

WEBINAR 03 — LAUNCH THE PILOT

90-Day Robotic Data Pilot Planner

Launch a Real-World Robotics Data Program in 90 Days

A practical roadmap from focused pilot to production scale.

Organizations do not need to solve their entire data problem at once. A focused pilot can validate data quality, operational feasibility, and model impact before a larger investment is made. This planner gives you the structure, gates, and worksheets to run that pilot well.



Duration: 45 minutes • Dates to be announced

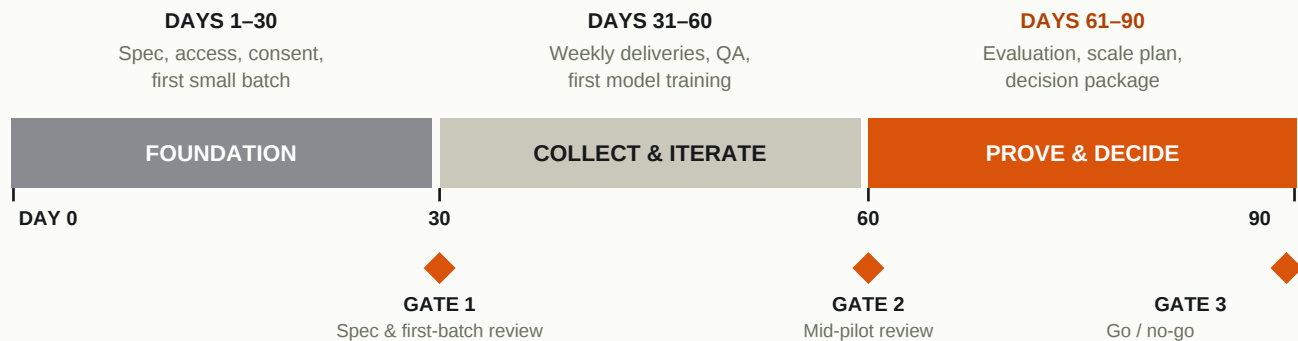
Robotic Data — Data for the Physical AI Revolution

OVERVIEW

Ninety days from idea to evidence

The fastest way to de-risk a robotics data program is not a bigger plan — it is a smaller one, run well. A focused 90-day pilot produces the three proofs a larger investment needs: the data can be collected at quality, the operation is feasible, and the model measurably improves.

The 90-day pilot at a glance



Three 30-day phases, each ending in a formal gate. Shorter engagements truncate the same structure: a 30-day pilot is a feasibility spike ending at Gate 1; a 60-day pilot adds first model results at Gate 2; the full 90 days delivers a complete go/no-go decision package.

The planner mindset

Everything in this resource works backward from Gate 3: what evidence will the go/no-go decision need, and what must happen in each phase to produce it?

SECTION 01

01 Why robotic data pilots stall

Most stalled pilots fail before the first sensor records anything. The recurring causes:

- **Scope too broad.** The pilot tries to solve the entire data problem at once instead of proving one task in one environment.
- **No acceptance criteria.** Without pre-agreed thresholds, no result can prove success — every review becomes a debate.
- **Compliance surfaced late.** Security review, site permissions, and consent frameworks discovered in week six block collection indefinitely.
- **Unusable deliveries.** Data arrives in formats the ML team can't ingest; integration work silently consumes the pilot window.
- **No owner, no gates.** Nobody is accountable for a decision, so the pilot drifts until budget or attention runs out.
- **Volume worship.** Success is reported in terabytes collected rather than model or operational impact — leadership can't act on it.

The pattern

Pilots rarely fail at collection. They fail at definition — scope, criteria, compliance, and ownership left unresolved until they become blockers.

SECTION 02

02 Select the right first use case

Score each candidate use case 1–5 on the criteria below. Favor the highest total that no single criterion fails.

Criterion	What good looks like	Score /5
Business value	A measurable pain with an executive sponsor who wants the result	
Data leverage	Performance is plausibly data-limited — more or better data should move the metric	
Feasibility	Accessible environment, manageable safety and permission requirements	
Bounded scope	One task, one environment type, defined modalities — completable in 90 days	
Expandability	Success generalizes: the playbook extends to other regions, sites, and tasks	

Anti-pattern

The most ambitious use case is almost never the right first one. Pick the pilot you can finish, not the one you'd most like to show off.

SECTION 03

03 Define the required data and delivery format

Everything below must be written down and agreed before collection begins — it becomes the contract the pilot is judged against.

The data specification

- Modalities, resolution, and frequency for every sensor; annotation types and label taxonomy; required metadata per episode.
- Volume targets and diversity targets (environments, conditions, behaviors) expressed as a coverage matrix.

Acceptance criteria

- Quality thresholds: time-sync tolerance, calibration residuals, label accuracy, completeness per episode.
- Coverage thresholds: percentage of the matrix filled; minimum episodes per cell for priority conditions.
- Rejection rules: what gets sent back, and how quickly it must be recollected.

Delivery format and cadence

- File formats and schema documented and tested with a sample batch before the pilot starts.
- Weekly incremental deliveries — never a single end-of-pilot dump — so integration and QA problems surface in week two, not week twelve.

