

Robot IQ: Teaching Robots with Human Know-How

Field IQ Ultra runs a Skill Pack for a person. Robot IQ runs the same Skill Pack on a robot.



Summary.....	2
1. Why robots need know-how, not just bodies.....	3
2. What Field IQ already teaches people.....	4
3. The ways to teach a robot, in the agreed order.....	4
3.1 First: human video from Field IQ.....	4
3.2 Second: simulation.....	5
3.3 Third: a little remote control on the target robot.....	5
3.4 How the three combine: one valve.....	5
4. Our approach: two birds, one stone.....	6
5. The capture database.....	6
5.1 Do we capture enough today? No.....	6
5.2 What one session stores.....	7
5.3 Why it is gold.....	7
5.4 Four shelves, never mixed.....	7
6. Raw and distilled: why we keep both.....	8
7. What a Skill Pack is and holds.....	9
8. One pack for a person or a robot.....	9
9. The flywheel to EON Delta.....	10
10. Safety and consent.....	10
11. Roadmap (proposed; Dan to confirm).....	11
12. Glossary.....	11

Summary

Robots are getting good bodies. What they lack is know-how: which valve to close first, how far to turn it, how to tell the job is done. Field IQ already teaches that know-how to people, one checked step at a time. Robot IQ uses it to teach robots. In one line: **Field IQ Ultra runs a Skill Pack for a person. Robot IQ runs the same Skill Pack on a robot.**

- **A Skill Pack** packages one industry's jobs (for example oil and gas). Inside, one module per machine family; each job is one approved SOP with its steps, what "done" looks like, the parts, expert demonstrations and the safety rules.
- **We teach robots in this order:** 1) human video from Field IQ, the cheapest lesson; 2) simulation, which turns a few examples into thousands; 3) a little remote control, which fits the result to one robot's body.
- **EON Genesis multiplies the practice.** Genesis will run thousands of simulations of the same skid and the same task, so a robot can practise one job, such as closing a valve on a pump skid, many thousands of times, with variations, before it ever touches the real machine. EON already has the 12 oil-and-gas skids in 3D, with their SOPs. The models still need physics: mass, friction and joint limits (section 3.2).
- **Two birds, one stone.** The sessions we record to train and check people also become lessons for robots.
- **Why EON is unique.** EON combines a product that teaches people at the real machine and records how skilled work is done, step by step, checked by Assess IQ (Field IQ); industrial

procedures and 3D machines (Genesis, the Situation Room skids, Mirror Direct); and simulation at scale in Genesis. All three feed one Skill Pack, for a person or a robot: work intelligence for people today, and for robots next.

- **We do not capture enough today.** Field IQ keeps steps, times and verdicts well, but throws away the motion the phone already measures.
- **So we build one capture library:** every guided session with its video, 3D hands, depth, phone position, parts, step labels and verdicts.
- **We keep each session twice:** raw (everything, for future models and new robots) and distilled (a small labelled summary, for today's training). We need both.
- **The flywheel:** every checked step is a record. Assess IQ labels it, Compound IQ joins it, a distillery makes training sets, and EON Delta, our own model built on an open-weights base, learns from them.
- **Where it starts:** valve isolation on pump skids (people fit every lock and tag); switchgear racking, keeping people away from arc-flash risk; routine inspection rounds (reading gauges, checking for leaks); taking samples at hazardous points; CNC machine changeover and setup; maintenance in confined or hot spaces. The robot takes the dangerous or repetitive step; people stay in charge.
- **Safety and consent come first.** No video or sound leaves the phone until the worker consents and the customer's contract allows learning. Robots start on mock-ups with no stored energy.
- **The long-term goal:** an off-the-shelf robot loads an oil-and-gas or aerospace Skill Pack and does the job better than a frontier AI model alone. It stays a goal until our own test set proves it.

1. Why robots need know-how, not just bodies

A robot arm can grip a handwheel. That does not mean it knows what to do with it.

