

The Skeptic's Field Guide

How to Evaluate Robotics Vendors When You Have Never Bought a Robot

The robotics deployment trilogy (REA·004–006) established what a robot is worth, why deployments stall, and what a RaaS contract actually prices. This paper turns to the moment that precedes all three: the purchase decision itself. Its premise is simple. The first-time robotics buyer faces one of the most asymmetric information environments in modern capital equipment — a market where vendor valuations run one hundred times ahead of audited revenue, where the demonstration is a choreographed artifact, and where the seller may not exist by the time the machine needs its first major service. The remedy is not enthusiasm, and it is not abstention. It is procedure.

Prepared by Kevin Nussbaum / Applied Futures Research

August 2026 · zspacedynamics.com

Epistemic register: claims are tagged GROUNDED (sourced fact), ANALYTIC (derived from stated assumptions), PROJECTION (forecast under uncertainty), or SPECULATIVE (structured conjecture).

Executive Summary

A buyer who has never purchased a robot is not purchasing a machine. They are purchasing a **bundle of dependencies**: hardware, a proprietary software stack that only the vendor can update, a spare-parts channel that only the vendor supplies, a support organization whose depth is unknowable from the outside, and — above all — a bet on the vendor's continued existence. In mature capital equipment categories (machine tools, lift trucks, conveyors), a century of market consolidation has made that bundle boring. In robotics, and especially in mobile manipulation and humanoids, the bundle is the entire risk.

The 2026 market presents the sharpest version of this problem the industry has seen. On one side, the established categories are healthy and measurable: North American companies ordered 36,766 industrial robots worth \$2.25 billion in 2025, up 6.6% in units and 10.1% in revenue over 2024, with collaborative robots the fastest-growing segment. On the other side, the frontier categories are priced for a future that has not arrived: over \$8 billion has flowed into more than 100 humanoid robotics companies, one of which carries a ~\$39 billion private valuation without publicly disclosed deployment economics — while the first humanoid maker to face public-market audit, Agility Robotics, went out at \$2.5 billion. That 15-to-1 spread between the most-hyped private valuation and the first audited one is the single best summary of the information asymmetry a buyer faces. **GROUND**

\$2.25B 2025 NA ROBOT ORDERS <i>36,766 units, +6.6% YoY (A3)</i>	\$8.1B+ HUMANOID FUNDING <i>100+ companies, mostly pre-revenue</i>	15 : 1 VALUATION SPREAD <i>Top private vs. first audited humanoid</i>	\$250K STRANDED-ASSET CASE <i>One customer's idle machines, 2026</i>
--	---	--	--

This paper delivers a complete evaluation procedure for that environment: a vendor-mortality base rate the buyer should treat as actuarial fact (Section 3); a method for decoding demonstrations and spec sheets that were engineered to be believed rather than verified (Sections 4–5); a balance-sheet diligence protocol that requires no finance background (Section 6); a reference-check and site-visit script (Section 7); the contract clauses that convert vendor risk into vendor obligation (Section 8); and a weighted scorecard with a disqualification taxonomy (Section 10). The verdict, developed in the closing scenario table, is deliberately unromantic: **buy the boring categories on process, pilot the frontier categories on contract, and let someone else's balance sheet absorb the mortality risk of the rest.** **ANALYTIC**

The core discipline of this field guide in one sentence: treat every vendor claim as a hypothesis, every demo as a press release, every spec sheet as marketing until reproduced on your floor, and every startup balance sheet as a countdown timer — then buy anyway, deliberately, from whoever survives that treatment.

1 Why This Purchase Is Different

Industrial buyers have well-worn instincts for capital equipment. Those instincts fail on robotics in four specific ways and naming them precisely is the first defensive act.

1.1 The asset is a subscription wearing a machine costume

A CNC machine purchased in 2010 still cuts metal today regardless of if its manufacturer exists. A modern robot — particularly anything with learned manipulation policies, fleet-management software, or cloud-tethered perception — degrades toward brick the moment its vendor stops shipping updates, security patches, and parts. REA-004 valued the humanoid as a capital asset; the honest amendment this paper adds is that the asset's useful life is **min (mechanical life, vendor life, software-support life)** — and for most frontier vendors, the binding constraint is not the first term. **ANALYTIC**

1.2 Information asymmetry is structural, not incidental

The vendor knows its intervention rates, its field failure modes, its true cost of support, and its cash runway. The buyer knows what the demo showed and what the salesperson said. In mature markets, this gap is closed by third-party data: resale values, independent service networks, decades of fleet records, public financials. In robotics the third-party data layer barely exists. Agility's SPAC filing is notable precisely because it will be the first time a pure-play humanoid company's claims are subject to quarterly public scrutiny — the entire rest of the category is self-reported. **GROUNDED**

1.3 The seller's survival is a spec

No one evaluates a forklift on the OEM's Series C. For a robotics purchase, the vendor's mortality risk belongs on the spec sheet next to payload and reach, because it directly determines expected asset life, parts availability, and software support. Section 3 quantifies why: the base rates are bad enough that ignoring them is not optimism, it is malpractice. **ANALYTIC**

