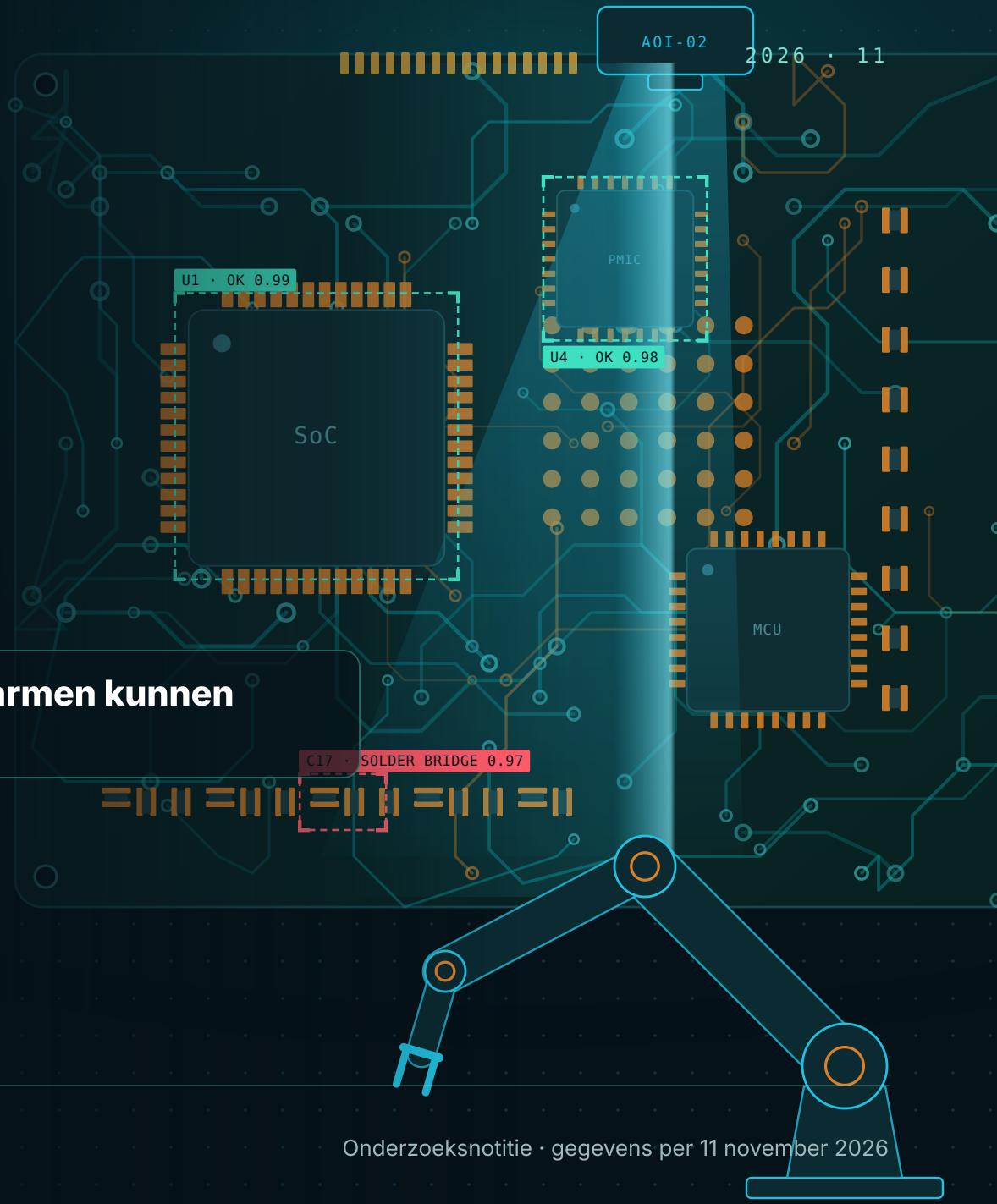


ELEKTRONICAPRODUCTIE · AI-INSPECTIE × ROBOTICA

AI op de SMT-lijn:

wat zich in 2026 terugverdient

AI-visuele inspectie en robotarmen in elektronicafabrieken

Vasteland-China · Taiwan · Zuidoost-Azië · Verenigde Staten

KERNBEVINDING

AI-review loont, in beperkte zin. Belichaamde armen kunnen wachten.

Tan Shuai
tanshuai.comi@tanshuai.com

Onderzoeksnotitie · gegevens per 11 november 2026

De moeite waard, maar smal, nu en met partners

KERNBEVINDING

Wat in 2026 volwassen is, en waarvoor fabrieken zullen betalen, is AI-review en niche-inspectie op AOI-beelden plus geprogrammeerde robot-machinebediening, niet "een groot model dat naar platen kijkt plus belichaamde robots die ze pakken".

0.46–0.65 yr

Terugverdientijd van AI-review in de VS, Taiwan en Shenzhen, bij RMB 50k per lijn

≥95% · <20 PPM

De interne AI-review van een top-EMS: de lat voor verkoop aan tier-1-EMS

100–700 RMB M

Adresseerbare derden-retrofitmarkt in vasteland-China, eenmalig, venster van ≈5 jaar

0 scores

Benchmark-scores voor industriedefecten van de vlaggenschepen van november 2026

1.1–1.5 yr

Terugverdientijd van een cobot-machinebedieningscel, twee ploegen

0.003–0.14 FTE

Arbeid die een uithaal-robotarm per ploeg bespaart: verdient zich nooit terug

Bron: managementsamenvatting en hoofdstukken 4–5 van het rapport; terugverdientijd- en marktcijfers zijn schattingen van de auteur.

Zes vragen, zes antwoorden

① Wat fabrieken draaien

3D SPI, AOI's voor en na reflow, steekproefsgewijze röntgeninspectie, ICT/FCT. Achter elke AOI beoordeelt een persoon; takt ≈13–22 s per plaat.

② Hoe ver de techniek reikt

Inspectie is industrieklasse als detector + AI-review + een paar mensen; tegen geen enkele prijs nul mensen. Vasteland: binnenlandse modellen; Taiwan: de klant beslist; VS: invoerrechten en certificering.

③ On-premises en luchtgescheiden

Praktisch; de grenzen zijn procedureel. Een ventilatorloze box ≤100 W aan de lijn, servers in de serverruimte, standaard 8-bit.

④ Leraar en leerling

Inspectie: frontier-modellen pre-labelen, mensen beoordelen, een golden set fungeert als poort. Armen: de leraren zijn nog altijd menselijke demo's en simulatie.

