softinator®
TECHLABS

// FRONTIER MODELS // AGENTIC SYSTEMS // EMBODIED AI

THE HUMAN INTELLIGENCE LAYER FOR ROBOTS & AI.

Secure data, expert reasoning, model development, and physical-intelligence infrastructure for frontier teams, enterprises, and governments.

// 01 // POSITIONING

REASONING IN MOTION.

Softinator.ai is the frontier AI and robotics wing of Softinator® TechLabs: built to connect expert human judgment, high-signal data, frontier-model engineering, and real-world autonomy.

FROM SOFTWARE DELIVERY TO FRONTIER INTELLIGENCE

A deliberate systems pivot

The next computing interface is not another dashboard. It is an agent that can reason across code, tools, vision, and the physical world. We concentrate our delivery model around the infrastructure required to train, evaluate, and deploy those systems.

WHAT STAYS CONSTANT

Engineering discipline at enterprise scale

Four decades of group engineering heritage: secure delivery, structured quality systems, rigorous review, technical program management, and the ability to mobilize interdisciplinary talent remain the operating core.

300+
DEVELOPERS AND RESEARCHERS IN-HOUSE

47K+
PHDS, PROFESSORS, AND ELITE DEVELOPERS IN NETWORK

44+
YEARS OF GROUP EXCELLENCE, EST. 1981

STRATEGIC THESIS

Human expertise becomes model infrastructure

Our sister-company ecosystem in academic research publishing across the USA, UK, and Australia gives Softinator TechLabs direct reach into a global scholar community numbering in the hundreds of thousands. Softinator.ai translates that reach into carefully governed data programs, expert evaluation, post-training signals, and the next generation of embodied demonstrations for robotics.

GLOBAL PRESENCE: USA // AUSTRALIA // INDIA | ISO 9001:2015 | ISO 27001:2022 | SOC 2 TYPE II | NDA-GOVERNED DELIVERY

// 02 // OPERATING MODEL

ONE FOUNDRY. FOUR LAYERS.

A modular operating architecture lets enterprises enter at the data, model, agent, or deployment layer without losing traceability across the system.

01 // HUMAN SIGNAL

Expert data and evaluation

Domain-aware creation, annotation, critique, adjudication, and red teaming for logic-dense model behaviors. STEM CODE VISION

02 // LEARNING

Pre-training and post-training

Corpus engineering, SFT, preference signals, verifier data, model evaluation, and targeted adaptation. SFT RLHF EVALS

03 // SYSTEMS

Agentic applications

Tool-using workflows, software engineering agents, multimodal task environments, and evaluation harnesses.

04 // PHYSICAL INTELLIGENCE

VLA and embodied AI

Teleoperation, egocentric capture, simulation, robot learning, and edge inference for autonomous systems.

DESIGN PRINCIPLE

Security and provenance are part of the product

Every layer is designed around least-privilege access, isolated workstreams, role-based review, controlled exports, traceable quality gates, and client-defined retention.

// 03 // FRONTIER ROSTER

TRUSTED BY FRONTIER TEAMS.

We work with research, data, and engineering teams across the world's leading AI and robotics organizations, from frontier model laboratories to industrial automation platforms.

Google DeepMind
ANTHROP\C
amazon ai services
Meta AI
Mistral AI
Alibaba Cloud
YASKAWA
GreyOrange
Ai surge
scale

INDIVIDUAL ENGAGEMENT SCOPE IS GOVERNED BY NDA. ALL TRADEMARKS AND LOGOS ARE THE PROPERTY OF THEIR RESPECTIVE OWNERS.

// 04 // ENTERPRISE HERITAGE

TWO DECADES OF ENTERPRISE TRUST.

Before the frontier AI pivot, Softinator® TechLabs earned its engineering reputation delivering for enterprises across consumer goods, insurance, real estate, education, media, and retail.

Colgate · Godrej · UDACITY
Hiranandani · ROOTS MICE PLATFORM · Mahendra's
MANTRA FM 91.9 · OMdemand · RoboKart Share what you do
GRIFFCON · Pelz World Chiang Mai · + many more
HDFC ERGO, ALERSAL, ARTVY, ROMEO AI …

AMONGST OTHER ESTEEMED COLLABORATIONS ACROSS INDIA AND WORLDWIDE.

