Applied AI LLC

FOUNDER STRATEGY / NOVEMBER 2026

Building AI Startups

A Guide for Founders

Turning continual learning AI and robotics into a business customers can rely on

Vasyl Harasymiv

Founder of Applied AI LLC

This practical leadership paper shows founders how to turn adaptive AI breakthroughs into viable companies by connecting research decisions, the founding team and capital requirements with proven performance, funded milestones and the ability to support the first customer and scale repeatable deployments without making the founder the permanent fallback.

Q4 2026

STRATEGY AND EXECUTION

Executive summary

This paper explains how to turn continually learning AI into a company that can deliver. The central argument is that scientific evidence, capital commitments and customer obligations must inform the same decisions. The leadership task is to make that connection dependable while preserving the team's ability to explore.

Vasyl Harasymiv, Founder of Applied AI LLC in Chicago, combines hands-on AI leadership with financial-services and capital-markets experience. His background includes an executive director role at JPMorgan, senior data science work at Grubhub, and AI advisory work across strategy, analytics, governance and implementation. [9] His published insights connect technical evaluation with customer understanding. This experience equips him to advise AI startups, fintech ventures and high-growth firms with authority on the decisions that shape a new company: what to build, how to organize delivery and where to commit capital.

Black Book Insights (London UK/Tampa FL USA) assisted in polling 150 AI industry stakeholders in October 2026, working with Applied AI LLC. Ten measures of operating practices and buying conditions inform the analysis; published research supplies technical and market context at www.blackbookinsights.com

What the findings mean

Retained performance after adaptation leads the buying conditions at 90%, followed by defined human intervention and restoration at 79% and documented operating cost at 53%. The commercial priority is useful learning that customers can continue to rely on, control and afford.

Fewer than half report full implementation of each operating practice: written acceptance criteria at 45%, a named deployment owner at 40%, milestone-linked budgeting at 36%, human-support costing at 16%, and tested containment and restoration at 11%. These findings direct attention to the work required to sustain a customer commitment.

Strategic priorities

Priority	Immediate management decision
Prove learning over time	Fund repeated adaptation and retention tests under conditions the customer recognizes.
Make recovery deliverable	Give monitoring, intervention and restoration an owner, acceptance criteria and a budget.
Expose deployment economics	Price supervision, failed attempts, integration and recovery into the cost of useful work.
Protect founder capacity	Assign an owner to each consequential commitment and require acceptance before expansion.

The first ninety days

Days 1–30: Agree the next research and customer decision, its evidence requirements and its owner. Establish the milestone budget, monthly cash forecast and baseline measures. **Days 31–60:** Run a bounded partner evaluation, test recovery and account for support effort. Reconcile the roadmap, fundraising case and forecast with the results. **Days 61–90**: Complete a material commitment through acceptance. Decide what to fund next and give investors or the board a clear account of the evidence and required action.

01 FOUNDER DECISIONS

Build the company around the next proof

An AI founder has to build a business while important parts of the technology remain uncertain. A strong demonstration may attract attention, but the next commitments require more: a customer problem worth solving, a test that can change the plan, and enough cash and team capacity to act on the result.

I argue for an approach that keeps research ambitious and commitments specific. The team should know what it needs to learn, which decision the result will change and what the company can afford to do next. Leadership becomes especially consequential when a system keeps learning after deployment: each technical advance can alter support, staffing and commercial obligations.

Choose a starting point a small team can own

A startup should use the wider AI market to identify an opening it can serve. The external indicators below describe the research and deployment setting. The fieldwork adds a more immediate commercial signal: 90% identify retained performance after adaptation as a buying condition. A first market should make that capability valuable and measurable.

Published signal	Implication for the founder
Over 90 percent of notable frontier models in 2025 came from industry, according to Stanford's 2026 AI Index. [1]	Define a capability the startup can establish with its own resources and explain why that focus matters.
IFR reports 603,000 industrial robot installations in 2025 and an operational stock of 5.079 million. [2]	Choose a reachable customer and use case; integration and economic value still have to be demonstrated.
A 2026 NIST report identifies unresolved practical challenges in monitoring deployed AI. [7]	Put evaluation and monitoring into the product, staffing plan and cash forecast from the beginning.

Match each stage to a concrete decision

Before a first pilot, resolve whether the learning advantage matters to a specific buyer and can be tested in realistic conditions. During the pilot, establish the cost of integration, supervision and recovery. Before expansion, test whether another customer or site can obtain a similar result without relying on exceptional founder involvement.

Choose the next milestone around the uncertainty most likely to change the company's direction. With milestone-linked budgeting reported at 36%, make the cost of that decision explicit. Agree an owner, acceptance criteria and the response to an unfavorable result. Keep the planning process light enough to update when evidence changes.

02 FOUNDER DECISIONS

Understand what your AI must keep proving

Continual learning aims to acquire useful knowledge over time while retaining earlier capabilities. In reinforcement learning, this includes maintaining decision-making competence as tasks and environments change. A recent survey treats adaptation, knowledge retention and resource constraints as connected research problems. [4]

Distinguish continual learning from adjacent capabilities. Retrieval can change an answer without changing the model; longer context can support adaptation within a session; periodic retraining updates a system between releases. Retained learning during operation creates a different evaluation problem because relevant state and history can change between similar tasks.