⑤ Kosten en terugverdientijd

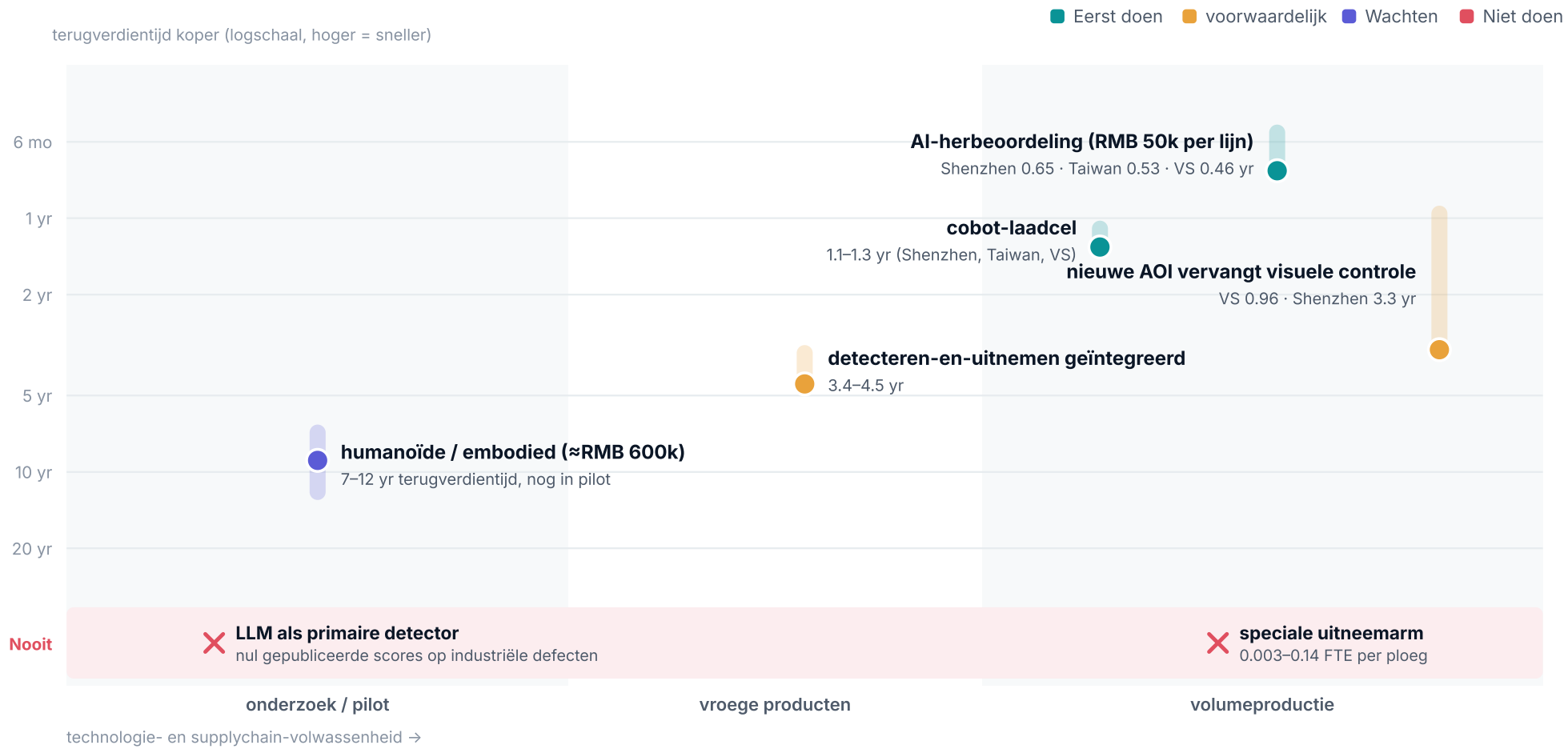
AI-review binnen een jaar in de VS, Taiwan en Shenzhen; cobotbediening 1.1–1.5 jaar; Zuidoost-Azië 2–6 jaar; een uithaalarm nooit.

⑥ Nu of wachten

Reële vraag, overvol aanbod; OEM-AI laat een retrofitvenster van ≈5 jaar over. Nu inspectie, smal; belichaamde AI na de data van 2027–28.

Bron: managementsamenvatting van het rapport.

Waar AI vandaag past op de fabrieksvloer



Bron: Terugverdientijd uit het model van de auteur (hoofdstuk 4); de volwassenheidsas is het oordeel van de auteur.

De mensen zitten achter de AOI. Daar komt AI binnen

Een volume-SMT-lijn

takt ≈13–22 s per board · 3–5 AOI's per lijn



Bron: Openbare jaarverslagen en leveranciersmateriaal (Huaqin, DBG, Sinic-Tek en anderen); samengesteld door de auteur.

Ongeveer 97–99% van de AOI-alarmen zijn false calls

per 100 AOI-alarmen



Vandaag: mensen controleren elke melding

42–55 s per keer

≈448 false calls per ploeg op één drukke lijn

→ 5.2–6.8 h per ploeg ≈ 1 fulltime reviewer

casus-bezetting: 2 per ploeg

AI-herbeoordeling: ruim false calls op, geef de rest aan mensen

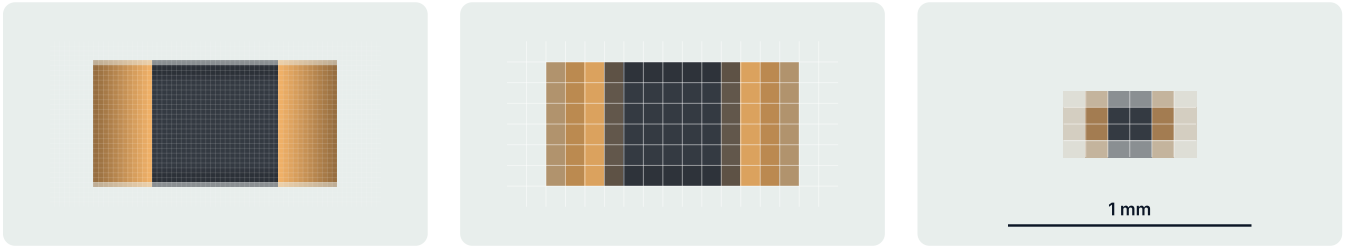
false calls of herbeoordelingsuren 60–90% omlaag (claims van leveranciers)

top-EMS in-house benchmark: dekt ≥95% van gemarkeerde onderdelen, escapes <20 PPM, menselijke herbeoordeling ≥85% omlaag

bij falen naar mensen, lijn blijft draaien

Bron: Siemens-casestudy, iFactory (claim van leverancier), DBG-jaarverslag 2025; herbeoordelingsuren geschat door de auteur.

Een universele camera is geen AOI



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

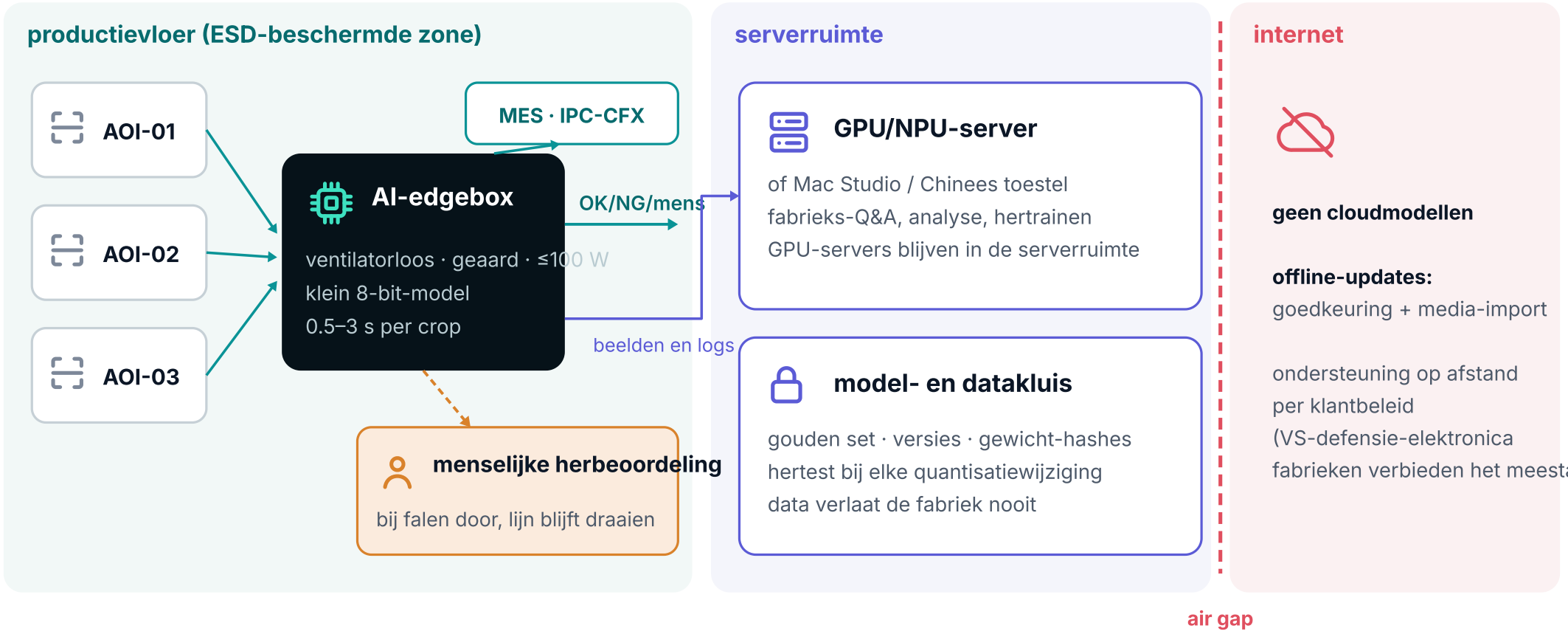
4K-camera over een paneel van 300 mm: ≈78 µm/px
0402 krimpt tot ≈13×6 px: genoeg voor aanwezigheid en grove offset

zelfde camera op een 01005
≈5×3 px: effectief onzichtbaar

- **AOI:** 8–20 µm per pixel, meerhoekige meerkleurige belichting en gestructureerd licht
- **4K over een paneel van 300 mm:** ≈78 µm per pixel; een 0402 is ≈13×6 pixels, genoeg voor aanwezigheid, niet voor soldeerverbindingen
- **01005:** ≈5×3 pixels, in feite onzichtbaar
- **Dus:** de AI werkt op de eigen beelden van de AOI; het vervangt de optiek niet

Bron: Circuits Assembly; berekening van de auteur voor een paneel van 300 mm en een 4K-sensor. Pixelrasters zijn illustratief.