STRATEGIC ALLIANCES

Microsoft Solutions Partner · Google Cloud · aws · SAP Partner

FROM ENTERPRISE SYSTEMS TO FRONTIER AI

The delivery organization that shipped enterprise platforms, commerce systems, and digital transformation programs for these clients now powers Softinator.ai: one quality culture, carried from ERP integrations to frontier model development and robotics.

// 05 // CAPABILITY MAP

ARCHITECTING THE AUTONOMOUS STACK.

A focused set of capabilities for model builders, governments, robotics teams, and enterprises moving from proof-of-concept to governed operations.

MODELS

Foundation model development

Data strategy; corpus engineering; SFT and preference data; evaluation; adaptation; safety testing.

AGENTS

Agentic AI systems

Tool use; software agents; multimodal workflows; environment design; trajectory review; reliability evaluation.

ROBOTICS

Embodied intelligence

VLA strategy; perception-action datasets; teleoperation; simulation; edge deployment architecture.

DATA

High-signal data engines

Expert creation; annotation; critique; rubric design; adjudication; provenance; controlled delivery.

ENTERPRISE

Secure deployment

Private environments; observability; human approvals; red teaming; governance; integration.

INFRASTRUCTURE

AI-ready compute

Directional program for sovereign compute, GPU lifecycle planning, resilient power, cooling, and connectivity.

ENGINEERING LANGUAGES AND RUNTIMES

PYTHON | C++ | RUST | GO | JAVA | TYPESCRIPT | CUDA | ROS 2

Built to meet the client where the bottleneck lives

Engagements can begin with a tightly scoped evaluation set, a production data program, an agentic application, a model adaptation workstream, or a robotics data pilot. They then scale through the same governance and quality architecture.

// 06 // INDUSTRY COVERAGE

ONE DISCIPLINE. EVERY SECTOR.

Frontier AI, agentic systems, and robotics data programs are delivered with sector-specific task ontologies, domain experts, and compliance boundaries.

- MANUFACTURING
- AUTOMOTIVE & MOBILITY
- WAREHOUSING & LOGISTICS
- HEALTHCARE & PHARMA
- BANKING & INSURANCE
- RETAIL & CONSUMER
- MEDIA & EDUCATION
- ENERGY & UTILITIES
- PUBLIC SECTOR
- LABORATORIES & RESEARCH
- AGRICULTURE & FOOD
- AEROSPACE & DEFENCE

WHY INDUSTRY DEPTH MATTERS

Generic annotation and generic automation fail at the domain boundary. Sector-qualified experts define the task families, recognize unsafe or implausible behavior, and evaluate long-horizon outcomes so that models learn the judgment each industry actually requires.

// 07 // DELIVERY SYSTEM

FROM BRIEF TO HIGH-SIGNAL OUTPUT.

A production system for complex programs where quality cannot be reduced to throughput alone.

01 // DECOMPOSE THE TARGET BEHAVIOR

Translate model failure modes, deployment constraints, and domain expectations into atomic task families and measurable rubrics.

02 // CALIBRATE EXPERTS AND REVIEWERS

Run shared examples, ambiguity reviews, qualification gates, and reviewer alignment before production.

03 // PRODUCE, VERIFY, AND ADJUDICATE

Layered author-reviewer workflows combine automated validation with specialist judgment and exception routing.

04 // MEASURE ERRORS, NOT JUST VOLUME

Track taxonomy-level defects, disagreement, rework, drift, acceptance, and evidence completeness.

05 // DELIVER WITH PROVENANCE

Versioned schemas, controlled exports, lineage, documentation, and client-specific retention complete the loop.

// 08 // MODEL DEVELOPMENT

PRE-TRAIN. POST-TRAIN. PROVE.

Model performance is a data-and-evaluation system. We support the complete learning loop while protecting every confidential client program.