1.4 The category rewards theatrical competence

Robotics is unusually demo-able. A machine that works 40% of the time can produce a flawless three-minute video; a vendor six months from insolvency can stage an immaculate facility tour. Because capital flows toward spectacle — the funding data in Section 2 shows capital rewarding demos at historic scale — the industry has been trained, rationally, to invest in the appearance of capability. The buyer's counter-move is to design evaluations the vendor cannot choreograph. That is Sections 4 and 5. **ANALYTIC**

A useful mental model: you are not a customer in a store. You are a counterparty in a negotiation with a party that has rehearsed, and you have not. This guide is the rehearsal.

2 The State of the Vendor Field, 2026

Vendor evaluation begins with an accurate map of the market's structure, because the evaluation procedure should differ radically by segment. The 2026 field is divided into three tiers with fundamentally different risk profiles.

2.1 Tier 1 — The measurable core: industrial arms, cobots, AMRs

The established market is real, growing, and boring in the best sense. Per the Association for Advancing Automation (A3), North American companies ordered 36,766 robots valued at \$2.25 billion in 2025 — a 6.6% unit increase and 10.1% revenue increase over 2024 — with non-automotive demand outpacing automotive for consecutive years and general industries taking the majority of units. Momentum built through the year: Q3 2025 alone saw 8,806 robots ordered at \$574 million, up 11.6% in units year-over-year. Collaborative robots are the standout sub-segment: A3 began tracking cobots as a distinct category in Q1 2025 (1,052 units, \$39.2 million, 11.6% of units); by Q4 2025 cobots hit a record 28.6% of quarterly orders, and in Q1 2026 companies ordered 1,637 cobots — up 55.6% year-over-year, with revenue up 78.2% to \$69.8 million, 18.1% of all units. Globally, industrial robot shipments are forecast to grow from roughly 550,000 units in 2025 to 760,000 by 2030. **GROUND**

For the buyer, Tier 1's practical meaning: vendors here (the established arm OEMs, the major cobot makers, the scaled AMR fleets) have installed bases in the tens or hundreds of thousands, dealer/integrator networks, published safety certifications under standards like the new ANSI/A3 R15.06-2025, and — for the public ones — audited financials. Vendor mortality risk is low; the evaluation problem shifts almost entirely to application fit, integration quality, and total cost of ownership. Most first-time buyers should be shopping here, and Section 10's scorecard weights reflect that. **ANALYTIC**

2.2 Tier 2 — The venture-scaled challengers

Between the industrial core and the humanoid frontier sits a broad band of venture-backed vendors in picking, palletizing, inspection, agriculture, construction, and food service. These companies often have genuinely deployed product and real revenue, but venture economics: they must grow into their valuations or die trying. Industry trackers put total robotics-sector funding at roughly €37.9 billion in 2025, up from €13.8 billion the year prior — an extraordinary expansion of the pool, which cuts both ways for buyers. Capital abundance extends runways; it also funds competitors into your vendor's niche and postpones the reckoning rather than canceling it. **GROUND**

2.3 Tier 3 — The humanoid frontier: priced for a different decade

The humanoid segment deserves its own risk category because its financial structure is unlike anything else in capital equipment. More than \$8.1 billion has been raised across 100+ humanoid companies. Figure AI closed over \$1 billion in Series C funding at a reported ~\$39 billion post-money valuation without disclosing comparable commercial deployment data. Apptронik raised more than \$935 million across its Series A at a valuation above \$5 billion, with

deployments at Mercedes-Benz and GXO. Stage labels have lost meaning: the average humanoid 'Series A' in the twelve months to mid-2026 was over \$212 million — manufacturing-scale financing wearing a product-validation label. **GROUND**

Against that backdrop, the most informative single event of 2026 is Agility Robotics' decision to go public via SPAC at a \$2.5 billion valuation — the first U.S.-listed pure-play humanoid. Agility's disclosed numbers are the closest thing the segment has to a floor of audited truth: roughly 1,000 robots tied to more than \$300 million in booked multi-year revenue (the filings themselves caution this is potential multi-year value tied to a three-year RaaS contract, subject to milestones — not current-period revenue), about 65,000 hours of Digit operation across nine customer facilities, and \$620+ million in expected proceeds, with Foxconn leading the PIPE. Third-party market sizing puts the entire global humanoid market at roughly \$1.4 billion in 2026, projected to \$19.6 billion by 2033. **GROUND**

Read those numbers together and the asymmetry becomes arithmetic. The segment's aggregate private valuation is plausibly north of \$60 billion against a total current market measured near \$1.4 billion — the sector is capitalized at roughly 40 times its own present-year revenue pool, *before* any single vendor's share. The most operationally proven vendor in the category priced at ~1/15th of the most richly valued one. For a buyer, the inference is not that humanoids are fake — 65,000 real hours at nine facilities is not fake — but that **valuation is noise and deployment hours are signal**, and every evaluation weight in this guide follows from that principle. **ANALYTIC**