Een luchtgescheiden implementatie: de box beslist, de serverruimte antwoordt

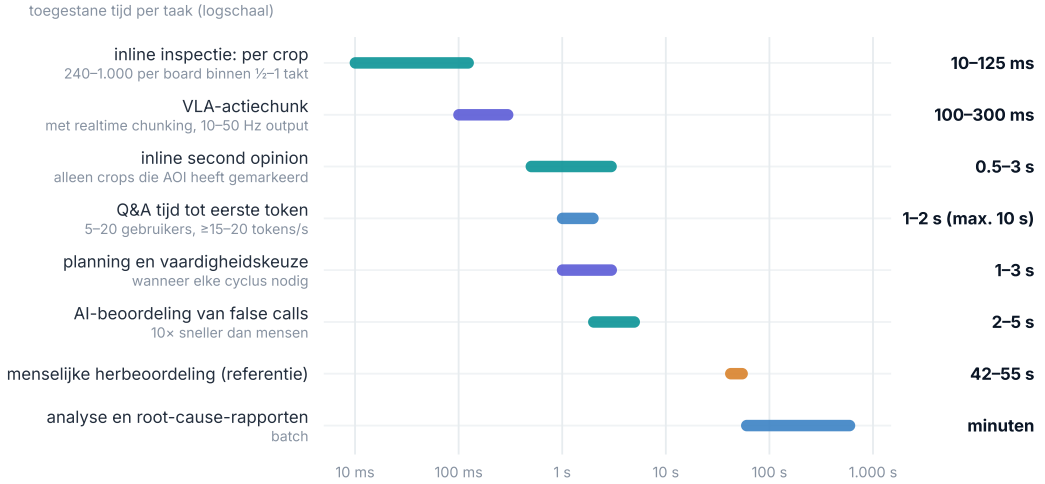


Bron: Synthese van de auteur (hoofdstuk 2).

Clean- en ESD-zones bepalen waar rekenkracht staat; snelheidslimieten hangen af van het gebruik

reinigheidseisen →

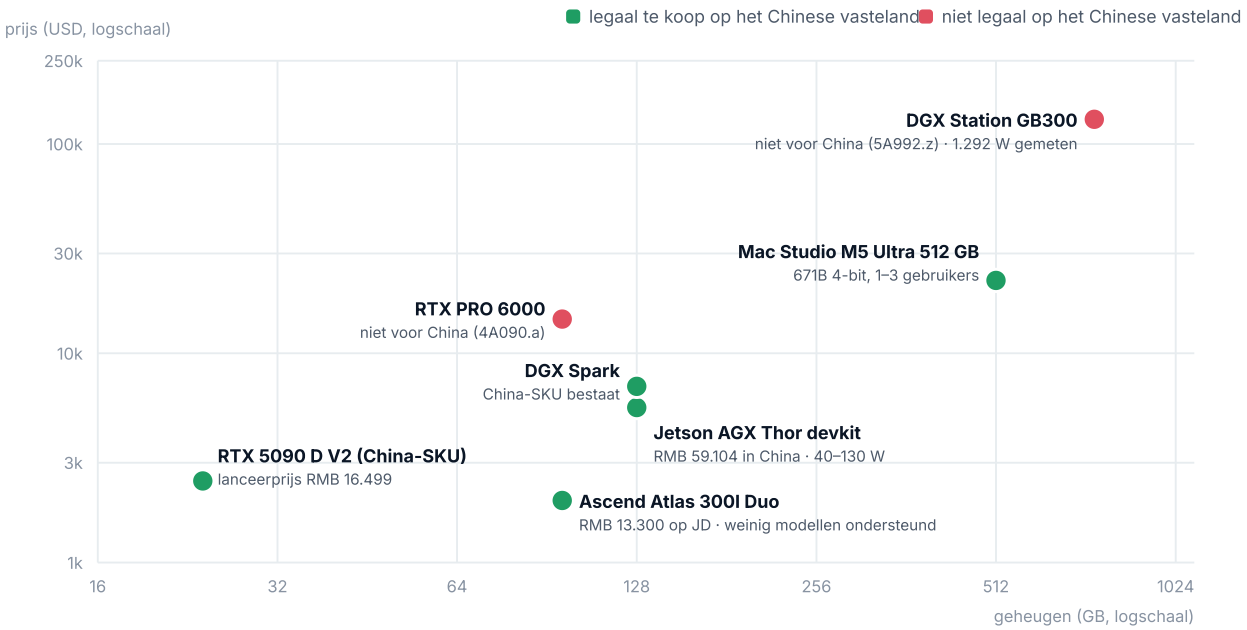
gewone SMT/PCBA-hal	cleanroom ISO 7–8	mini-omgeving ISO 5–6	fab, ISO 5 of schoner
ESD-beschermde zone (≈ISO 8) RV 40–60%	medisch, cameramodules, EV cleanroom + ESD	kale die, wire bond, optica mini-omgeving	SEMI E78 ESD · E187/E188 cyber fab-specifieke plannen
line-side compute ventilatorloze, geaarde metalen box naast de AOI of herbeoordelingsplek	**line-side compute** ventilatorloos, afgedicht IP65/66, makkelijk af te vegen; vraag ISO 14644-14-data	**line-side compute** alleen camera's en licht; rekenbox buiten	**line-side compute** alleen tool-geïntegreerde, fab-gecertificeerde controllers
GPU/NPU-server serverruimte of geklimatiseerde IT-rack	**GPU/NPU-server** buiten de cleanroom, netwerk naar binnen	**GPU/NPU-server** serverruimte	**GPU/NPU-server** vloeistofgekoelde racks in de subfab of het datacenter



Kernpunt Een gewone SMT-hal is een ESD-beschermde zone en geen geclassificeerde cleanroom: een ventilatorloze box in een geaarde metalen behuizing staat bij de AOI of de reviewplek, en GPU-servers komen nooit op de lijn.

Bron: ANSI/ESD S20.20, IEC 61340-5-1, ISO 14644, SEMI-normen en leveranciersmateriaal; samengesteld door de auteur. Synthese van de auteur (hoofdstukken 1–2).

Exportregels bepalen de hardware-shortlist eerder dan benchmarks



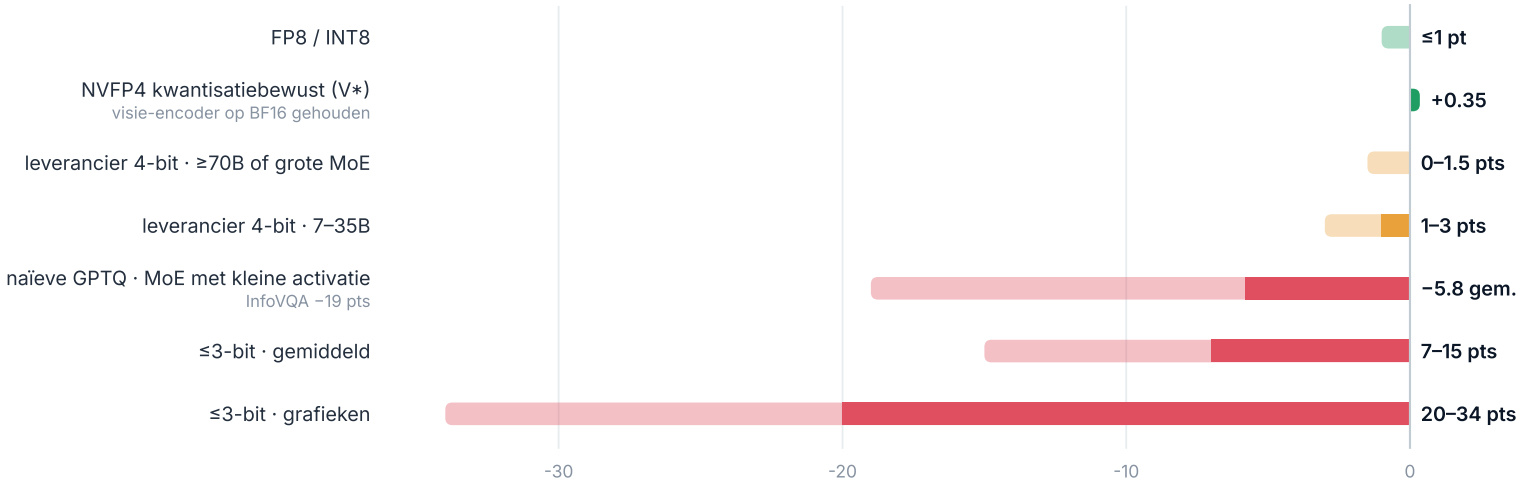
- **Jetson Thor:** 128 GB, 40–130 W; 2–8B vision-modellen op AOI-uitsnedes, VLA's van 3B-klasse in realtime; RMB 59,104 in China
- **Mac Studio 512 GB:** DeepSeek 671B op 4-bit voor 1–3 gebruikers
- **DGX Station:** 5A992.z, niet voor het vasteland; een 743B-model geeft 32 gebruikers elk ≈4 tokens/s
- **Binnenlandse accelerators:** twee Ascend-servers met acht kaarten voor volledige multi-user Q&A

Bron: NVIDIA-exportclassificaties, Apple, RS China, JD.com en reviewprijzen (november 2026). StorageReview (aug en okt 2026); ServeTheHome. Onafhankelijke tests uit één bron.

