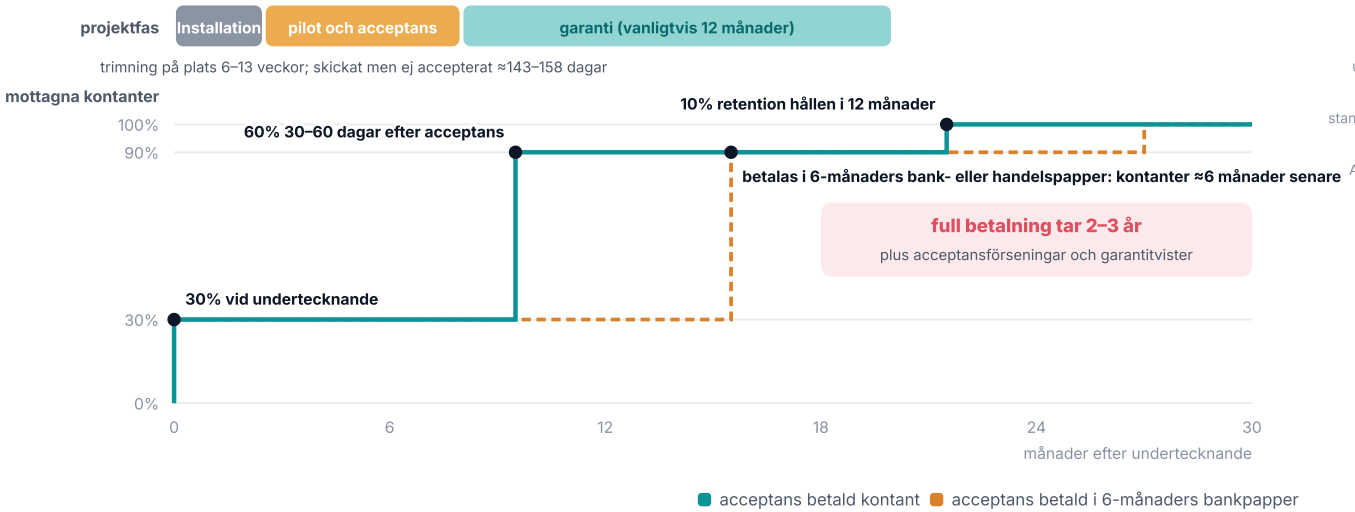
Nischer AOI täcker dåligt: THT och handlödning, konformbeläggning, kablar/kontakter/FPC, slutmontering och skruvar

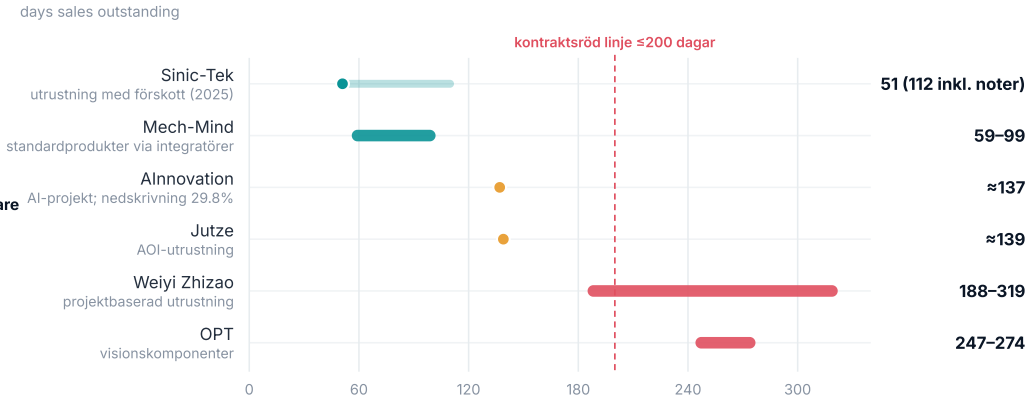


Bevis Den ledande AI-seende-leverantören hade 2025-intäkter på RMB 1.086 miljarder och en nettoförlust på RMB 991 miljoner; det närmaste börsnoterade jämförelsebolagets 3C-intäkter föll 79.9% år över år i januari–maj 2026.

Källa: Jutze, Sinic-Tek och DBG:s årsredovisningar; Hikvision och SmartMore publikt material; sammanställt av författaren. Årsredovisningar eller prospekt 2025.

Köpare betalar 30/60/10; full betalning tar två till tre år





Nyckelpunkt 30% kommer med beställningen, 60% ungefär fyra till nio månader senare och de sista 10% ungefär 16–21 månader senare; från ett första pilotprojekt hos ett stort EMS till full betalning kan det ta två till tre år [estimate].

Källa: Sinic-Tek:s svar på förfrågan, Mech-Mind, GKG, JT och andra handlingar; schema på sedvanliga villkor, tidpunkt uppskattad. Företagshandlingar; fordringar (inkl. kontraktstillgångar) ÷ omsättning × 365.