Take a pump skid. To make it safe for work, someone must switch off and lock the motor's breaker, close and lock the suction and discharge valves, hang tags, bleed the pressure, and press start to prove nothing moves. Order matters. "Done" has a look: the handwheel stops turning, the lock is through the wheel, the gauge reads zero. Skip a step and someone can get hurt. In the US this lockout/tagout work is governed by OSHA 29 CFR 1910.147

(www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.147).

None of that is in the robot's body. It is know-how, and it lives in experienced workers' heads.

Big robot models such as NVIDIA's GR00T N1 learn from human video, synthetic data and real robot data (arxiv.org/abs/2503.14734). That gives a robot general grasping, not your plant's procedure. Our dataset search on 15 October 2026 found no public dataset of oil-and-gas, pump-skid or lockout work.

So the scarce thing is not the robot. It is checked, step-by-step demonstrations of real industrial jobs. That is what Field IQ makes.

2. What Field IQ already teaches people

Field IQ guides a worker through a job on a phone, headset or glasses. At each step it shows a card (“Step 4: close the suction valve clockwise until it stops”), and a frontier AI model checks a photo. The step passes, or the worker looks again. Each session leaves a record:

- **Steps**, each with an instruction, a caution and one of 16 action verbs (open, close, press, turn left and so on).
- **Times** for each step, to a tenth of a second.
- **A verdict per check**: done, not done or unclear, with what the AI saw and how sure it was.
- **Parts**, named, with a box around each on the photo.
- **The expert’s gold run**. With Rehearse, an expert records the job once; Field IQ then compares each worker’s steps with it.
- **A tamper-evident chain**. Each record is sealed to the one before, so any change shows.

Take a simple job: setting an air conditioner with its remote. Field IQ shows “Step 3: press MODE until the snowflake shows”, checks a photo of the display, and logs “done in 2.1 seconds, verdict done, 92% sure”. A robot could learn the step order and the check from that. It could not learn how the thumb moved, because we keep a photo, not the motion.

3. The ways to teach a robot, in the agreed order

Labs use five ways: remote control, human video with hand tracking, simulation, big robot models fine-tuned on a few demonstrations, and planners that turn written steps into robot steps. The strongest systems mix them. Our agreed order is human video, then simulation, then remote control. An open robot model is the base brain; the Skill Pack’s steps are its plan.

3.1 First: human video from Field IQ

A person does the job while a camera records their view and their hands. Software tracks the hands in 3D, and the robot learns to copy the motion. Why first:

- **It is the cheapest lesson**. In the EgoMimic study, one hour of human hand data gave about 1,400 demonstrations, against about 135 from an hour of robot data. The authors write that “one additional hour of hand data is significantly more valuable than one hour of robot data” (arxiv.org/abs/2410.24221).
- **It is fast to record**. In the UMI study, bare hands managed 231 demonstrations an hour; remote control managed 35 (arxiv.org/abs/2402.10329).
- **The field has moved this way**. NVIDIA’s EgoScale pre-trained on 20,854 hours of human head-camera video and raised average success by 54% compared with the same model without that pretraining (abstract; whether relative or absolute is not stated; arxiv.org/abs/2602.16710). Physical Intelligence reported about twice the success on tasks its robot had seen only in human video (pi.website/research/human_to_robot).
- **Our labels come free**. Field IQ knows which step card was showing and whether it passed. Public datasets mostly guess; Apple’s EgoDex warns that its written labels may be wrong (github.com/apple/ml-egodex).

The catch: a human hand is not a robot gripper. Researchers call this the embodiment gap. It is why step three exists.

3.2 Second: simulation

A simulator is a video-game copy of the world with real physics. A robot practises there thousands of times, on thousands of slightly different valves, at no risk. NVIDIA Isaac Sim and Isaac Lab are open tools for this (github.com/isaac-sim/IsaacSim, github.com/isaac-sim/IsaacLab).

Simulation multiplies. MimicGen made over 50,000 demonstrations from about 200 human ones (arxiv.org/abs/2310.17596). DexMimicGen made 21,000 from 60 (arxiv.org/abs/2410.24185).