Make retained performance part of the product promise

Dohare and colleagues demonstrated loss of plasticity, or diminished ability to learn new information, in the deep continual learning settings they studied. That problem differs from forgetting previously learned information. Their work supports treating both as evaluation questions; it does not establish that every architecture will behave the same way. [3]

Evidence dimension	Question the founding team should answer
Acquisition	What new capability was learned, from how much experience, against which baseline?
Retention	Which earlier capabilities were retested after learning, and where did performance deteriorate?
Continued adaptability	Does the system still learn effectively after extended operation?
Resource growth	How do memory, compute, energy and response time change as experience accumulates?
Transfer and coordination	Which results carry to a new platform, task or group of agents, and which require additional work?

Use these dimensions to organize the evidence behind the 90% retained-performance requirement. Technical teams should define task-specific measures and report test conditions, uncertainty and unsuccessful trials alongside headline results. Retesting earlier skills should be part of each material learning evaluation, with agreed limits on acceptable regression.

Make the technical advantage specific

A technical advantage earns commercial weight when its effect can be measured. Test whether a novel architecture improves learning, resource use or diagnosis under relevant conditions. An explanation of how a system works does not validate an explanation of an individual decision; a lower compute requirement in one experiment does not establish a lower cost of service.

Connect each technical claim to a testable customer benefit. Compare learning speed, retained skills and resource growth across repeated changes. Set a research objective and spending limit for unproven advantages. Reuse the evaluation record for product choices, partner discussions and financing.

03 FOUNDER DECISIONS

Win a first customer you can learn from

Start with a recurring problem, a buyer who can define success and access to the conditions needed to test it. A narrow first use case gives the founding team a way to connect a technical advantage with a purchase decision. Interview the operating users, the budget owner and the person who will accept the result before committing to a pilot.

Choose the deployment form that fits that problem. A humanoid may suit an environment designed for people; a purpose-built machine may offer a simpler route to reliability and cost control. Test where adaptation creates value before taking on the hardware, facilities and integration costs of a particular product strategy.

Define the value of adaptation

Consider an industrial handling workflow in which objects or layouts change frequently. Agree on whether adaptation should reduce reconfiguration effort while preserving throughput, quality and an operating boundary. Compare technician time, failed attempts, downtime and recovery costs. Those measures help the founder explain why a buyer should pay for the product.

Stage	Evidence to establish	Decision enabled
Scientific validation	Repeatable advantage on a defined task, relevant baseline and retained prior skills	Continue, redirect or stop the research line
Engineering validation	Usable interfaces, measured latency, resource limits and recovery behavior	Commit to a bounded integration
Partner evaluation	Performance in representative conditions and an agreed cost comparison	Offer a supported pilot
Repeatable deployment	Similar outcomes at another site with manageable support effort	Expand the product and commercial commitments

The stages can overlap, but the evidence should remain distinct. A paid pilot establishes willingness to fund an evaluation. It does not by itself establish repeatable product demand. Nor should a strong benchmark result automatically trigger a sales expansion.

Set boundaries around the design partnership

A design partner should offer representative tasks, evaluators and a route to a buying decision. The 53% operating-cost requirement makes access to real supervision, downtime and integration costs especially valuable. Agree who funds the work, what the pilot includes, who owns support and what learning may be retained or transferred. Set an end date and decision point.

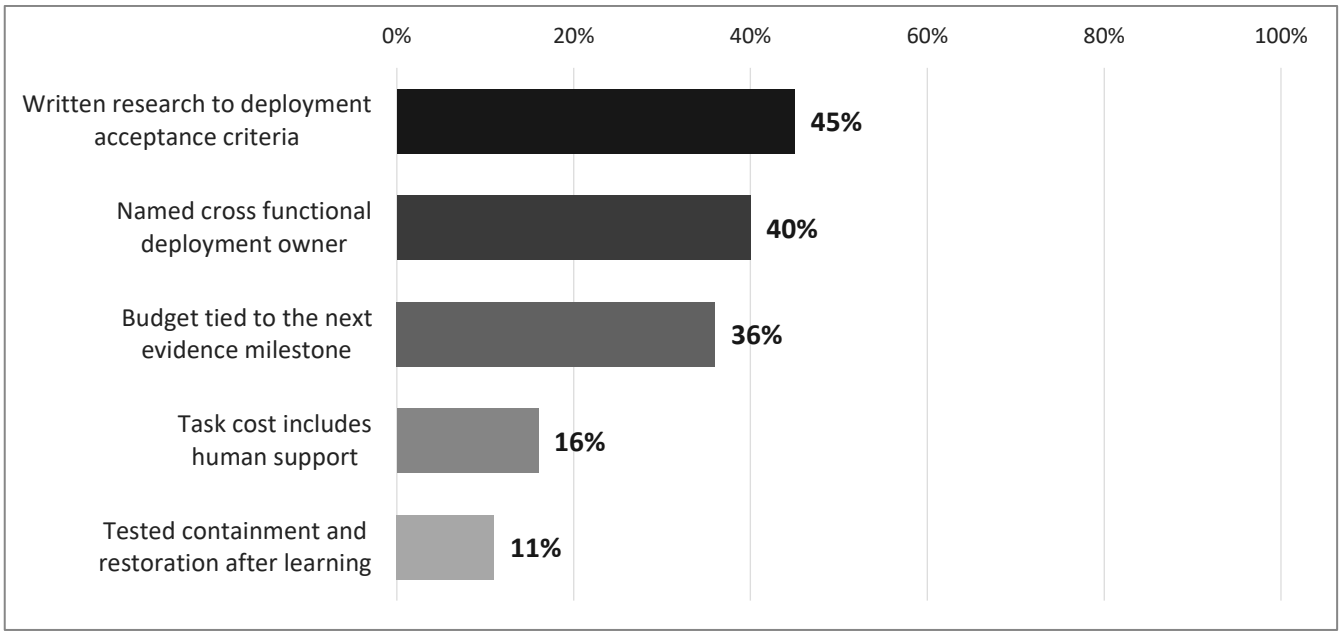
Keep interfaces and deployment records consistent enough to evaluate a second platform. Google DeepMind has reported transfer of selected learned tasks across robot embodiments. [5] This supports cross-platform capability as a research direction, without establishing unrestricted transfer or continual learning during deployment. Build a reusable learning component where evidence justifies it; select hardware according to the first task and its operating constraints.