Tier	Examples of segment	Typical vendor profile	Dominant buyer risk	Evaluation posture
Tier 1 — Measurable core	6-axis arms, cobots, scaled AMRs	Public or long-established; huge installed base; audited financials	Application fit, integration cost	Buy on process; negotiate on TCO
Tier 2 — Venture-scaled	Piece-picking, palletizing, inspection, ag/construction	Series B–D; real revenue; must grow into valuation	Vendor mortality; support depth	Full diligence per this guide; contract armor
Tier 3 — Humanoid frontier	General-purpose bipeds/mobile manipulators	Mega-funded, mostly pre-revenue; self-reported metrics	Everything at once	Structured pilot only; assume vendor turnover

3 Vendor Mortality: The Actuarial Reality

The polite fiction of vendor selection is that all finalists will exist in five years. The data says otherwise, and three well-documented deaths — one incumbent, one application startup, one humanoid — show the distinct ways robotics vendors fail and what each failure mode did to customers.

3.1 Base rates

Startup mortality generally: Carta recorded 966 U.S. startup shutdowns among its customers in 2024, up 25.6% from 769 in 2023, and analysts expected the wave to continue as companies funded in the 2020–21 boom exhausted reserves. Robotics adds hardware's burn rates and long sales cycles on top. In China — the deepest robotics ecosystem by company count — 2025 produced a striking split: roughly 610 robotics financing events (nearly triple the prior year, exceeding ¥57 billion) at the same time the industry's first wave of humanoid shutdowns arrived, with leading firms absorbing capital while trailing firms were eliminated at accelerating pace. Polarization, not tide-lifts-all-boats, is the pattern a buyer should internalize. **GROUND**

3.2 Case one — Rethink Robotics: the incumbent that misread the requirement

Rethink was the most credentialed failure in the category's history: founded by an icon of the field, lavishly funded, and the company that effectively created the cobot category with Baxter and Sawyer. It died in 2018 because its design philosophy — ease of use, flexibility, an approachable face — diverged from what industrial customers actually specify: precision and reliability. Post-mortems consistently converge on the same lessons: technological novelty means little if misaligned with the customer's core requirement, and closed ecosystems lose to open ones (Universal Robots won the category in part by welcoming third-party components and integrations). **GROUND**

Buyer lesson: pedigree is not a spec. A famous founder, category-creator status, and deep funding did not make the product meet requirement. Your evaluation must test the vendor against your requirements, not their narrative — and vendors whose ecosystems are open (third-party grippers, standard fieldbuses, published APIs) leave you salvageable options that closed ecosystems do not. **ANALYTIC**

3.3 Case two — Picnic: the stranded customer

Picnic Technologies built pizza-assembly robots, took real customers, pivoted under a new CEO toward a 'hospitality-first automated pizza chain' in late 2025, and by 2026 had shut down entirely and liquidated — its Seattle headquarters found empty, neighbors reporting the company gone months earlier. The detail that belongs in every procurement file is what it left behind: a Seattle restaurant owner, an early public champion of the technology, described being stuck with a \$250,000 'robot aquarium' — his phrase for the idle machines still sitting in his restaurant. **GROUND**

Buyer lesson: this is the stranded-asset scenario in its purest form, and it illustrates the asymmetry of early-hardware risk: the vendor's downside was bounded by its investors' capital; the customer's downside was a six-figure ornament plus the operational plans built around it. Every contract protection in Section 8 — escrow, parts guarantees, wind-down obligations, refund triggers — exists to redistribute exactly this asymmetry before signature, because after the shutdown there is no counterparty left to negotiate with. **ANALYTIC**

3.4 Case three — K-Scale Labs: the frontier startup as engineering culture warning

K-Scale Labs, a YC-backed startup attempting low-cost humanoids, never closed its Series A and was gone by late 2025, subsequently open-sourcing its IP. Its former COO's published post-mortem contains a diagnostic gem for buyers: the team spent a 'genuinely long time' debating whether to install mechanical end stops — the most basic safety redundancy in actuated hardware — because one faction argued the AI policy should learn the joint limits. His rebuttal is the entire hardware discipline in three sentences: end stops exist because software fails; the model might be right 99.99% of the time; in the physical world, the 0.01% is the only number that matters. **GROUNDED**

Buyer lesson: this debate is a litmus test you can administer. Ask any frontier vendor's engineers — not sales — how the machine behaves when the policy produces a bad inference: what physically prevents harm, independent of software? Vendors with crisp answers (mechanical stops, torque limits, certified safety-rated monitored stops, independent e-stop chains) have hardware culture. Vendors who answer with model accuracy statistics have demo culture. Agility's positioning of Digit v5 around 'cooperative safety' and its selection as launch partner for NVIDIA's Halos safety stack shows the mature end of the spectrum: safety as architecture, not aspiration. **ANALYTIC**