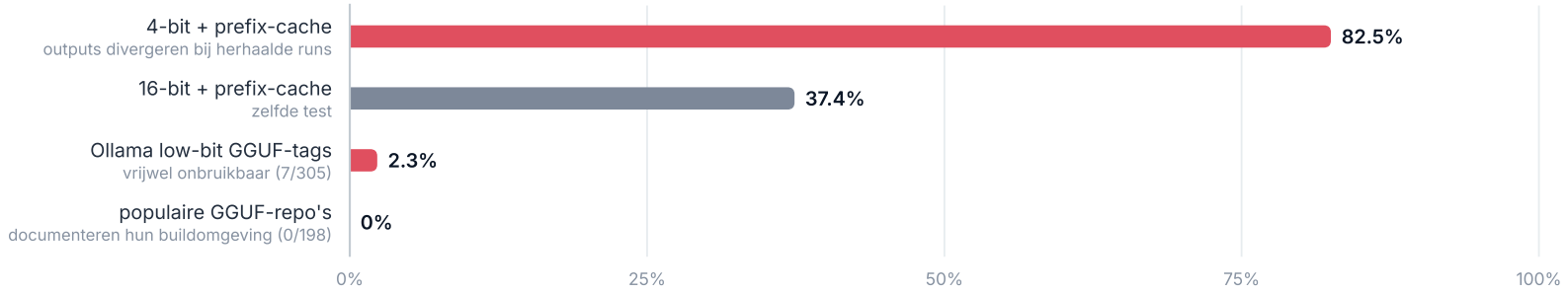
Quantisatie: 8-bit is veilig, 4-bit hangt van het recept af, 3-bit en lager valt af

nauwkeurigheid: 8-bit is veilig, 4-bit hangt van het recept af, ≤3-bit valt af

scoreverandering t.o.v. volledige precisie (punten)

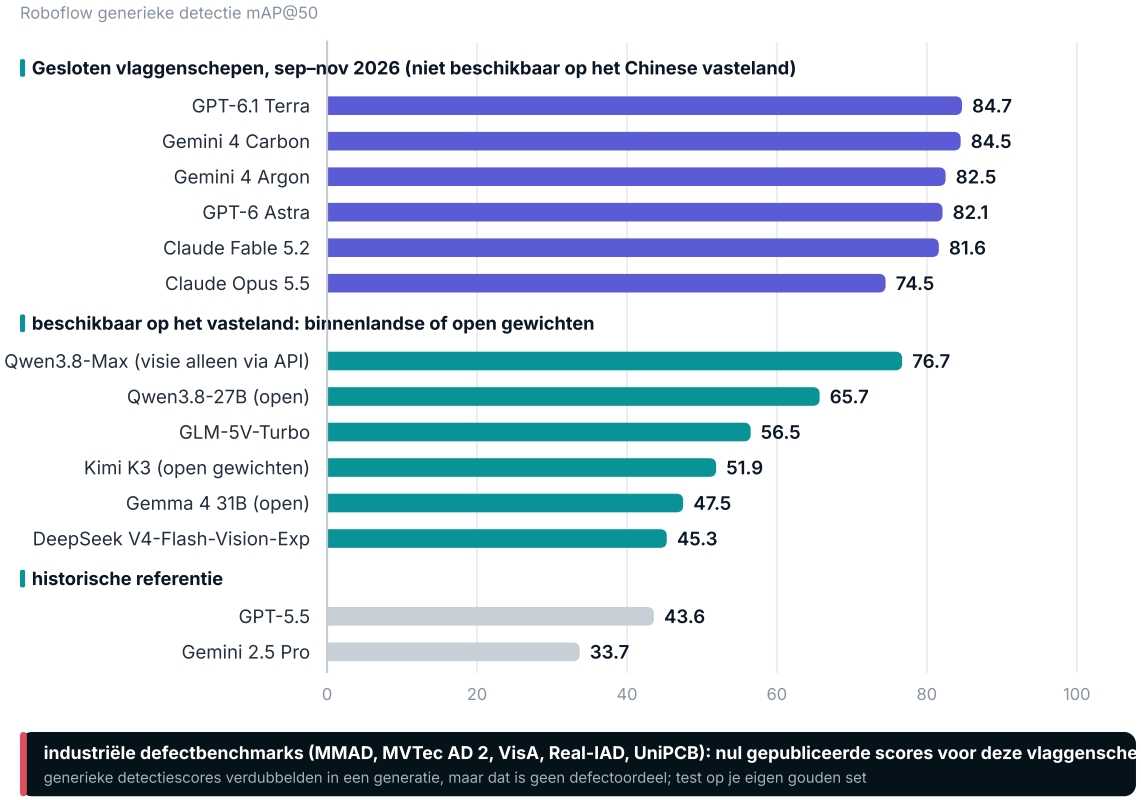


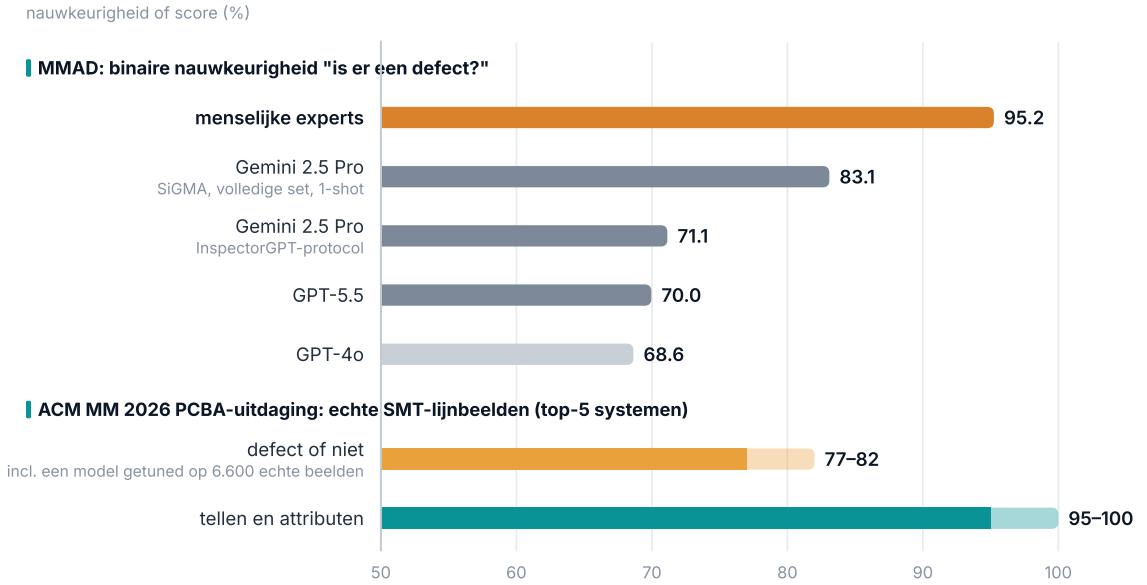
engineeringfouten wegen zwaarder dan quantisatiefouten



Bron: NVIDIA Nemotron-modelkaart, VLMQ, VEQ, arXiv 2609.05881, 2610.19498, 2609.04748 en anderen.