BLACK BOOK INSIGHTS FIELDWORK

Operating practices need to reach deployment

The five reported operating practices range from 45% for written acceptance criteria to 11% for tested containment and restoration after learning. Formalizing the initial commitment is more common than demonstrating how the team will support and recover the deployed system.

Figure 1 Reported fully implemented operating practices



Source: Black Book Insights, AI Startup Leadership and Deployment Readiness Study, October 2026. Reported study cohort: 150 AI industry stakeholders. See the methodology for interpretation and percentage bases.

The founder should close the operational loop

Written acceptance criteria at 45% and a named deployment owner at 40% provide a starting point. The lower rates for milestone budgeting (36%), human-support costing (16%) and tested recovery (11%) show where the plan needs concrete deliverables. A pilot should produce an accepted result, a cost record and a tested response to deterioration.

These measures describe different practices and do not track a single group through successive stages. They nevertheless identify useful priorities for management review: connect each product promise to the people, cash and operating controls required to sustain it.

04 FOUNDER DECISIONS

Build a founding team that can make decisions

A founder becomes a constraint when routine decisions wait for the same person. With 40% reporting a deployment owner, responsibility for the whole customer outcome needs attention. Define which judgments stay with the founder, which move to a team lead and which require a specialist.

Choose people who improve the quality of judgment

Hire against the next difficult obligation. A researcher must distinguish a repeatable result from a favorable trial; a delivery lead must carry an integration through acceptance. Operational leadership requires the judgment to challenge an assumption, model its financial consequences and explain the decision clearly. Assess these abilities through completed work, not seniority alone.

Invite disagreement before a commitment hardens. Ask for the strongest objection and the evidence that would change the recommendation. Keep sensitive research, investor and personnel matters confidential while making setbacks visible to the people who must respond. Trust depends on discretion and the willingness to surface an inconvenient fact early.

Give each decision an owner

Let technical leads choose experimental methods within agreed objectives and resources. Give a product owner responsibility for customer requirements and acceptance. Reserve changes in direction or material commitments for the founder. Qualified safety leaders need authority to stop or escalate physical deployment work.

One person may hold several responsibilities, but each decision needs an owner. Record the evidence, choice and follow-up. Reallocate responsibility when delays reveal overload; additional approval steps should resolve a specific problem.

Use short written decisions to reduce repetition

A short decision note should state the recommendation, evidence, strongest objection, cost of delay and next action. Record who can decide and who will carry the choice through to completion. This gives the founder a basis for judgment without requiring every discussion to be reconstructed, and preserves technical qualifications as results become commercial commitments.

Prepare external communications in the founder’s voice while retaining the evidence behind each claim. Technical owners should verify scientific language; the financial model should support the promised timetable and use of funds. One consistent account of progress should serve the team, customers and investors at the level of detail each needs.

Protect the work only the founder can do

Reserve founder time for direction, key people, capital and consequential customer commitments. A trusted operating partner should bring options, a recommendation and a route to completion, then reconcile the resulting roadmap, resources and dependencies. Establish authority to resolve routine conflicts so that coordination returns capacity to the founder instead of creating another approval queue.

Judge delegation by decision waiting time, rework and accepted commitments. The founder should gain capacity while retaining a clear view of execution.

05 FOUNDER DECISIONS

Keep a small team moving toward accepted work

A company plan should change how people spend their next week. The 45% acceptance-criteria result makes the first task clear: define finished work before increasing commitments. Translate a few quarterly outcomes into weekly dependencies and decisions. Keep research assumptions visible so that a changed result leads to a deliberate revision of scope, spending or timing.

Separate experiments from delivery commitments

An experiment can succeed as a learning exercise even when its hypothesis fails. A delivery commitment succeeds when an agreed result is accepted by its intended recipient. Mixing the two encourages optimistic schedules and makes it difficult to understand why work slipped. Plans should show which dependencies are scientific unknowns and which are execution obligations.