3.5 Pricing mortality into the evaluation

Treat vendor survival as a probability, not a vibe. A defensible starting prior for a five-year horizon, absent specific diligence: Tier 1 established OEM ~95%+ survival (of the product line, via the company or an acquirer honoring support); Tier 2 venture-scaled with real revenue ~60–75%; Tier 3 pre-revenue frontier ~35–50%, with enormous variance — and note that 'acquired' frequently means 'product sunset,' which for the buyer is functionally death with better press. The expected-value consequence is mechanical: a machine with a 7-year mechanical life bought from a vendor with a 50% five-year survival probability has a materially shorter expected supported life, and the NPV math of REA-004 must be run on the supported life, not the nameplate life. **PROJECTION**

Underwriting rule of thumb: for any Tier 2–3 vendor, price the deal as if you will need to exit or re-platform at the midpoint of the contract and let every dollar of the vendor's contract concessions buy that assumption down.

4 The Demo Is Choreography

A robotics demonstration is a produced artifact, engineered under the vendor's control to display the machine's best reachable behavior. That is not an accusation of fraud; it is the job description of a demo. The buyer's task is to understand the standard production techniques and then design an evaluation that removes them.

4.1 The standard production toolkit

- **Teleoperation and human-in-the-loop assistance.** The single most important question of the humanoid era: was the machine autonomous, teleoperated, or 'supervised autonomy' (a human approving or correcting plans)? Vendors vary widely in disclosure norms. Ask for the autonomy ratio in writing for anything you are shown.
- **Selection effects.** One flawless run can be the survivor of forty attempts. Ask: how many takes? Is this footage continuous? What is the success rate across the last 100 attempts of this exact task — and will you show me the log?
- **Speed manipulation.** Video played above real time flatters cycle rates. Insist on timestamped, continuous, single-take footage — or better, live performance.
- **Environmental staging.** Perfect lighting, high-contrast fixtures, pre-sorted parts, empty aisles, dry floors. Every deviation between the stage and your floor is unquantified risk you are absorbing.
- **The hero unit.** The demo machine may be a hand-tuned, engineer-maintained flagship unrepresentative of production units. Ask the serial number, its build revision, and whether production units match its configuration.
- **Task cherry-picking.** The machine performs the task it is best at, which is presented as representative. Your evaluation must use your task mix, weighted by your actual volume distribution.

4.2 The counter-demo: a buyer-designed test protocol

The remedy is to move the burden of proof: you design the test, the vendor performs it. A credible counter-demo protocol has six elements:

1. **Your inputs.** Your parts/SKUs/totes/pallets, sampled from real production including the ugly ones — damaged cartons, shrink-wrap glare, mixed-case pallets, the 5% of items that cause 50% of exceptions.
2. **Your environment, or a faithful proxy.** Your lighting, floor conditions, rack spacing, temperature, dust, and traffic. If testing at the vendor's site, you specify the environmental setup and verify it.
3. **Continuous duration.** Minimum 4–8 hour unbroken runs; the failure modes that matter (thermal drift, battery degradation curves, perception drift as conditions change, error accumulation) do not appear in 10-minute windows.

4. **Defined metrics, agreed in writing beforehand.** Throughput at your task mix; success rate; and above all the intervention economics of Section 5.3 — missions per intervention and minutes of human attention per operating hour.
5. **Your people at the controls.** By the final session, your operators — trained by the vendor within a defined number of hours — should run the machine. If only the vendor's PhD can operate it, you are not buying a product; you are renting their PhD.
6. **Full logs, contractually deliverable.** Raw event logs from the test become an appendix to the purchase agreement, and the acceptance test in Section 8 references them as the baseline the production system must reproduce.

Expect resistance proportional to how much the demo depends on choreography. A vendor with real field maturity will negotiate the details and agree; 65,000 field hours across nine facilities is exactly the kind of operation that can afford to say yes. A vendor who refuses buyer-designed testing outright has answered your diligence question at zero cost to you. **ANALYTIC**

5 The Spec Sheet Decoder

Robotics spec sheets are written in a dialect where every term is technically true and systematically flattering. This section translates the load-bearing terms.

5.1 Terms that require interrogation

Spec-sheet term	What it implies	What it actually means / what to ask
“Up to X picks/hour”	Sustained throughput	Peak burst on ideal items. Ask: sustained rate over 8 hours on my SKU mix, including exception handling and recovery time.
“99% uptime”	Nearly always working	Often power-on availability, not task availability. 99% still = ~87.6 hrs/yr down. Ask: task-level availability, planned + unplanned, measured how, over what fleet.
“Autonomous”	No humans involved	May include remote teleop assist. Ask: interventions per 100 missions, remote-operator staffing ratio, and who pays for the remote operators.
Payload (kg)	It can carry that	Usually at zero reach/speed derating. Ask for the payload-reach-speed envelope curve, not the corner point.
MTBF (hours)	Reliability	Ask the definition of “failure” (hard fault vs. any stoppage), the fleet size and hours behind the number, and MTTR with parts on hand vs. parts shipped.
“AI-powered / learns over time”	Improves autonomously	Ask: what specifically updates, on whose data, validated how, and can an update degrade my performance? (Demand a rollback right — Section 8.)
Cycle time (sec)	Real-world speed	Lab conditions, best-case path. Multiply by 1.3–2.0 for field conditions until proven otherwise on your floor.