Generieke visie verdubbelde in één generatie; scores voor industriedefecten ontbreken



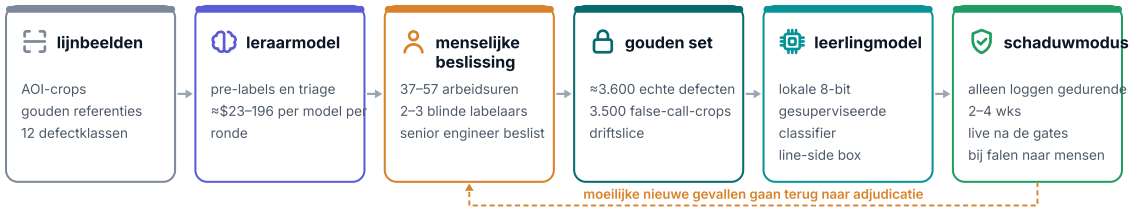


industriële defectbenchmarks (MMAD, MVTec AD 2, VisA, Real-IAD, UniPCB): nul gepubliceerde scores voor deze vlaggensche
generieke detectiescores verdubbelden in een generatie, maar dat is geen defectoordeel; test op je eigen gouden set

Bron: Roboflow Vision Evals; releasenotes van modellen; MMAD, MVTec AD 2 en andere benchmarkpagina's (per 11 nov 2026). MMAD; SiGMA; InspectorGPT; ACM MM 2026 PCBA-uitdaging.

Frontier-modellen pre-labelen, mensen beoordelen, een golden set fungeert als poort

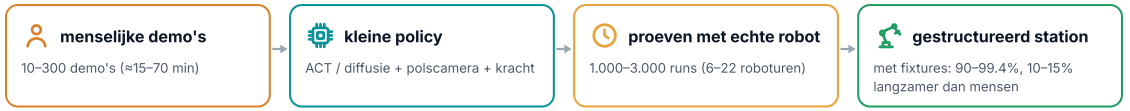
inspectie: het frontiermodel pre-labelt, mensen beslissen, een gouden set poort

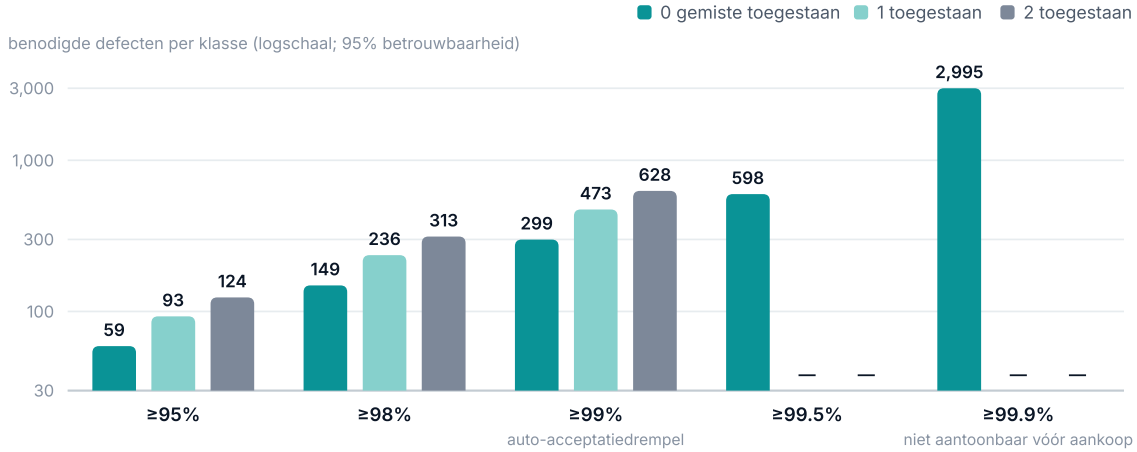


drempel voor automatische acceptatie van leraarlabels

per klasse: ≥299 defecten, nul gemist (recall ≥99% bij 95% betrouwbaarheid), overeenstemming niet slechter dan mens–mens; anders alleen pre-labeling

armen: de leraren zijn nog steeds menselijke demo's, RL op echte robots en simulatie — modellen helpen alleen offline: programmeren en ombouwen



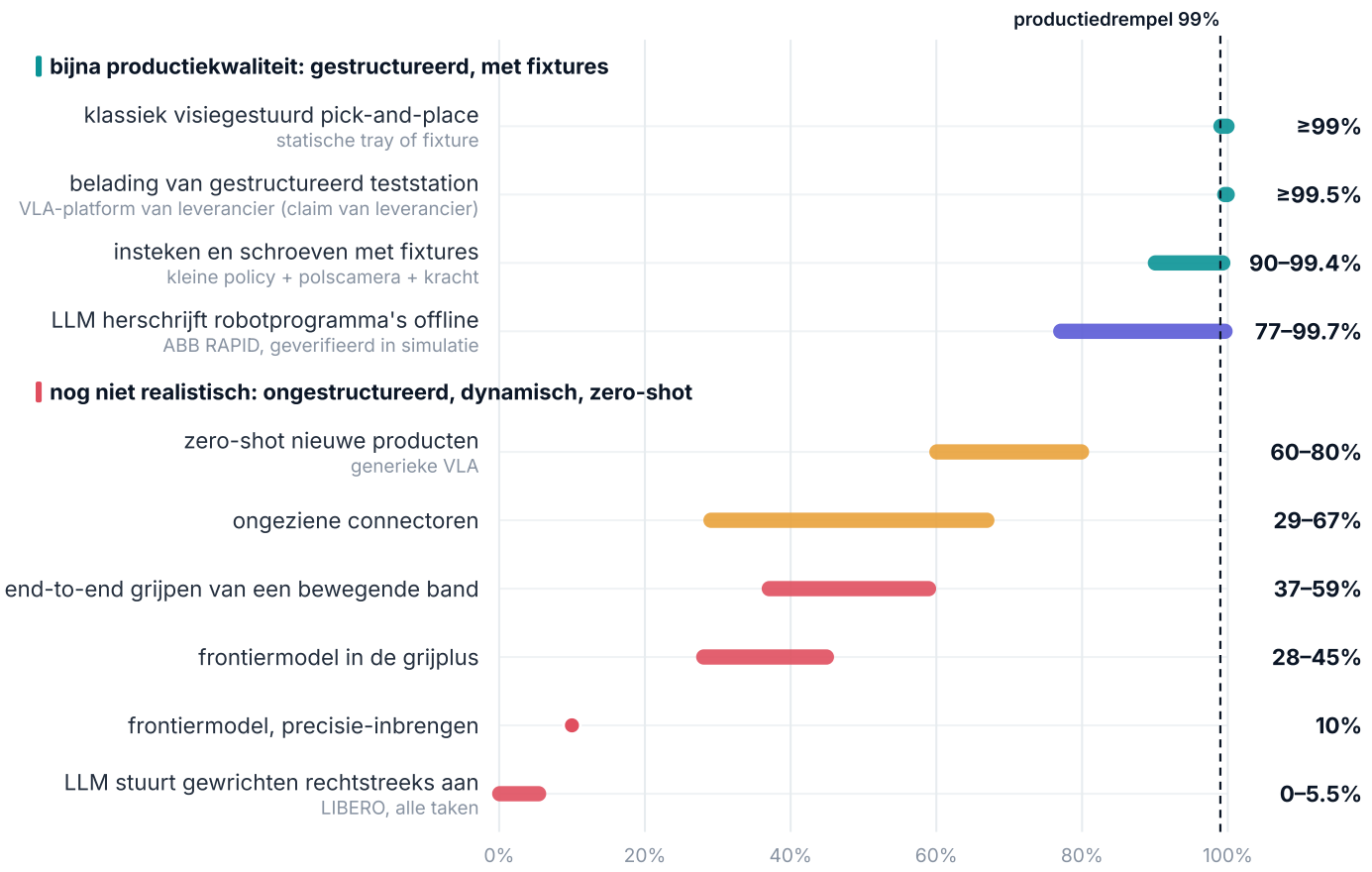


Kernpunt Accepteer de labels van de leraar alleen automatisch voor een klasse met nul missers op ≥299 golden defects (99% recall bij 95% betrouwbaarheid); kritieke klassen zoals kortsluitingen, ontbrekende onderdelen en polariteit vereisen meer.