Review	Primary question	Required output
Weekly delivery review	What is blocked, and which decision would unblock it?	Named action, decision owner and next checkpoint
Research evidence review	What did the latest results change about our beliefs?	Updated evidence, uncertainty and next test
Monthly company review	Are spending and commitments consistent with the strategy?	Revised forecast and explicit priority changes
Milestone review	Is the evidence sufficient for the next commitment?	Recorded approval, restriction, redirection or stop

Use one set of records for these reviews. A slipped experiment, delayed hire or new customer requirement should update both the delivery plan and forecast. Keep meetings that produce decisions or resolve dependencies. Measure accepted work rather than the volume of status reporting.

Own strategic projects through acceptance

A fundraising process, platform evaluation or deployment can fail between functions even when specialists finish their tasks. Name one owner for the whole outcome. Define acceptance criteria, dependencies and resource limits. Escalate conflicts in priorities without making the founder the default project manager.

Completion requires acceptance and a transfer of ongoing responsibility. A pilot needs its evaluation record, operating instructions, escalation owner and cost assumptions. A financing process needs consistent materials and resolved diligence questions. Reuse those records across reviews so reporting does not become a separate job.

Reward the disclosure of bad news

Make unfavorable results safe to report early. Stop an unproductive experiment when the evidence warrants it. Ask whether it answered the intended question, who needs the result and what changes next. Concealed uncertainty creates avoidable commitments.

06 FOUNDER DECISIONS

Use runway to reach the next decisive milestone

A startup budget should show which uncertainties the company can afford to resolve. The 36% milestone-budgeting result suggests that this connection needs explicit attention. Link headcount, compute and equipment spending to the next consequential test. Model delays in evidence, customer evaluation and financing before cash pressure narrows the company's choices.

Build the forecast from the work

Model hiring by start date and total employment cost. Connect experimentation to expected compute usage, simulation capacity and specialist time. For embodied systems, include equipment, replacement parts, facilities, test access, field engineering and travel. Put monitoring, customer support and incident response into the deployment economics rather than treating them as future overhead.

Separate research exploration, reusable product engineering and customer-specific work. A paid pilot can consume the people needed to build the wider product. Include that tradeoff in the forecast. Promotional credits or borrowed equipment can extend near-term runway without establishing the eventual cost of the service.

Plan for decisions under different outcomes

Planning case	Assumption to test	Management response
Expected progress	Evidence and integration arrive near the planned dates	Fund the agreed next milestone
Scientific or integration delay	A critical result requires another experimental cycle	Protect essential work and defer optional commitments
Stronger evidence or demand	Results support a larger opportunity earlier than planned	Check delivery capacity before accelerating hiring or sales
Capital arrives later	A financing process takes longer than the operating plan assumes	Rephase spending before cash pressure forces rushed decisions

Show cash by month and compare actual spending with the assumptions that drove it. Decide in advance which evidence permits additional hiring or equipment purchases and which commitments can be deferred. Review timing differences in customer receipts and financing alongside the cost of another test cycle. A forecast should make the next choice visible before cash pressure makes it urgent.

Measure the cost of useful work

Only 16% report task costing that includes human support, while 53% identify documented operating cost as a buying condition. For a robot, track total cost per accepted task, including supervision, failed attempts and recovery. For a learning platform, track the cost of adding a validated capability and retaining it through later updates. Use the measures appropriate to the application.

Preserve room to change course. Count monitoring, integration and support in the milestone budget, with an allowance for another test cycle. Distinguish spending that builds reusable capability from customer-specific work. The purpose is to fund a decision the team can act on, with enough capacity left to support the result.

07 FOUNDER DECISIONS

Raise capital on a claim you can substantiate

A founder must explain both what the startup has demonstrated and why the opportunity could become larger. Keep that distinction consistent across the pitch, financial model, customer discussions and diligence materials. A clear account of uncertainty lets investors assess the plan without turning every ambition into a promise.

Explain what the next round will prove

Connect the customer problem, technical advantage, strongest evidence and next milestone to the use of funds. The 90% retained-performance and 79% intervention-and-restoration findings support a case for evaluation and deployment control. Explain the output that each investment will produce, its accountable owner and the result that would change the plan.

Describe the first commercial route separately from the broader opportunity. Estimate the customers and deployments the startup could realistically serve through its chosen product and delivery model. The number of robots operating globally, for example, provides context but does not automatically become the company's addressable revenue.

Maintain one record of material claims

For each material capability or economic claim, retain its approved wording, source, test conditions, responsible technical or financial owner and last review date. Indicate whether it is demonstrated, under evaluation or an ambition. This lightweight discipline helps prevent an experimental result from becoming an unsupported promise as it is shortened for different audiences.

Investor materials should explain uncertainty without burying it in appendices. If a result depends on a particular task set, hardware configuration or amount of human support, the narrative should say so. If a deployment opportunity depends on a partner's future decision, distinguish the relationship from a committed customer contract.