Planera offline, agera med små policyer, håll rörelse och säkerhet deterministiska

Lager	vad den gör och hur långt bevisen räcker	tidsskala · vem äger det
Planering L4	**frontmodeller (offline)** skriva eller ändra robotprogram offline, omställningar, förklara fel skriver om ABB RAPID-program till 77–99.7%, utrullade efter simulering styrloopen bara 0.2–0.4 Hz: håller sig utanför realtid	**sekunder till minuter** modelljättar
färdigheter / policy L3	**ACT, diffusionspolicies, små VLA** strukturerade stationer med fixturer 90–99.4% edge-chips: Jetson Thor, D-Robotics RDK, Ascend, Rockchip	**åtgärdschunkar 100–300 ms** chiptillverkare · robottillverkare
rörelse-styrning L2	**robotstyrenhet** banor, servo, kraftreglering, kollisionskontroller deterministisk kod; inlärda policies sätter bara mål	**millisekunder (realtid kräver >83 Hz)** robottillverkare
Säkerhet L1	**säkerhets-PLC, skannrar, kraft- och hastighetsgränser** certifierad enligt ISO 10218-1/-2:2025 och lokala regler ingen ML i säkerhetsfunktioner: EU-bedömning av anmält organ från 20 jan 2027	**hård realtid** certifierad robot + integratör

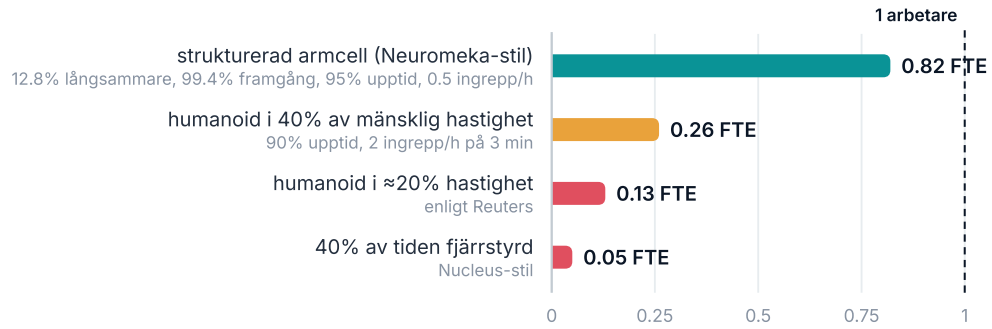
små företag vinner i integration och sensorhårdvara: sätt inlärd perception och små policies i certifierade robotceller, station för station

	framgångsgrad	takt	ingrepp / fjärrstyrning	upptid / fel	volymkörning
Figure 02 @ BMW	endast mål >99%/skift	endast mål 84 s	endast mål 0/skift	u/a	1 250 h
Agility Digit @ GXO	u/a	u/a	u/a	u/a	100 000 backar
Apptronik Apollo @ Mercedes	u/a	u/a	u/a	u/a	u/a
Agile ONE @ Audi	u/a	u/a	u/a	u/a	u/a
Nucleus (ej namngiven fabrik)	u/a	u/a	≈40% fjärrstyrd*	u/a	≈2 h video
AgiBot G2 @ Longcheer	99.99%	18–20 s per steg	u/a	ett fel, 2 min	64+ h
hybridcell från Neuromeka	322/324	12.8% långsammare	u/a	u/a	324 cykler
π0.6 @ kundanläggning	—	—	96.4% autonom	u/a	1 skift

* metod ej offentliggjord

offentliggjort (mest leverantörsuppgifter) · endast mål · u/a

netto ersatt arbete per robot och skift



Källa: författarens syntes (kapitel 3, 6) och uppskattningar; ISO 10218:2025; EU:s maskinförordning; offentligt material från BMW, Figure, Agility, Mercedes-Benz och andra.

Bygg smalt, bevisa det på fabrikens egna bilder, skala sedan

Bygg	AI-granskning och defektklassificering på AOI-bilder för medelstora EMS med äldre AOI eller blandade märken; inspektion som AOI:n täcker dåligt; cobot-celler där varje del hanteras
Bygg inte	En generell AI-plus-arm-plattform; en LLM som primär inspektör; en utplockningsarm på en SMT-linje; humanoider och robots-as-a-service; frontiermodeller i styrkretsar eller säkerhetsfunktioner

Pilotprojekt

1. **Vecka 0–2:** dataavtal, defektklasser, baslinje
2. **Vecka 1–4:** guldset, märkt blint och avgjort
3. **Vecka 2–4+:** skuggläge, jämfört per klass
4. **Vecka 4–8:** assisterad drift endast på godkända klasser
5. **Sedan:** skriftligt godkännande, skala linje för linje, omvalidera vid varje förändring

Säkerhet Ingen maskininlärning i säkerhetsfunktioner: EU kräver bedömning av anmält organ från och med den 20 januari 2027, och AI-certifieringsvägar kommer inte att fastställas före 2027–2029.

Källa: rapportens kapitel 6–7 och sammanfattning; ISO, IEC, UL, EUR-Lex, SAC och nationella regleringsdatabaser.

OM FÖRFATTAREN

Tan Shuai

Teknolog och grundare, med mer än tjugo års arbete inom chip, hårdvara och systemmjukvara samt 20 beviljade patent. Hans arbete har fört honom genom Silicon Valley, New York, Chicago, Taiwan, Greater Bay Area (Shenzhen och Hongkong), Vietnam och Malaysia, med ett långt fokus på avancerad tillverkning och leveranskedjor. Idag arbetar han som AI-forskare och -utvecklare, med fokus på hur AI används på fabriksgolvet.

Den fullständiga rapporten (alla figurer, tabeller och källor) och artikeln finns på tanshuai.com. Han välkomnar samtal med personer som arbetar med AI inom tillverkning: i@tanshuai.com