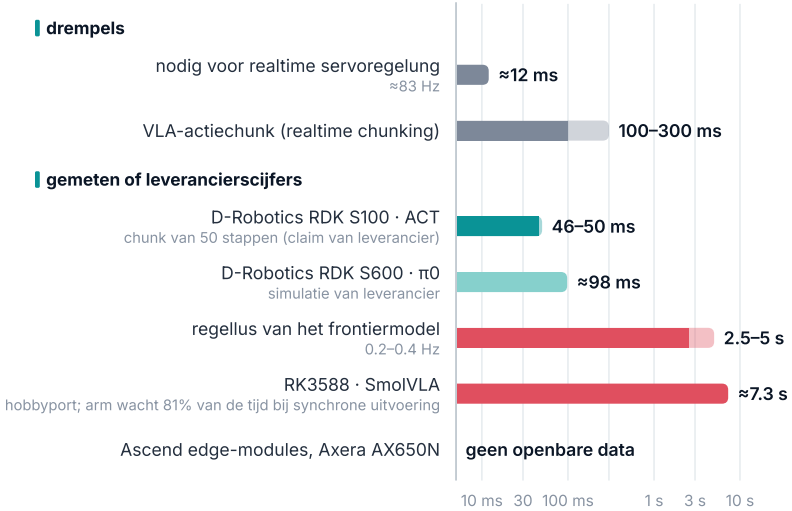
Bron: Synthese van de auteur (hoofdstuk 3); kosten en steekproefgroottes zijn schattingen. Exacte binomiale betrouwbaarheidsintervallen (95%).

Robotleerlingen: bijna productiekwaliteit alleen als het station gestructureerd is

taakslagingspercentage



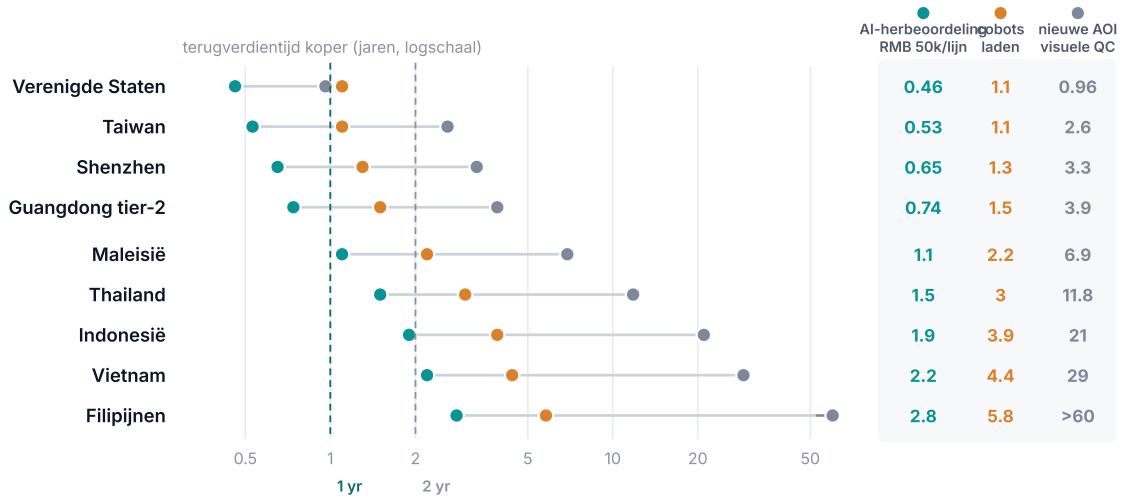
inferentielatentie per actiechunk of regelcyclus (logschaal)



Bron: Papers en leveranciersmateriaal uit hoofdstuk 3; sommige cijfers zijn claims van leveranciers. D-Robotics-documentatie en -forum (leverancierscijfers); community-ports op GitHub; onderzoek van Anthropic.

AI-review verdient zich in maanden terug; machinebediening in ongeveer een jaar



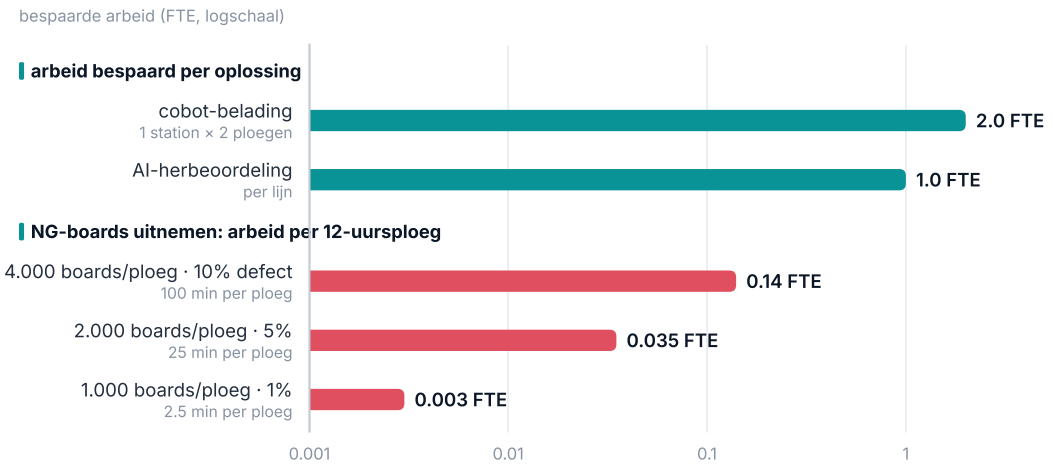


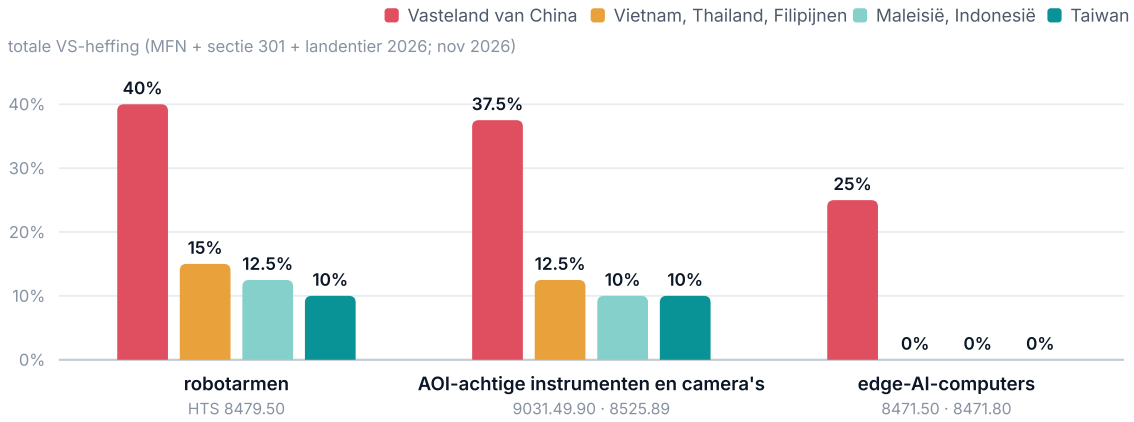
een robotarm enkel om NG-boards uit te nemen betaalt zich nergens terug
bij first-pass yield boven 90% kost uitnemen slechts 0.003–0.14 FTE per ploeg

Kernpunt Bij RMB 50,000 per lijn verdient AI-review zich in 0.46 jaar terug in de VS, 0.53 in Taiwan, 0.65 in Shenzhen en 2.2 in Vietnam; een cobotbedieningscel in twee ploegen in 1.1–1.5 jaar [estimates].

Bron: Officiële loonrichtlijnen van Shenzhen, Ministerie van Arbeid van Taiwan, JETRO, BLS; schattingen van de auteur. Model van de auteur: RMB 50k AI-herbeoordelingsbox per lijn; cobot-beladingscel van RMB 200k of $140k, twee ploegen.

Een uithaalarm verdient zich nooit terug; waar een arm wordt gemaakt, verandert zijn Amerikaanse invoerrecht





Kernpunt Boven 90% first-pass-opbrengst kost het uithalen van NG-platen 0.003–0.14 van een persoon per ploeg, dus een NG-afleider of -buffer doet het werk; een in het vasteland gemaakte arm betaalt ook ongeveer 40% gecombineerd Amerikaans invoerrecht.

Bron: Schattingen van de auteur. USITC Harmonized Tariff Schedule hoofdstuk 99; Federal Register; USTR (november 2026).