Spec-sheet term	What it implies	What it actually means / what to ask
"Safety certified"	Safe around people	Certified to what standard, which configuration, by whom? For arms/cobots ask for conformity to ANSI/A3 R15.06-2025 (or ISO 10218/TS 15066 lineage) in your exact deployment configuration — certification of a base platform ≠ certification of your cell.

5.2 Availability arithmetic

Availability claims compound badly and deserve two minutes of arithmetic every time. A '99% available' machine on two-shift operation (4,000 hrs/yr) is down 40 hours — a full work-week of the labor it displaced. At 97%, 120 hours. Now add the distinction the spec sheet elides: **mean time to repair depends on parts logistics, not engineering.** A 2-hour MTTR with parts on site becomes a 2-week outage when the failed actuator ships from a single overseas facility. Ask where the parts depot is, what the guaranteed parts-dispatch SLA is, and what on-site spares kit the vendor recommends — then price that kit into the deal. **ANALYTIC**

5.3 The metric that matters most: intervention economics

For any mobile or manipulation system, the honest unit of performance is not success rate but **missions per intervention (MPI)** and its cost twin, **minutes of human attention per operating hour.** The labor math is unforgiving: net labor displaced = gross task-hours automated – (supervision hours + exception-handling hours + rework hours + the fractional cost of remote operators). A system that automates 100 hours of work but requires 35 hours of skilled attention has displaced 65 hours — and if the attention must come from your best technician, the effective number is lower still. Demand MPI measured on a fleet, over months, on tasks like yours; a vendor that operates at real customer sites for tens of thousands of hours has this number. Whether they will show you is a diligence result in itself. **ANALYTIC**

6 Balance-Sheet Diligence for Non-Bankers

You do not need a CFA to underwrite vendor survival. You need four questions, the discipline to ask them directly, and a decoder for the answers you get when the vendor is private and coy.

6.1 The four questions

- 1. Runway:** "What is your current cash position plus committed capital, divided by monthly net burn?" The only honest formats are a number of months or a statement of profitability. For a 5-year support commitment, you want either 24+ months of runway with a credible path to the next raise, demonstrated ability to raise in the current window, or revenue durability that makes the question moot.
- 2. Revenue quality:** "Of your announced revenue, how much is recognized, how much is booked-but-contingent, and how much is LOI/pipeline?" The industry norm is to headline the largest number available. Agility's own SPAC filings model the honest version: \$300M+ described explicitly as potential multi-year value tied to one three-year RaaS contract,

subject to milestones, not current-period revenue. If a public filing under securities law hedges that carefully, apply harder discounting to conference-stage claims.

3. **Concentration:** “What share of revenue comes from your top customer? Top three?” Nine facilities is a real fleet; nine facilities where one strategic partner is most of the volume is a different risk. Also ask the mirror question: will I be more than 10% of your revenue? Being a vendor's biggest customer sounds like leverage and is actually exposure — their failure modes become yours.
4. **Strategic investors — signal and noise:** Foxconn leading Agility's PIPE, Amazon in its cap table, NVIDIA and Microsoft in Figure's — strategic capital signals partner conviction and sometimes manufacturing or distribution muscle. But it is not a survival guarantee (Amazon abandoned its \$1.7B iRobot acquisition under regulatory pressure, and strategics routinely write off minority positions). Treat a strategic investor as one point of validation, never as the underwriting.

6.2 Distress signals visible from outside

- Funding-round age: venture companies typically raise every 18–24 months. A Tier 2/3 vendor whose last round is 30+ months old and who is suddenly flexible on price is telling you something.
- Executive churn, especially CFO/COO departures and revolving CEOs. Picnic's terminal phase featured a new CEO and a pivot to a “hospitality-first automated pizza chain” whose flagship pop-up never materialized — pivots announced mid-decline are usually epitaphs.
- Layoffs reported by trade press; delayed conference presence; sales engineers who stop returning calls; a demo facility that has visibly stopped changing.
- Aggressive discounting and unusually long free-pilot offers — startups near the edge buy logos with product because logos help the raise.
- Public-market comparables repricing: when the first audited humanoid trades at 1/15th the top private mark, the private marks will eventually reconcile, and the reconciliation is rarely gentle to the weakest balance sheets in the cohort.