But simulation needs a seed. A simulator cannot invent which valve to close first. That comes from people, so simulation comes second. Dan's idea of distilling from simulation is right for the multiplying.

We have a head start. EON Genesis models already say, per step, which part moves and how: axis, distance, angle. They lack physics: mass, friction and joint limits.

3.3 Third: a little remote control on the target robot

Remote control (teleoperation) means a person drives the robot while the robot records its own moves. It gives perfect robot data, but slowly. Since every robot body differs, a few runs on the real robot close the last gap:

- EgoScale's final stage used about 50 hours of human data and 4 hours of robot remote control over 344 tasks (arxiv.org/abs/2602.16710).
- On a second humanoid with only 20 robot demonstrations, adding another humanoid's data and human data clearly beat training on those 20 alone (arxiv.org/abs/2503.13441). The paper gives no absolute success level.
- OpenVLA, an open robot model, is fine-tuned with 10 to 150 demonstrations per task (arxiv.org/abs/2406.09246).

The tools are cheap. A GELLO controller, a small 3D-printed copy of the robot arm, costs under \$300 in parts (wuphilipp.github.io/gello_site/).

3.4 How the three combine: one valve

Layer	Comes from	Gives the robot
What to do, in what order, what "done" looks like	Field IQ, distilled	"Step 4: close the suction valve clockwise until it stops."
How the hand moves	Field IQ, raw	The hand's 3D path on the handwheel, every frame.
Many versions	Simulation	Thousands of valves: sizes, angles, lighting, stiffness.
This robot's body	Remote control	50 to 100 runs on the real arm (our pilot plan).
The base brain	An open robot model	General grasping, fine-tuned on all of the above.
The checker	The same Field IQ check	A photo after the move, plus the robot's own end-stop reading: done, not done or unclear.

4. Our approach: two birds, one stone

Robots learn from demonstrations. But most people do not know how to do an industrial job well. Field IQ solves that: it teaches people and checks them. So every guided session is a demonstration with a teacher's marks on it. One capture does two jobs: it trains the person today, and it becomes a lesson for a robot later, once we keep the motion as well as the photo.

Field IQ's loop is "do a step, check it, move on". Robot research uses the same loop: Inner Monologue fed a success check back to the robot's planner after each step (arxiv.org/abs/2207.05608). Big robot models split "which step next" from "how to move" (arxiv.org/abs/2503.14734). A Skill Pack's steps fill the first slot; its checks keep the robot honest. The Ego-Exo4D authors make the same point: step data can help robots "learn the structure of a procedure from human demonstrations" (arxiv.org/abs/2311.18259).

5. The capture database

Dan: "That video library is very important."

5.1 Do we capture enough today? No.

To copy a motion, a robot needs much more than a photo (arxiv.org/abs/2410.24221, arxiv.org/abs/2602.16710): continuous video at about 30 frames a second (we keep 1 to 3 small stills per step); where the camera was, every frame; depth, so distances are real; the hand's joints in 3D, commonly 21 per hand (facebookresearch.github.io/projectaria_tools/docs/data_formats/mps/hand_tracking); where the parts are; and one clock lining it all up. Force and touch matter too, but a phone cannot record them.

The good news: the phone already works out much of this, then throws it away. At every check it computes the phone's position, the lens numbers and LiDAR depth. Apple's hand detector already runs during Watch me and can report 21 joints per hand (developer.apple.com/documentation/vision/vnhumanhandposeobservation/jointname), but we keep only "hands in view: yes or no". Keeping these is the cheapest first fix.

Some things need something special:

- **The right camera spot.** A phone on a chest band, back camera, gives a near first-person view. "Facing me" mode uses the front camera, with no position and no depth, so it is not robot-grade.
- **Depth** needs a LiDAR phone or a headset. Whether Meta's glasses give apps depth or hand tracking is to confirm (wearables.developer.meta.com/docs/develop/dat/build-integration-ios/). The glasses show the steps; the chest phone records.
- **A second view** from the side, where we can, like Ego-Exo4D's outside cameras (arxiv.org/abs/2311.18259). A device test for heat and battery comes before any video.
- **Force and touch** need instrumented tools or remote control. Not phase one.