Make board updates useful for decisions

Open an investor or board update with the decision that matters and what changed. Show the evidence, consequences for the roadmap, spending against plan, cash outlook and requested action. Distinguish matters requiring a decision from progress reported for information. Describe setbacks precisely enough for supporters to help without rebuilding the underlying analysis.

Maintain a diligence repository containing reproducible results, the operating model, key agreements and current risks. Keep the pitch, projections and board materials consistent with that record. An accountable owner should resolve open questions and record commitments as ordinary work proceeds, protecting research time when an investor or customer deadline arrives.

Communicate with judgment

Founders need to change the level of detail without changing the underlying claim. Researchers need experimental precision, buyers need operating consequences and investors need a credible route to value. Clear writing helps the team make commitments it can meet and protects time otherwise spent correcting avoidable misunderstandings.

08 FOUNDER DECISIONS

Build support and recovery into the product

Defined intervention and restoration is a buying condition at 79%; tested containment and restoration is implemented at 11%. The measures answer different questions, but both make recovery evidence a leadership priority. NIST identifies practical gaps in monitoring deployed AI. [7] Fund evaluation and recovery as part of the product, with named owners and acceptance criteria.

Define what may change and who can intervene

Before a bounded deployment, define what the system may learn, which behaviors remain constrained and what evidence is required before learning reaches a wider deployment. Assign responsibility for detecting deterioration, investigating regressions and deciding whether to continue. Give unfamiliar conditions an explicit escalation path instead of assuming the system can always adapt successfully.

Record the learning state, history and operating conditions needed to investigate an outcome. A version number may be insufficient when instances accumulate different experience. Test how an instance is isolated, constrained or restored. For physical systems, qualified engineering and safety teams must establish controls appropriate to the equipment and hazards.

Evaluate the system as well as the agent

Multi-agent research examines how agents cooperate to complete shared tasks and the difficulties that arise across different decision-making methods. [6] Test coordinated systems against the joint task, including changing team composition, delayed communication and conflicting local objectives. An individually strong agent does not establish dependable performance by the group.

Before learning moves between machines, require acceptance under the receiving platform’s conditions. Record differences in environment and prior experience that may affect performance. Transfer is a capability to test, with conditions attached.

Count human work accurately

Anthropic’s study of its coding product and API activity found that more experienced coding users allowed more automatic actions while also interrupting agents more often. The study describes software usage, not robotic deployment, but illustrates why autonomy and oversight should be measured together. [8]

For an industrial pilot, report who watches the system, who intervenes, how often intervention is required and how long recovery takes. Increased autonomy is commercially useful only if the resulting performance and oversight burden meet the customer’s requirements.

Use failures to improve the operating model

A retention failure can change the pilot scope, evaluation budget and support assumptions. The responsible lead should bring the founder a recommendation, revise the cash forecast and customer timetable, and carry agreed changes into delivery. Follow-through turns technical learning into management information.

Demonstrate detection, intervention and restoration with the customer’s operating team where appropriate. Establish the people, time and cost needed for that response before increasing commitments.

09 FOUNDER DECISIONS

Understand why an enterprise will trust a startup

Research contribution by Black Book Insights, Tampa, Florida, and London, United Kingdom.

The fieldwork places retained performance after adaptation at 90% and defined human intervention and restoration at 79%. These priorities make the buyer's central concern concrete: can a learning system remain useful as it changes, and can the organization manage an unfavorable result?

Sell a dependable operating result

A learning claim becomes commercially relevant when the buyer can connect it to a recurring problem and an accountable supplier. Faster changeovers, less specialist reprogramming or fewer manual exceptions may justify adoption. Demonstrate the outcome under the buyer's actual constraints, then show the conditions under which that outcome holds.

Make retained capability visible

Agree on the variation that matters to the buyer. Measure new learning alongside earlier capabilities and show what happens after repeated changes. Use the same definitions in demonstrations, acceptance tests and operating reports. This gives the customer a way to evaluate the 90% priority throughout the relationship, rather than relying on a single successful demonstration.

Define the response to deterioration

For the 79% intervention-and-restoration condition, explain who watches the system, what triggers intervention and which operating state can be recovered. Include the customer's responsibilities and any limits on remote support. Where learning is shared across machines or sites, a local result should pass an acceptance process before wider use.

Make the economics reviewable

The 53% documented-cost result reinforces the need for an operating model that includes customer staff, field engineers and remote supervisors. Record their time as well as compute, hardware, integration and recovery costs. Price the support obligation with the customer, and distinguish reusable product work from services needed for one installation.

Prepare for the next site before selling it

A named integration and delivery owner is a reported condition at 37%; reproducible second-site performance is reported at 31%. These lower rates do not make either discipline optional. A first evaluation and a broader rollout can have different requirements. Founders should establish the owner, evidence and support capacity needed for the commitment actually being sold.