None of these is disqualifying alone; three at once is a pattern. Log them at first contact and re-check at every negotiation milestone — vendor health is a time series, not a snapshot. **ANALYTIC**

6.3 The RaaS special case

REA-006 argued that Robotics-as-a-Service pricing is fundamentally an insurance product: the subscription bundles hardware risk, obsolescence risk, and performance risk into a premium. The corollary for diligence is sharp: **in RaaS, the vendor's balance sheet is your insurance carrier's balance sheet.** You would not buy a policy from an insolvent insurer at any premium. RaaS from a fragile vendor is the worst of both worlds — you avoided the capex but still built your operation atop their survival, and unlike an owned asset, a repossessed or bricked fleet leaves you with nothing to salvage. RaaS therefore demands *more* balance-sheet diligence than purchase, not less, plus the continuity clauses in Section 8. **ANALYTIC**

7 Reference Checks That Actually Work

Vendor-provided references are curated by definition. They are still worth taking — but the value is in how you run them and who else you find.

7.1 Who to talk to

- **Not the VP who championed the purchase.** Champions defend their own decision; it is human nature and useless data. Ask to speak with the maintenance technician who services the machine and the line supervisor who lives with it daily. Their candid candor is the single highest signal input in this entire guide.
- **A customer at 12+ months, not 3.** The honeymoon-phase reference has not yet met the second-year failure modes, the first major software migration, or the renewal pricing.
- **A churned customer.** Ask the vendor directly: “Give me a customer who stopped using the product and permission to call them.” A confident vendor has a de-installation story they can live with (application mismatch, site closure). Refusal to acknowledge any churn is a claim that the product has never failed anywhere, which you should treat accordingly.
- **The integrator’s techs,** if an integrator is involved — they see the vendor’s escalation behavior across many sites and have less stake in your decision.

7.2 The question script

1. “Walk me through the last three times it went down. What broke, how long to diagnose, how long for parts, who fixed it?” (Specifics resist embellishment; vagueness is data.)
2. “How many hours per week does someone touch, babysit, or clean up after this machine?” (This is the field-truth version of Section 5.3’s intervention economics.)
3. “What does the vendor’s support actually look like at 2 a.m. on a Saturday? Have you tested it?”
4. “What did the software update process do to you? Ever had an update make things worse? Could you roll back?”
5. “What did you budget for integration and what did it actually cost — in dollars and in your own engineering hours?”
6. “If you were doing it again: same vendor, different vendor, or no robot?” Then be quiet and let the pause do the work.

7.3 The site visit

One visit to a production deployment outranks every reference call. Protocol: visit a site that has run 12+ months; arrive during a normal shift, not a scheduled showcase; watch a full hour of operation and count interventions yourself; find the whiteboard, the log book, or the CMMS entries — maintenance reality is written down somewhere on every industrial floor; and note the workarounds (cones, tape lines, “don’t stack these two SKUs” signs) because workarounds are the fossil record of failure modes the vendor didn’t mention. If the vendor cannot produce a

single visitable production site running your approximate use case — as opposed to a customer innovation lab — then whatever you are buying, it is not a proven product. **ANALYTIC**

8 Contract Architecture: Converting Risk Into Obligation

Everything diligence uncovers has exactly one place to go: the contract. The clauses below are the buyer's standard armor; for Tier 2–3 vendors, treat the first five as walk-away items. Vendors will resist several — resistance is information about which risks they know they carry. (Standard disclaimer: this is a procurement framework, not legal advice; run final language past counsel.)

Clause	What it does	Negotiation notes
Software escrow with release triggers	Source code, build environment, calibration data, and update tooling held by a third-party escrow agent; released on insolvency, discontinuation of support, or failure to meet SLA cure periods.	K-Scale open-sourced its IP after death — an accidental gift to its ecosystem. Escrow makes that outcome contractual rather than charitable. Verify deposits are updated each release; stale escrow is theater.
Spare-parts & wind-down guarantee	Guaranteed parts availability for N years (target 7 for Tier 1–2 hardware); on end-of-life, a last-time-buy window plus transfer of supplier list and drawings for wear parts.	The Picnic customer's \$250K aquarium is the alternative. For fragile vendors, negotiate a consigned on-site spares kit purchased up front.
Performance-based acceptance	Payment milestones gated on reproducing the Section 4 counter-demo metrics on your floor: throughput at your mix, MPI floor, task-availability floor, over a 30-day continuous acceptance run.	Holdback 20–30% until acceptance. Define the metrics, the measurement method, and the logging source in an appendix — disputes die in ambiguity.
SLA with teeth	Task-level availability commitment with escalating remedies: service credits → extended term → termination for cause with refund schedule if availability persists below floor.	Credits alone are a rounding error to the vendor; the termination-with-refund tier is what changes their staffing decisions.
Update governance & rollback	No forced updates during your peak season; buyer right to defer non-security updates; guaranteed rollback to last-known-good; vendor liability if an update degrades agreed metrics.	“The model learns over time” must not mean “your throughput is a random walk.”
Data ownership & portability	You own operational data generated on your floor; vendor use for model training requires consent; on exit, data export in documented format.	Your exception logs are the training curriculum for their next model — that has value; price it or restrict it.