5.2 What one session stores

Item	Why it matters	Size (estimate)
Video per step, no overlays, 30 frames a second	The motion itself	About 1.8 GB an hour
Phone position and lens numbers, every frame	Cancels body movement	About 21 MB an hour
3D hand joints, both hands	What the robot copies	About 49 MB an hour
Depth at each check	Real distances	About 246 KB per check
Parts in 3D, named from CAD where we can	What is touched	Small
Steps with times, action verb, tool	The plan	Small
Verdicts, retakes, reviewer labels	Right or wrong, and why	Small
Voice and transcript, lined up with steps	The expert's reasons	To measure
Device, mount, calibration	Makes it usable later	Small
Consent, training flag, contract, site, pseudonymous worker id	Proves we may use it	Small

On these estimates, one hour is about 2 GB raw, a 20-minute lockout session about 0.66 GB, and 1,000 sessions about 0.66 TB (worked out in the Robot IQ specification, section 2.12).

Video goes to cloud storage, never into the main database, which keeps the index and labels. We export in LeRobot formats: LeRobotDataset v3 for LeRobot and SmolVLA (huggingface.co/docs/lerobot/lerobot-dataset-v3), and GR00T's flavour of LeRobot v2 with a `meta/modality.json` file (github.com/NVIDIA/Isaac-GR00T). Which version openpi reads is to confirm. Anyone allowed can search it ("every 'close valve' step that failed its first check") and play the video with the steps, verdicts and hand skeleton on top. Each customer sees only its own sessions.

5.3 Why it is gold

Public datasets show kitchens, sports and bike repair. Ours would show real industrial machines and procedures, every step labelled and checked, expert runs next to novice runs, with consent on file. Mistakes are gold too: a retake shows a wrong move and its correction, and newer robot models learn from corrections (www.pi.website/blog/pistar06). And variety (more places, more objects) helps a robot more than repeating one scene (arxiv.org/abs/2410.18647). Field IQ works across many sites by design.

5.4 Four shelves, never mixed

Customers only ever see the Visible shelf.

- **Visible** EON's own video and consented customer sessions. Field IQ Ultra, Genesis, EON Live and customers may show these.
- **Internal** open research sets (IndEgo, IndustReal, HoloAssist), for Aiden's team only, with credit. Never shown to customers and never put into a Skill Pack as data: their people agreed to research, not to our product.

- **Locked** Ego-Exo4D and Ego4D, only for training and testing Assess IQ. No viewer, no export, no partners, no outside AI service. The Ego-Exo4D licence says its data “may not appear in or be visible in any program, dataset, or product” and bars any third-party access; Ego4D has the same terms (ego4d-data.org/pdfs/Ego-Exo4D-Model-License.pdf, ego4d-data.org/pdfs/Ego4D-Licenses-Draft.pdf).
- **Out** non-commercial or unclear licences, such as EPIC-KITCHENS, Assembly101, IKEA ASM and EgoDex. Not used.

6. Raw and distilled: why we keep both

Back to the air-conditioner remote.

- **Raw:** 8 seconds of video of a thumb pressing MODE three times, with hand joints, depth and phone position.
- **Distilled:** “Step 3: press MODE until the snowflake shows. Done in 2.1 s. Verdict: done, 92% sure.”

Why keep raw:

- **Future models** will want things we cannot name today, and we cannot film the job again.
- **Relabelling.** When Assess IQ improves, we rerun it over old sessions.
- **New robot bodies.** A two-finger gripper today, a five-finger hand tomorrow: each needs the motion translated again from the original. EgoScale found wrist-only labels did poorly and finger-joint labels worked best (arxiv.org/abs/2602.16710). Had we kept only the wrist, we could never fix it.
- **Audits.** The original proves what happened.