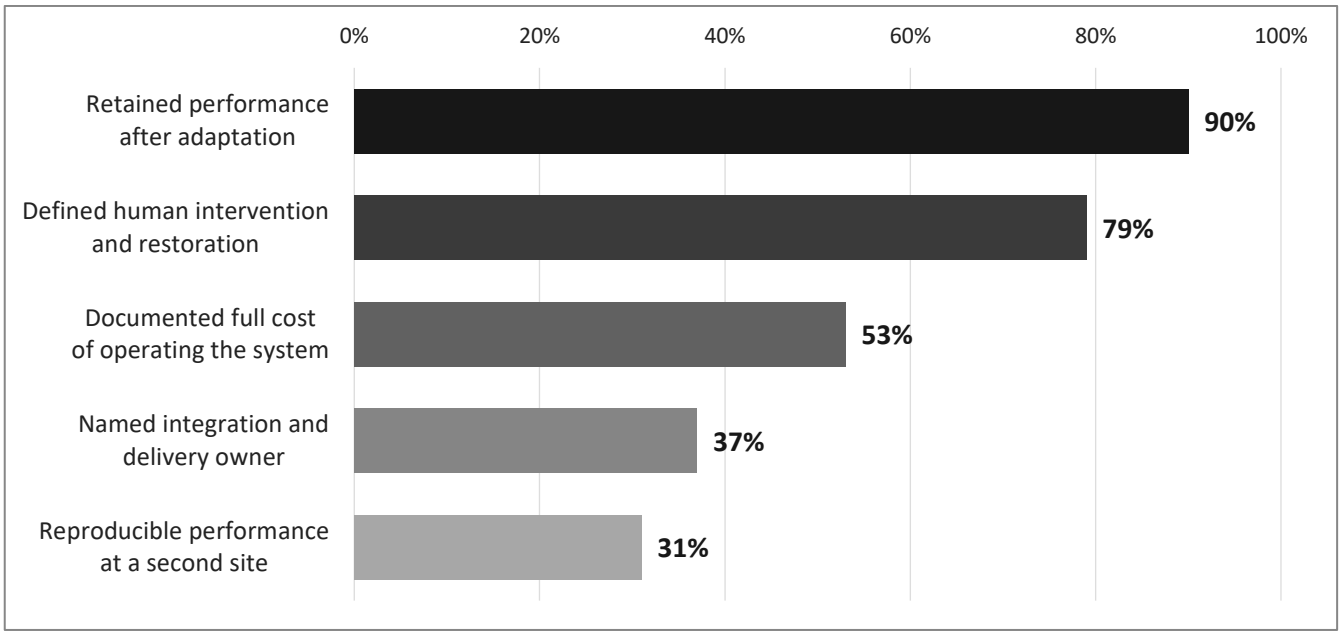
A useful first customer relationship produces a shared record of performance, exceptions and costs. Use that record to revise the roadmap and pricing. Expansion becomes more credible when another team can obtain a comparable result without exceptional access to the founder or a proportional increase in specialist support.

BLACK BOOK INSIGHTS FIELDWORK

Buyers prioritize retained performance and recovery

Retained performance after adaptation leads the reported buying conditions at 90%. Defined human intervention and restoration follows at 79%, ahead of documented operating cost at 53%. The ordering supports a product plan that makes learning dependable and economically visible.

Figure 2 Reported mandatory buying conditions



Source: Black Book Insights, AI Startup Leadership and Deployment Readiness Study, October 2026. Reported study cohort: 150 AI industry stakeholders. See the methodology for interpretation and percentage bases.

Translate the ordering into delivery work

Lead the evaluation package with evidence of retained skills and a tested intervention-and-restoration procedure. Add full operating costs before a commercial commitment expands. Name the integration owner and define the evidence needed at the next site, even when a buyer’s immediate decision concerns only the first installation.

A lower mandatory-condition rate does not measure lack of interest. It can reflect the decision stage or the scope of the initiative. The results guide the sequence of proof; they do not justify removing delivery ownership or repeatability from the company plan.

10 FOUNDER DECISIONS

Set a ninety day agenda for the founding team

Use the next ninety days to make direction, evidence and obligations visible. Adapt this agenda to the company's stage, available cash and next deployment commitment.

Days one through thirty

Agree with cofounders on the customer problem, technical claim and next consequential test. Establish baseline performance and retained-skill measures. Assign owners for the five practices in Figure 1. Complete a strategy brief, decision register, milestone budget and monthly cash forecast. State which decisions require the founder and which the team can make within agreed limits.

Days thirty one through sixty

Run a bounded partner evaluation with acceptance criteria and spending limits. Test intervention and restoration, and record human-support work. Resolve the interfaces, test access and delivery responsibilities needed for the pilot. Reconcile the product roadmap, financing narrative and forecast with the evidence. Close unresolved dependencies or bring the founder a specific decision.

Days sixty one through ninety

Complete one material commitment through acceptance and transfer its ongoing responsibilities. Package the learning results, recovery record, full cost and operating instructions. Decide whether to fund a second deployment, another test cycle or a narrower scope. Issue an investor or board update that explains the evidence, financial consequences and next decision.

Management measure	What it helps reveal
Time from material issue to decision	Whether priorities and decision rights reduce waiting
Milestones accepted with stated evidence	Whether delivery is supported by usable results
Cash required to the next evidence point	Whether the strategy remains financially executable
Intervention and recovery burden per task	Whether autonomy produces a useful operating benefit
Retained performance after learning	Whether improvement preserves earlier capability
Support effort at the next deployment	Whether the commercial model can become repeatable

Set targets from actual baselines and customer requirements. At day ninety, review capability, cost and support burden against the plan. Confirm that the next commitment has an owner and a funded delivery path. The outcome should be a defensible next decision, with less dependence on constant founder intervention.