Security, privacy, consent, and ownership

- PII handling: face and plate blurring, audio scrubbing, and retention limits agreed in writing.
- Consent framework for demonstrators, workers, and bystanders; site NDAs and permissions secured in phase one.
- Data ownership, license terms, and permitted uses (training, evaluation, redistribution) settled contractually up front.

Week-two test

By the end of week two you should have ingested a small real batch end-to-end: collected, processed, delivered, loaded, and QA-scored. If that path doesn't work small, it won't work big.

SECTION 04

04 The Robotic Data 90-day pilot framework

Phase	Focus	Exit gate
Days 1–30 FOUNDATION	Lock the use case, data spec, and acceptance criteria. Secure site access, permissions, and the consent framework. Stand up the pipeline and run a small first collection batch end-to-end.	Gate 1 — spec locked and first batch accepted: quality metrics computed, format ingested by the ML team, compliance sign-off complete.
Days 31–60 COLLECT & ITERATE	Run full collection cadence with weekly deliveries. Track QA and coverage metrics against the spec. Train or fine-tune the first model on pilot data; feed early findings back into the collection plan.	Gate 2 — mid-pilot review: coverage on track, quality stable, first model signal measured; spec adjusted where reality disagreed with the plan.
Days 61–90 PROVE & DECIDE	Complete deliveries and final QA. Evaluate the model against baseline on held-out real-world data. Compile operational metrics, unit costs, and a scale plan with budget scenarios.	Gate 3 — go/no-go: a decision package with model impact, operational evidence, cost model, and a scaling roadmap leadership can act on.

Roles that make it move

- **Pilot owner** — one named person accountable for gates and the final decision.
- **Executive sponsor** — clears permissions, budget, and cross-team blockers.
- **ML lead** — owns ingestion, training runs, and the evaluation protocol.
- **Data partner lead** — owns collection operations, QA, and delivery cadence.

SECTION 05

05 Evaluate quality, coverage, diversity, and model impact

Gate 3 needs evidence across four lenses — plus the operational numbers that determine whether scaling is affordable.

Lens	What to measure
Quality	Sync tolerance, calibration residuals, label accuracy, episode completeness, rejection rate over time
Coverage	Percentage of the specified coverage matrix filled; episodes per priority condition; gaps explicitly listed
Diversity	Spread across environments, sites, lighting, weather, and human behaviors — measured, not asserted
Model impact	Delta vs. baseline on held-out real-world data; error-cluster reduction in targeted conditions; sim-to-real gap change
Operations	Cost per accepted episode, collection throughput, QA turnaround, intervention and recollection rates

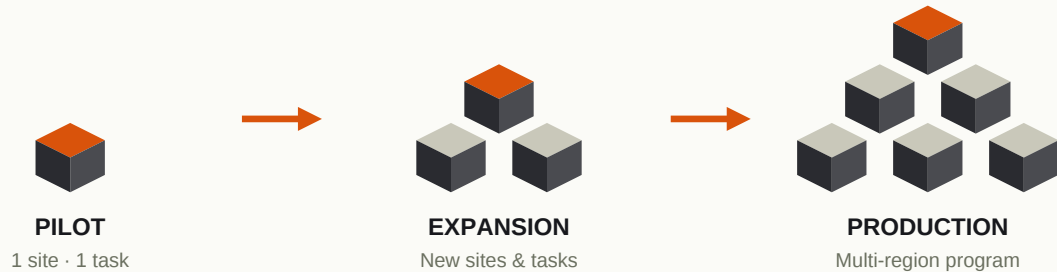
Evaluation discipline

Define the evaluation protocol — held-out sets, baselines, and metrics — in phase one, before anyone has an incentive to move the goalposts.

SECTION 06

06 Move from pilot to continuous production

A successful pilot's real product is a proven playbook. Scaling means replicating it — across sites, regions, and tasks — and converting the one-off engagement into a continuous data operation.



- ☐ Scale plan names the next sites, regions, and tasks in priority order
- ☐ Unit economics from the pilot project affordably to production volumes
- ☐ Delivery cadence becomes a standing contract with SLAs, not a project
- ☐ Evaluation suite from the pilot becomes the permanent regression gate
- ☐ Deployment failures feed a targeted recollection loop (the data flywheel)

SECTION 07

07 Your pilot, on one page

Fill this in during the session. The final agenda item is live audience Q&A — the most useful questions name a specific task, site, or blocker from this worksheet.

PILOT USE CASE (ONE SENTENCE)

EXECUTIVE SPONSOR / PILOT OWNER

TARGET START DATE · PILOT LENGTH (30 / 60 / 90 DAYS)

ACCEPTANCE CRITERIA (TOP THREE)

SECURITY, PRIVACY, OR CONSENT ITEMS TO RESOLVE

GATE 3 DECISION-MAKERS

QUESTIONS FOR THE LIVE Q&A

Webinar 03 — Launch the Pilot

45 minutes • Dates to be announced • www.roboticdata.com — Robotic Data creates the real-world datasets that power autonomous vehicles, humanoid robots, drones, and the next generation of Physical AI.