Why keep distilled:

- **Small and fast to search.** About 10 KB of text and labels per step; with the hand path and one blurred check photo, up to about 35 MB an hour, against about 2 GB raw (estimates).
- **Kinder to privacy:** faces blurred, no voices.
- **What we train on now:** step planners, step checkers, Skill Pack structure, EON Delta.
- **Easier to combine** across sites and customers, where contracts allow.

Raw is the master tape; distilled is the sheet music. Raw moves to cold storage after a set time (to decide) and is deleted on request; distilled is kept long term. Raw needs the stronger consent, because it holds video and sound. Both need training named as a purpose.

7. What a Skill Pack is and holds

A Skill Pack is one industry’s jobs, packaged so that a person or a robot can run them. Inside, one module per machine family, such as a pump skid; each job is one approved SOP. Proposed contents (the format is set in milestone M4):

- **The pack:** industry, version, who approved it, and its modules.
- **Each job:** name, machine type, the approved SOP and its version.
- **Steps:** instruction, action verb, part, tool, caution, expected time and end state (“valve closed, lock through the wheel”).
- **What “done” looks like:** look-for text, gold-run reference photos, check rules.
- **Parts:** real names from CAD, and where each sits in 3D.
- **Expert demonstrations:** gold runs, raw and distilled, from the visible library only.
- **Motion targets:** the Genesis part motion per step.
- **Safety rules:** hold points, stop conditions, who must approve, the lockout rules, and person-only steps (such as fitting a lock or a tag).
- **Variations:** a simulation scene and what to vary (valve size, lighting, stiffness).
- **Robot adapters (optional):** add-ons fitted to one robot body from remote-control runs.
- **Provenance:** which sessions it learned from, under which consent and contract, sealed.

8. One pack for a person or a robot

Part of the pack	Field IQ Ultra (a person)	Robot IQ (a robot)
Steps	A card on the screen or glasses	The robot’s plan, one step at a time
What “done” looks like	A photo check after the step	The same photo check, plus the robot’s own end-stop reading
Parts	Names drawn on the machine	Targets to reach and grip
Expert demonstrations	The gold clip the worker can replay	Training data for the robot’s motion
Motion targets	The Genesis animation	The simulator’s goal
Safety rules	Caution cards and hold points	Hard stops, and a person approves each step

Until now Robot IQ has been described only loosely. This makes it concrete: Robot IQ is not a separate product with separate content. It is a second player for the same Skill Packs.

The long-term picture: a plant buys an off-the-shelf robot, loads the oil-and-gas valve Skill Pack, and the robot does the job better than a frontier AI model alone, because the pack brings industrial demonstrations and checked steps no public dataset has. Research already uses robots you can buy; EgoScale ran on a Galaxea R1 Pro and a Unitree G1 (arxiv.org/abs/2602.16710). It stays a goal until our test set shows it.

9. The flywheel to EON Delta

Dan's words: "like a Tesla". The more it is used, the more it learns.

1. **Field IQ Ultra captures;** the capture library keeps every step, raw and distilled.
2. **Assess IQ labels it:** which step, which part, right or wrong and why, how skilled.
3. **Compound IQ joins the labels** across sessions, sites and machines: the fastest safe way, the common mistakes (our working meaning; to confirm).
4. **The distillery makes training sets:** step plans, step checks and hand paths.
5. **EON Delta learns.** EON's own model, built on an open-weights base, starts from general frontier AI data, then learns from EON's own: Field IQ, EON Live, Genesis, Tacit, Control Room and Situation Room. Thousands of interactions, then millions.
6. **Skill Packs improve** from what EON Delta and Compound IQ learn.
7. **Field IQ Ultra runs the pack for a person; Robot IQ runs it on a robot.** Every run makes new records. Back to step 1.

Two conditions. We use only AI models whose terms allow this. And each customer's contract must allow learning from its sessions, recorded as a training flag on each session. No flag, no training.