11 PRODUCT AND COMPANY DELIVERABLES

Translate the findings into a coherent product plan

The fieldwork points to evidence and operating capabilities that make adaptation a dependable service. Translate that priority into deliverables with an owner, an acceptance test and a funded place in the plan. A company can retain an ambitious research agenda while giving customers, engineers and investors a concrete basis for the next commitment.

Deliverable	Acceptance evidence	Accountable owner
Learning evaluation package	New skill acquisition, retention of prior skills, continued adaptability and resource use against a relevant baseline.	Research lead
Deployment control package	Monitoring triggers, intervention authority, tested containment and restoration, and a record of learning state.	Evaluation and safety lead
Reusable integration package	Documented sensor and action interfaces, timing limits, representative tests and reproducible setup instructions.	Engineering lead
Supported pilot package	Customer acceptance, full task cost, service responsibilities, operating limits and a defined commercial decision.	Product and delivery lead
Strategy and execution plan	Priorities, dependencies, decision rights, delivery owners and accepted milestones, updated when evidence changes.	Founder and operating lead
Capital and communication plan	Milestone budget, monthly cash forecast, scenarios, substantiated pitch and concise investor or board decisions.	Operating and finance leads

Keep the learning capability reusable

Where the evidence supports it, separate reusable learning capability from customer-specific integration. Define stable interfaces and controlled evaluations so another task or machine can be assessed without rebuilding the product. Reuse tests, operating records and integration tools as well as model components. This creates a practical basis for expansion even when the first contract is narrow.

Broaden only after the next constraint is understood

Transfer across platforms and coordination among multiple agents should have their own tests, owners and budgets. A result on one machine does not establish transfer to another, and a good individual agent does not establish coordinated performance. Use each expansion to learn which part of the system is reusable and which requires additional engineering.

These responsibilities can be combined while the team is small. Add specialist depth when the next obligation demands it, and keep one accountable owner for the whole company plan. A deliverable is useful when another person can rely on it to make a decision or operate the result without reconstructing the founder’s intentions.

12 FINDINGS AND INTERPRETATION

What the findings change about company strategy

Operating discipline is uneven across the measured practices

Written acceptance criteria (45%) and a deployment owner (40%) are the most frequently implemented practices. Milestone-linked budgeting follows at 36%, with human-support costing at 16% and tested recovery at 11%. The pattern identifies a management priority: extend accountability beyond the initial promise to the cost and work of sustaining it.

The results describe separate practices, not a sequence of failures within the same companies. They do not establish whether funding, stage or deployment exposure explains the differences. A founder can act on the findings by examining each obligation before expanding a customer commitment, then tracking whether the company's own interventions improve delivery.

The buyer evidence supports dependable adaptation

Retained performance after adaptation (90%) and intervention and restoration (79%) lead the buying conditions. This supports treating dependable adaptation as a product requirement: new capability must coexist with retained skills and a usable response to deterioration. The commercial implication is to sell an operating result supported by continuing evidence, rather than relying on a demonstration alone.

Human effort belongs in the value proposition

Documented operating cost is a buying condition at 53%, while 16% report task costing that includes human support. These are different measures, so their percentages do not quantify a matched market gap. Their management implication is shared: supervision, recovery and customer labor belong in the cost of useful work and in the service commitment.

The first sale and expansion require different proof

The 37% delivery-owner and 31% second-site findings are consistent with a need to distinguish the first evaluation from a broader rollout. They do not establish why respondents ranked those requirements lower. Founders should prepare repeatability before promising scale, with a named owner and evidence of the resources required at another site.

External research gives context to the priorities

Three developments shape the opportunity. Research is examining sustained adaptation and retained learning; robotics work is testing transfer across embodiments; and deployment research is giving more attention to monitoring and human oversight. [3–8] The practical implication is a broader product obligation: evaluate what the system keeps learning, where that learning transfers and who supports the resulting behavior.

My recommendation is to fund a narrow customer application that builds reusable evaluation, learning and integration assets. Connect each new result to the roadmap, cash forecast and delivery commitments before broadening the offer. The fieldwork informs that priority; subsequent operating results must determine whether it is working for the company.

13 CONCLUSION

Build a company that can sustain what it learns

An adaptive AI venture needs company execution that can respond as quickly as its evidence changes. The founder must be able to trust the connection between the research result, the financial plan and the customer commitment. The fieldwork gives that leadership task substance: retained capability and recovery lead buying conditions, while the implementation of support costing and tested recovery remains limited.

Retained performance after adaptation is a buying condition at 90%; defined intervention and restoration follows at 79%. Build longitudinal evaluation and usable operating controls into the customer offer. Their value must be demonstrated over repeated changes and carried into the budget, product roadmap and support responsibilities.

Make the next proof commercially meaningful

Choose a recurring customer problem where adaptation can produce a measurable benefit. Agree on the baseline, variation, acceptance criteria and cost record before the pilot begins. Preserve earlier skills while adding new ones, and record the time and resources required. Use the result to decide what to fund next rather than treating every successful demonstration as permission to expand.