Reëel maar klein: 16,000 adresseerbare lijnen, ≈RMB 320M eenmalig, een venster van vijf jaar

SMT-lijnen in China
40.000 regels
35k–46k

× 80% × 50%
nog geen AI: 75–85%
bereikbaar: 40–60%

serviceerbaar
16.000 regels
10.5k–23.5k

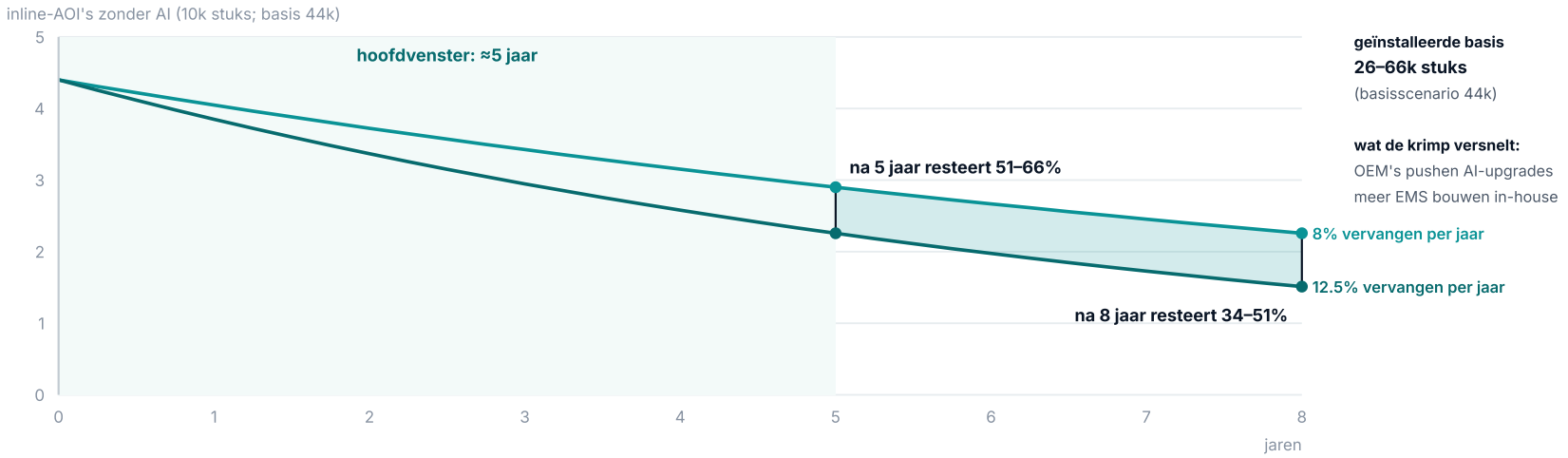
× RMB 20k
RMB 10–30k eenmalig

eenmalige markt
RMB 320M
RMB 105–705M

÷ ≈5 jaar
krimpt 8–12%/yr

venster van 5 jaar
RMB 64M/yr
RMB 21–141M

een nieuwkomer die in 3 jaar 2–5% van de serviceerbare lijnen wint: ≈320–800 lijnen, ≈RMB 2–5M per jaar
als abonnement (RMB 5k per lijn per jaar): RMB 1.6–4M recurring, slechts 0.3–2.3% van IDC's prognose voor industriële AI-QC in China



Bron: NVR (basis 2020); schattingen van de auteur. Schattingen van de auteur; geïnstalleerde basis en aandelen niet onafhankelijk geverifieerd.

De reviewwig is omsingeld; de marge zit in geproductiseerde hardware

top-EMS bouwen zelf

DBG (2025): dekt ≥95% van gemarkeerde onderdelen
escapes <20 PPM, menselijke herbeoordeling −85%
dit is de lat voor verkoop aan tier-1 EMS

AI ingebouwd in AOI door OEM's

Jutze AI 3D AOI in productie, "verwijdert menselijke herbeoordeling bij inspectie"
ViTrox V510Ai; Sinic-Tek ontwikkelt samen met Tencent
nieuwe machines komen met AI; oude hebben upgrades nodig

AI-herbeoordeling door derden

box + software op bestaande AOI geschroefd

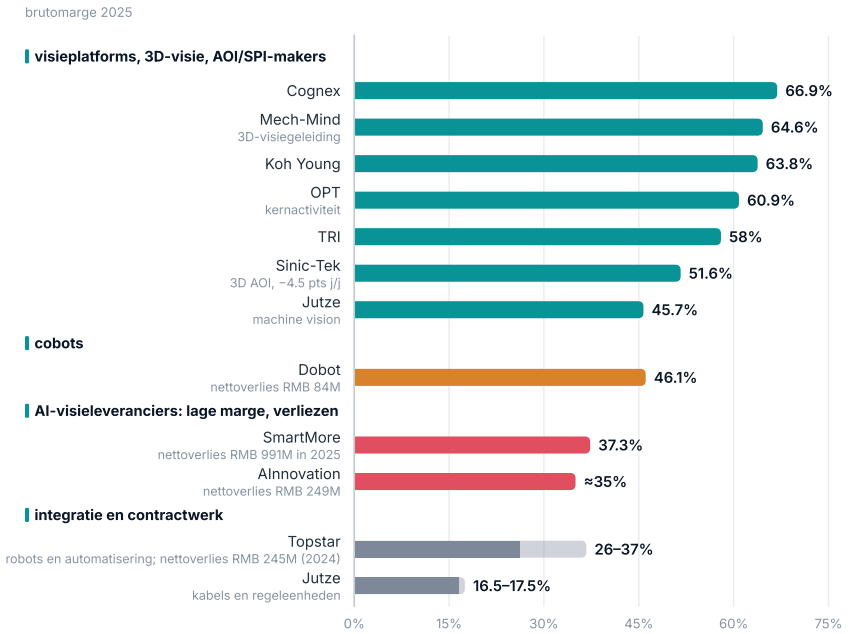
platformleveranciers

Hikvision Guanlan en SmartMore IndustryGPT verkopen
leidende AI-leverancier: omzet 2025 RMB 1,086B
nettoverlies RMB 991M: prijzenoorlog

wat overblijft voor derden

middelgrote EMS (≈10–40 lijnen) met oudere of gemengde AOI en geen OEM-AI-upgrade

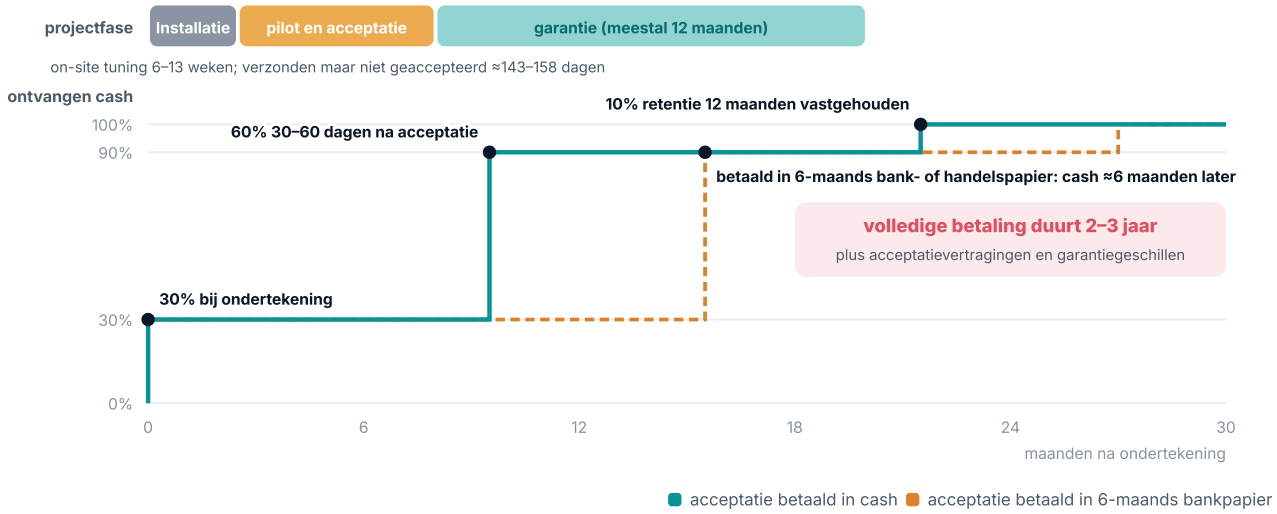
Niches die AOI slecht dekt: THT en handsolderen, conformal coating, kabels/connectoren/FPC, eindassemblage en schroeven