Clause	What it does	Negotiation notes
Assignment & change-of-control	Support obligations survive acquisition; you gain termination or re-negotiation rights if the acquirer sunsets the product line.	“Acquired” is how robotics products most often die. This clause is the difference between an acquirer honoring your SLA and mailing you a deprecation notice.
RaaS continuity terms	For subscriptions: fleet continues operating N months post-vendor-insolvency (escrowed operation runbooks, local-mode operation without cloud), plus purchase option at pre-agreed schedule.	Per Section 6.3: in RaaS the vendor is your insurer. These clauses are the guaranty fund.
Insurance & liability	Vendor carries product liability at specified limits; indemnity for IP infringement (increasingly relevant as vision/manipulation patent suits spread); clear allocation for injury scenarios in collaborative deployments.	Ask for certificates, not assurances. Verify the safety certification covers your configuration (Section 5.1).
Price & renewal protection	Caps on support/subscription escalation (CPI + margin); renewal terms locked at signature; consumables and EOAT pricing schedules attached.	The list-price robot with uncapped support pricing is a loss-leader with a tail.

A note on leverage: the first-time buyer often assumes they have none against a hot vendor. In 2026 the opposite is frequently true. Frontier vendors need referenceable production deployments more than they need any single check — deployment hours are the currency their next raise is priced in. A buyer offering a serious production site, honest data sharing, and a public case study is providing something scarce, and should charge for it in contract terms.

ANALYTIC

9 The Integrator Question: Who Owns Success?

For Tier 1 equipment, the practical vendor is often not the OEM but the systems integrator who designs the cell, writes the application logic, and holds the service relationship. This changes the diligence target.

- **Single throat to choke.** Demand one contractually accountable party for end-to-end performance. The classic failure mode of integrator-led projects is the triangle blame cycle: OEM blames integration, integrator blames the OEM, both blame your operations. The acceptance-test clause must bind whoever holds the paper.
- **Diligence the integrator like a vendor.** Years in business, engineer headcount and tenure, certified partnership tier with the OEM, references on projects of your scope — integrators die too, and a dead integrator strands your application code the way a dead OEM strands parts. Application source code and cell documentation belong in your escrow terms.

- **Beware the vendor who is also the integrator on their first ten sites.** Frontier vendors integrating their own early deployments are simultaneously learning the product and the trade. Their engineers may be brilliant; the question is whether deployment #11 — yours — gets the A-team or the newly hired B-team. Ask who, by name and résumé, will be on your floor.
- **Open ecosystems compound in your favor.** The Rethink/Universal Robots lesson generalizes: platforms with broad third-party ecosystems (grippers, vision, fleet software, independent service shops) give you exit ramps and second opinions. Every proprietary interface is a future toll booth.

10 The Scoring Framework

Diligence outputs must compress into a decision. The AFR vendor scorecard weights five dimensions; scores are 1–5 per line item, weighted, summed to 100. The weights encode this paper's core argument — half the decision is survival and proof, not features.

10.1 Weighted scorecard

Dimension	Wt.	What earns a 5	What earns a 1
Vendor viability	25	Public/profitable or 24m+ runway disclosed; diversified revenue; stable leadership; strategic backing as bonus	Pre-revenue; 30m+ since last raise; exec churn; discounting to close
Demonstrated field performance	25	10,000's of fleet hours at multiple third-party production sites; MPI and availability data shared; buyer-designed test passed	Video demos only; refuses counter-demo; no visitable production site
Economics (TCO, per REA-004)	20	NPV positive on supported life with conservative intervention costs; consumables and support priced and capped	Payback requires spec-sheet throughput and nameplate life
Contract protections	15	Escrow, parts guarantee, performance acceptance, SLA-with-teeth, change-of-control all agreed	Vendor paper, as-is, arbitration in their county
Support & ecosystem	15	Regional parts depot; 24/7 support tested; open interfaces; third-party service exists; integrator bench	Single overseas depot; proprietary everything; founder answers support email

Decision bands: 80+ proceed to negotiation; 65–79 proceed only with compensating contract terms on the weak dimensions; 50–64 structured pilot only, no production dependency; below 50, decline and calendar a re-look in 12 months. Calibration note: under these weights, a hypothetical vendor with dazzling technology, no field fleet, and a burn-heavy balance sheet cannot score above the pilot band no matter how good the demo — which is precisely the point.

ANALYTIC

10.2 Red-flag taxonomy

Class	Flags	Treatment
Disqualifying	Refuses buyer-designed testing; refuses escrow AND parts guarantee; no visitable production customer; safety certification doesn't cover your configuration; evasive on autonomy ratio in writing	Walk. No price fixes these.
Heavily discountable	Booked/pipeline revenue presented as recognized; last raise 30m+ old; single-customer concentration; support depot single-sourced overseas; demo-culture answers to the end-stop question (§3.4)	Proceed only with maximum contract armor and holdbacks; assume mid-contract exit (§3.5).
Discussable	Young company with strong strategies; closed ecosystem with strong roadmap; premium pricing vs. Tier 1 alternative; integrator doubling as OEM on early sites	Negotiate; verify; weight the scorecard honestly rather than rationalizing.