Build the operating capability alongside the science

The 36% milestone-budgeting, 16% human-support-costing and 11% tested-recovery findings identify obligations to address before commitments expand. Give each one an owner and a funded path to acceptance. Maintain financial forecasts and investor communications from the same evidence, so a change in technical progress reaches every consequential company decision.

Preserve the broader opportunity through disciplined delivery

A focused first application can create reusable learning, evaluation and integration capabilities. Preserve that opportunity by separating general capability from work specific to one customer. Expand across tasks, platforms or coordinated agents when transfer, retained performance and support economics have been demonstrated under the new conditions.

The immediate deliverable is a company plan the team can execute: a substantiated claim, a funded milestone, clear decision rights and an outcome the customer can accept. Leadership succeeds when difficult choices are made promptly and carried through to completion. Each completed commitment should leave the business more capable, the evidence more useful and the founder freer to shape what comes next.

STUDY METHODOLOGY

Fieldwork scope and interpretation

Black Book Insights conducted the AI Startup Leadership and Deployment Readiness Study in collaboration with Applied AI LLC. Fieldwork in October 2026 covered 150 AI industry stakeholders. The ten study percentages are reproduced in Figures 1 and 2 and form the primary research basis for this paper.

What the measures report

Measure group	Reporting definition	Coverage
Startup operating practices	Fully implemented practices	Acceptance criteria, deployment ownership, milestone budgeting, human-support costing, and tested containment and restoration.
Buying conditions	Conditions reported as mandatory	Retained performance, intervention and restoration, documented operating cost, delivery ownership, and second-site reproducibility.

Each percentage describes a separate item. The measures are not mutually exclusive and are not intended to sum to 100%. Operating practices and buying conditions answer different questions; they are presented separately.

How the findings are interpreted

The study records stakeholder-reported practices and requirements. It does not independently test models, certify deployment safety or measure return on investment. Recommendations combine the survey findings with the author's analysis and the published research listed in the references.

The study total is 150 stakeholders. Separate founder and buyer counts, item-level response bases, respondent characteristics, recruitment and weighting are not reported in the study summary. Percentages are presented as reported, without inferred respondent counts, confidence intervals, significance tests or claims of population representativeness.

The analysis does not pool the two measure groups, infer that one response causes another, or calculate matched gaps between different questions. A low reported mandatory-condition rate does not establish that a requirement lacks commercial importance. Decision stage, use case and deployment exposure can influence its relevance.

Basis for the recommendations

The analysis draws on the substance and ordering of the ten measures, particularly the emphasis on retained performance and recovery in buying conditions and the lower implementation of support costing and tested recovery. External research provides technical context. The product and leadership recommendations are management judgments informed by these findings, rather than additional survey results.

Sources and publication details

Research reviewed through November 2, 2026. The completed Black Book Insights fieldwork supplies the reported findings. The sources below provide technical and industry context.

[1] Stanford Institute for Human-Centered Artificial Intelligence. The 2026 AI Index Report. 2026. Context on the concentration of notable frontier-model development in industry.

2026 AI Index Report

[2] International Federation of Robotics. President's Report by Jane Heffner. September 24, 2026. World Robotics 2026 figures for 2025 installations and operational stock.

IFR President's Report

[3] Dohare, S., Hernandez-Garcia, J. F., Lan, Q., Rahman, P., Mahmood, A. R., and Sutton, R. S. Loss of plasticity in deep continual learning. Nature 632, 768-774. August 21, 2024.

Nature article

[4] Pan, C., Yang, X., Li, Y., Wei, W., Li, T., An, B., and Liang, J. A Survey of Continual Reinforcement Learning. arXiv:2506.21872, revised April 7, 2026. Preprint.

Continual reinforcement learning survey

[5] Parada, C. Gemini Robotics 1.5 brings AI agents into the physical world. Google DeepMind, September 25, 2025. Developer-reported research on embodied reasoning and transfer across robot platforms.

Google DeepMind research announcement

[6] Jin, W., Du, H., Tang, S., Zhao, B., and Yang, G. A Comprehensive Survey on Multi-Agent Cooperative Decision-Making: Scenarios, Approaches, Challenges and Perspectives. arXiv:2503.13415, revised August 28, 2026. Preprint.

Multi-agent decision-making survey

[7] Rao, A., et al. Challenges to the monitoring of deployed AI systems: Center for AI Standards and Innovation. NIST AI 800-4, March 6, 2026.

NIST AI 800-4

[8] Anthropic. Measuring AI agent autonomy in practice. February 18, 2026. Provider research on Claude Code and public API activity; findings relate to software use.

Anthropic research study

[9] Author background and published work. LinkedIn professional profile; product-leadership article (2020); KDnuggets author biography (2015); eFinancialCareers professional coverage (2023); Applied AI LLC advisory profile (2026).

Author and research contribution

Vasyl Harasymiv, Founder of Applied AI LLC, Chicago. Vasyl.Harsymiv@gmail.com 608.622.5327
Research contribution by Black Book Insights, Tampa, Florida, and London, United Kingdom.