10. Safety and consent

- **Robot safety.** First runs are on mock-ups with no stored energy: no pressure, no power, no chemicals. A person approves each robot step, and the Field IQ check runs after every one. Lockout follows OSHA 29 CFR 1910.147 (www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.147). Which robot safety standards apply is to confirm with a safety engineer.
- **Some steps stay with people.** OSHA says lockout or tagout devices "shall be affixed to each energy isolating device by authorized employees" (29 CFR 1910.147(d)(4)(i), same link). So a robot never fits a lock or a tag. The Skill Pack marks those steps "person only"; the robot stops and hands them to a person.
- **Turning until it stops.** A robot closing a valve stops when it feels the end stop, at a torque measured on a real handwheel, never on time alone. The photo check then confirms it.
- **Worker consent.** Today's consent screen promises "Never video, never sound" and names no training purpose. Before any video, sound, hand data or learning we need a new consent version that names training, an updated privacy policy, a pseudonymous worker id (never a name), deletion on request, and a decision on face blurring.
- **Customer contracts.** A clause that allows learning, and a training flag per session.

- **Dataset licences.** Ego-Exo4D and Ego4D stay on the locked shelf, never shown or shared, and never put into an outside AI service. The licence gives no right to use their names or the licensors' names without written permission, except in the required citation. All copies are destroyed if the licence ends (ego4d-data.org/pdfs/Ego-Exo4D-Model-License.pdf). Counsel to confirm.
- **Honest claims.** We say “goal” until our test set says “result”.

11. Roadmap (proposed; Dan to confirm)

Milestone	When	What
M0 Paperwork	Oct–Nov 2026	Dan, or a director he names, signs Ego-Exo4D and Ego4D; counsel; new consent; contract clause; locked shelf.
M1 Keep what we compute	Nov–Dec 2026	Save position, depth, hand joints, clock and part positions; training flag; pseudonymous id.
M2 Capture library v1	Q1 2027	Video per step, chest mount, device test, store, search, viewer, LeRobot export.
M3 Assess IQ v2	Q1 2027	Step models trained on public sets, tested on EON clips; video with voice becomes an SOP.
M4 Skill Pack v1	Q2 2027	CAD to SOP; one pack format for both products; first Compound IQ, distillery and EON Delta runs.
M5 Robot IQ lab pilot	H2 2027	One valve on a mock-up with no stored energy; Isaac Lab scene from Genesis; 50 to 100 remote-control runs (our pilot plan).
M6 First industry pack	2028	One Skill Pack on an off-the-shelf robot, with a customer under contract.

12. Glossary

Word	Plain meaning
Skill Pack	One industry's jobs, packaged (for example oil and gas): one module per machine family; each job an approved SOP with steps, “done” looks, parts, demonstrations and safety rules.
Field IQ Ultra	The next Field IQ: guides a person through a Skill Pack, checking each step.
Robot IQ	Runs the same Skill Pack on a robot.
Assess IQ	Labels recorded work: step, part, right or wrong, skill.
Compound IQ	Joins labels across sessions and sites (working meaning; to confirm).
Distillery	Turns labelled sessions into training sets.
EON Delta	EON's own model, built on an open-weights base model that EON fine-tunes and runs. Whether EON publishes Delta's own weights is a decision for Dan; it needs contracts that allow it if customer data trained it.
Capture library	The one database of every guided session.

Word	Plain meaning
Raw / distilled	Everything recorded / the small labelled summary.
Gold run	An expert's recorded run, used as the standard.
Remote control (teleoperation)	A person drives the robot; it records its own moves.
Embodiment gap	A robot's body is not a human's, so motion must be translated.
Depth, LiDAR	How far away each point is; the phone sensor that measures it.
LeRobotDataset	A common file format for robot training data.
Shelves	Visible (video we may show: EON's own and consented sessions), Internal (open research sets, Aiden's team only), Locked (Ego-Exo4D, Ego4D: train and test only), Out (not used).