The Verdict: Three Buyer Scenarios

Per house practice, the paper closes with the three-scenario table. The scenarios are buyer postures, not market forecasts; each is rational for a different risk budget.

	Scenario A — The Conservative Core	Scenario B — The Structured Adopter	Scenario C — The Frontier Bettor
Posture	Buy Tier 1 only: arms, cobots, scaled AMRs from established OEMs via vetted integrators	Tier 1 for production; one Tier 2/3 structured pilot under full Section 8 armor	Early production commitment to a frontier vendor for strategic learning
Rationale	Category is measurable and growing (+6.6% units 2025); vendor mortality $\approx$ solved; ROI math is real today	Captures learning-curve position while contract terms and holdbacks cap downside; pilots priced as tuition, not ROI	If general-purpose robotics compounds, early operational capability is the moat; deployment-hour scarcity gives buyers leverage now
Key risk	Ceding learning-curve position if frontier categories mature faster than expected	Pilot purgatory (REA-005); organizational fatigue; pilot costs treated as ROI failures	Vendor mortality at 35–50% over 5 yrs; stranded assets; the \$250K aquarium at industrial scale
Who should choose it	Most first-time buyers; anyone whose downtime cost is existential	Operations with in-house engineering, exec patience, and a real exception-heavy use case	Only buyers who can afford total loss of the program and are buying knowledge, not labor
Verdict weight	The default. $\approx$ 70% of first-time buyers belong here.	The recommended ceiling of ambition for $\approx$ 25–30%.	Legitimate for a few percent; a hobby with an invoice for everyone else.

Bottom line. The skeptic's edge in 2026 is not cynicism — the established market is genuinely strong, and even the frontier now contains at least one vendor with 65,000 audited-adjacent hours and public-market accountability arriving. The edge is procedural: insist on your test, your metrics, your contract, and the vendor's real numbers. Vendors built to survive audit will meet you there. The rest were never selling you a robot; they were selling you a demo, and demos make poor capital assets. **ANALYTIC**

Epistemic Ledger

Key load-bearing claims and their register, per AFR practice:

Claim	Tag	Basis
2025 NA orders 36,766 units / \$2.25B (+6.6% / +10.1%); Q1 2026 cobots +55.6%	GROUNDED	A3 quarterly and annual statistics, 2025–2026
Humanoid funding \$8.1B+/100+ cos.; Figure ~\$39B; Appteronik \$935M@\$5B+; Agility SPAC \$2.5B, \$300M booked (milestone-contingent), 65,000 hrs/9 sites	GROUNDED	SPAC transaction filings and contemporaneous financial press, 2025–2026
Rethink (2018), Picnic (2026, \$250K stranded customer), K-Scale (2025) failure narratives	GROUNDED	Trade-press post-mortems incl. first-person COO account
Sector capitalized ~40× present-year humanoid revenue pool; 15:1 top-private vs. first-audited spread	ANALYTIC	Arithmetic on the grounded figures above
5-yr vendor survival priors: Tier 1 ~95%+, Tier 2 ~60–75%, Tier 3 ~35–50%	PROJECTION	Base rates (Carta shutdown series, China polarization data) adjusted by segment structure; wide error bars acknowledged
Humanoid market \$1.4B (2026) → \$19.6B (2033) at 30.4% CAGR	PROJECTION	Third-party market research cited in SPAC coverage; treat as scenario input, not fact
Buyer leverage from deployment-hour scarcity persists through the frontier's public-market reconciliation	SPECULATIVE	Structured conjecture from funding-cycle dynamics; monitor as Tier 3 vendors list

Sources & Series Note

- Association for Advancing Automation (A3), North American robotics market statistics: annual 2025 report and quarterly releases Q1 2025–Q1 2026 (automate.org; therobotreport.com; mmh.com; designnews.com).
- Agility Robotics / Churchill Capital Corp XI SPAC transaction: Forbes (Jun 2026), ION Analytics/Mergermarket, TechCrunch (Jul 2026), Tech Funding News; figures as characterized in transaction filings.
- Humanoid funding landscape: Humanoid Index funding tracker; New Market Pitch funding analysis (Jul 2026); NBC News (Figure AI rounds); TechCrunch (Appteronik, AI2 Robotics).
- Vendor failures: The Robot Report — R. Xu, “6 lessons I learned watching a robotics startup die from the inside” (2026); TheStreet/GeekWire on Picnic Technologies (May 2026); RoboticsBiz on Rethink Robotics; 36Kr on China’s 2025 humanoid shutdown wave; TechCrunch/Carta shutdown data (Jan 2025).

- Series cross-references: REA·004 “The Humanoid as Capital Asset” (valuation framework); REA·005 “The Pilot Trap” (deployment stall patterns); REA·006 “Own, Lease, or Rent the Robot” (RaaS-as-insurance thesis). All at zspacedynamics.com.

— *End of REA·008* —