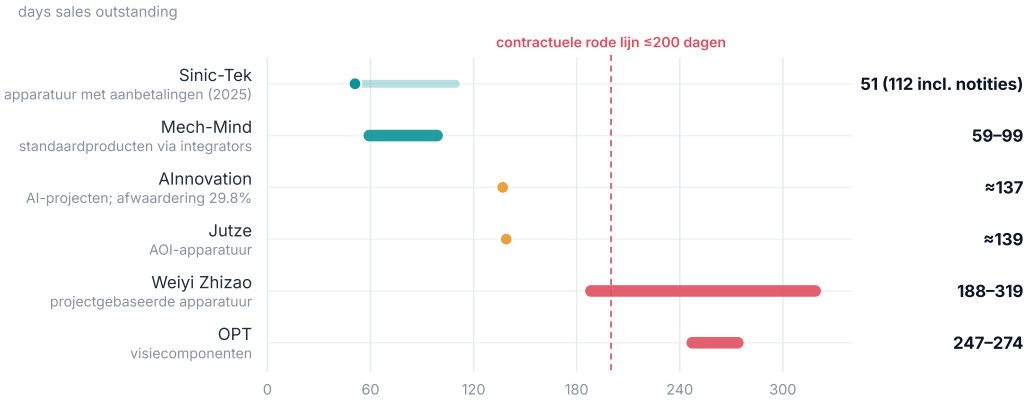


Bewijs De leidende AI-visieleverancier had in 2025 een omzet van RMB 1.086 miljard en een nettoverlies van RMB 991 miljoen; de 3C-omzet van het dichtstbijzijnde beursgenoteerde vergelijkingsbedrijf daalde in januari–mei 2026 met 79.9% op jaarbasis.

Bron: Jaarverslagen van Jutze, Sinic-Tek en DBG; publiek materiaal van Hikvision en SmartMore; samengesteld door de auteur. Jaarverslagen of prospectussen van 2025.

Kopers betalen 30/60/10; volledige betaling duurt twee tot drie jaar





Kernpunt 30% komt met de bestelling, 60% ongeveer vier tot negen maanden later en de laatste 10% ongeveer 16–21 maanden later; van een eerste pilot bij een grote EMS tot volledige betaling kan twee tot drie jaar duren [estimate].

Bron: Antwoord op vragen van Sinic-Tek, Mech-Mind, GKG, JT en andere stukken; schema op gebruikelijke voorwaarden, timing geschat. Bedrijfsstukken; vorderingen (incl. contractactiva) ÷ omzet × 365.

Plan offline, handel met kleine policies, houd beweging en veiligheid deterministisch

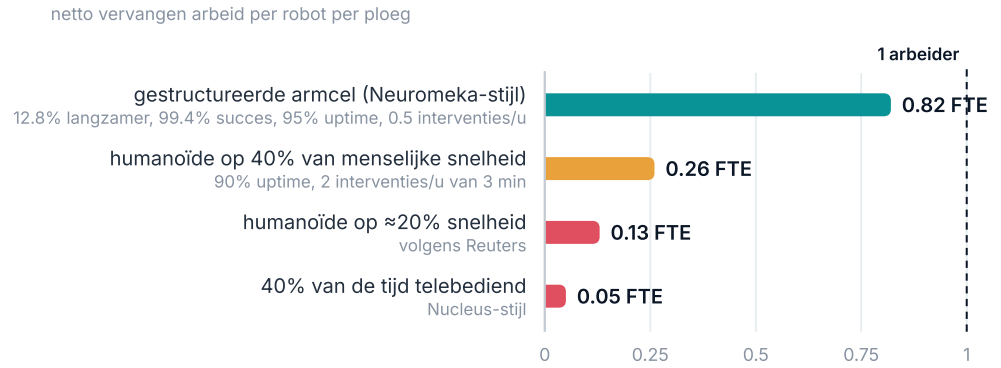
Laag	wat het doet en hoever het bewijs reikt	tijdschaal · wie bezit het
Planning L4	**frontiermodellen (offline)** robotprogramma's offline schrijven of wijzigen, ombouwen, fouten uitleggen herschrijft ABB RAPID-programma's op 77–99.7%, uitgerold na simulatie regellus slechts 0.2–0.4 Hz: blijft buiten realtime	**seconden tot minuten** modelgiganten
vaardigheden / poli L3	**ACT, diffusiebeleid, kleine VLA's** gestructureerde stations met fixtures 90–99.4% edge-chips: Jetson Thor, D-Robotics RDK, Ascend, Rockchip	**actiechunks 100–300 ms** chipmakers · robotmakers
bewegingsregeling L2	**robotcontroller** trajecten, servo, krachtregeling, botsingscontroles deterministische code; geleerde policies stellen alleen doelen	**milliseconden (realtime vereist >83 Hz)** robotmakers
Veiligheid L1	**veiligheids-PLC, scanners, kracht- en snelheidslimieten** gecertificeerd volgens ISO 10218-1/-2:2025 en lokale regels geen ML in veiligheidsfuncties: EU-beoordeling door aangemelde instantie vanaf 20 jan 2027	**hard realtime** gecertificeerde robot + integrator

kleine bedrijven winnen in integratie en sensorhardware: zet geleerde perceptie en kleine policies in gecertificeerde robotcellen, station voor station

	slagingspercentage	takt	interventies / teleoperatie	uptime / storingen	volumeproductie
Figure 02 @ BMW	alleen doel >99%/ploeg	alleen doel 84 s	alleen doel 0/ploeg	n/b	1.250 h
Agility Digit @ GXO	n/b	n/b	n/b	n/b	100.000 bakken
Apptronik Apollo @ Mercedes	n/b	n/b	n/b	n/b	n/b
Agile ONE @ Audi	n/b	n/b	n/b	n/b	n/b
Nucleus (niet nader genoemde fabriek)	n/b	n/b	≈40% telebediend*	n/b	≈2 h video
AgiBot G2 @ Longcheer	99.99%	18–20 s per stap	n/b	één storing, 2 min	64+ h
hybride cel van Neuromeka	322/324	12.8% langzamer	n/b	n/b	324 cycli
π0.6 @ klantlocatie	—	—	96.4% autonoom	n/b	1 ploeg

* methode niet bekendgemaakt

bekendgemaakt (meestal claims van leveranciers) · alleen doelen · n/b



Bron: synthese van de auteur (hoofdstukken 3, 6) en schattingen; ISO 10218:2025; EU-machineverordening; openbaar materiaal van BMW, Figure, Agility, Mercedes-Benz en anderen.

Bouw smal, bewijs het op de eigen beelden van de fabriek, schaal dan op

Bouwen

AI-review en defectclassificatie op AOI-beelden voor middelgrote EMS met oudere of gemengde AOI-merken; inspectie die de AOI slecht dekt; cobotcellen waar elk onderdeel wordt gehanteerd

Niet bouwen

Een algemeen AI-plus-armplatform; een LLM als primaire inspecteur; een uithaalarm op een SMT-lijn; humanoïden en robots-as-a-service; frontier-modellen in regelkringen of veiligheidsfuncties

Pilot

1. **Weken 0–2:** dataovereenkomst, defectklassen, nulmeting
2. **Weken 1–4:** golden set, blind gelabeld en beoordeeld
3. **Weken 2–4+:** schaduwmodus, per klasse vergeleken
4. **Weken 4–8:** geassisteerd draaien, alleen op goedgekeurde klassen
5. **Daarna:** schriftelijke acceptatie, lijn voor lijn opschalen, opnieuw valideren bij elke wijziging

Veiligheid Geen machine learning in veiligheidsfuncties: de EU vereist een beoordeling door een aangemelde instantie vanaf 20 januari 2027, en AI-certificeringstrajecten worden niet vóór 2027–2029 duidelijk.

Bron: hoofdstukken 6–7 en managementsamenvatting van het rapport; ISO, IEC, UL, EUR-Lex, SAC en nationale regelgevingsdatabases.

OVER DE AUTEUR

Tan Shuai

Technoloog en oprichter, met meer dan twintig jaar ervaring in chips, hardware en systeemsoftware en 20 verleende patenten. Zijn werk voerde hem langs Silicon Valley, New York, Chicago, Taiwan, de Greater Bay Area (Shenzhen en Hongkong), Vietnam en Maleisië, met een langdurige focus op geavanceerde productie en toeleveringsketens. Vandaag werkt hij als AI-onderzoeker en -ontwikkelaar, gericht op hoe AI op de fabrieksvloer wordt gebruikt.

Het volledige rapport (alle figuren, tabellen en bronnen) en het artikel staan op tanshuai.com. Hij verwelkomt gesprekken met mensen die aan AI in de productie werken: i@tanshuai.com