PRE-TRAINING

Corpus and domain signal

- Source strategy and taxonomy
- Deduplication and contamination controls
- Domain-expert content programs
- Metadata, provenance, and sampling

POST-TRAINING

Behavior shaping

- Instruction and supervised fine-tuning data
- Preference, critique, and revision signals
- Verifier-oriented reasoning traces
- Safety, policy, and adversarial evaluation

EVALUATION

Failure-mode intelligence

Benchmark-style tasks, regression suites, expert grading, rubric analytics, and slice-based reporting turn model quality into an operating discipline.

ADAPTATION

Enterprise specificity

Targeted data and evaluation for domain language, workflows, codebases, tool interfaces, visual artifacts, and constrained deployment contexts.

CODING | MATHEMATICS | STEM | MULTIMODAL | AGENTIC | SAFETY

// 09 // CODING INTELLIGENCE

ENGINEERING DATA FOR SOFTWARE AGENTS.

Experience patterns span code generation, comprehension, critique, repository context, verifier design, and environment-aware tasks across the modern language spectrum.

TASK FAMILIES

Beyond single-turn code generation

Repository-grounded comprehension, defect localization, patch reasoning, test construction, multi-language evaluation, code review, and tool-using trajectories, including systems work in Rust and Go. PYTHON C++ RUST GO JAVA TYPESCRIPT

QUALITY SYSTEM

Verifier-aware production

Executable checks where appropriate, rubric-based expert review where execution is insufficient, structured critiques, calibrated revisions, and adjudication for hard disagreements.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

Program pattern: advanced code verifier production

A frontier model team needs high-difficulty coding tasks with defensible solutions and machine-checkable evidence. Our workflow combines calibrated authors, independent reviewers, environment validation, issue taxonomies, exception handling, and daily quality telemetry while keeping the customer's model, benchmark, and target behavior confidential.

// 10 // MULTIMODAL QUALITY

WHEN MODELS MUST SEE, REASON, AND PRESENT.

We design evaluation systems for outputs that combine analytical correctness, visual fidelity, interaction, and professional judgment.

DATA VISUALIZATION

Correctness is multidimensional

A chart can be numerically valid yet misleading, visually broken, inaccessible, or inconsistent with the source. Evaluation therefore spans data fidelity, encodings, labels, hierarchy, responsiveness, and narrative clarity.

VISUAL ARTIFACTS

Professional task environments

Dashboards, documents, UI flows, and structured visual work products serve as realistic agent targets, supported by rubrics that separate functional, analytical, and presentation quality.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

Program pattern: multimodal visualization repair

A frontier multimodal model produces visualizations from business data. Our program isolates recurring failure modes, pairs the source with rendered evidence, routes samples through analyst and design review, captures structured critiques, and produces revision targets, all without identifying the model provider or proprietary dataset.

// 11 // FLAGSHIP PROGRAM

WE HELPED BUILD GEMINI 3.0.

Softinator.ai teams worked inside Google's Gemini 3.0 program: engineering expert data, multimodal evaluation, and repair signals that shaped release quality at frontier scale.

GOOGLE // PUBLIC ROSTER CLIENT | FRONTIER MULTIMODAL PROGRAM | SCOPE UNDER NDA

MULTIMODAL QUALITY

Data visualization at frontier standard

Analysts and visual reviewers evaluated and repaired model-generated charts, dashboards, and analytical artifacts across data fidelity, encodings, interaction, and accessibility.

EXPERT SIGNAL

Post-training data that teaches

Calibrated experts authored preference, critique, and revision signals for logic-dense behaviors, converting failure modes into precise training targets.

SCALE AND RIGOR

A network mobilized

Qualified specialists were mobilized through qualification gates, shared calibration, independent review, and daily quality telemetry across geographies.

gemini-viz-eval // production window

DELIVERY GLIMPSE // SANITIZED

```
$ evalrun -program gemini-3.0-viz -stage analyst+design-review
artifact classes: charts :: dashboards :: analytical documents :: interactive
rubric axes: data fidelity :: encodings :: hierarchy :: interaction :: accessibility
repair targets authored :: dual review complete :: batch telemetry exported
provider datasets and volumes remain confidential under program NDA
```

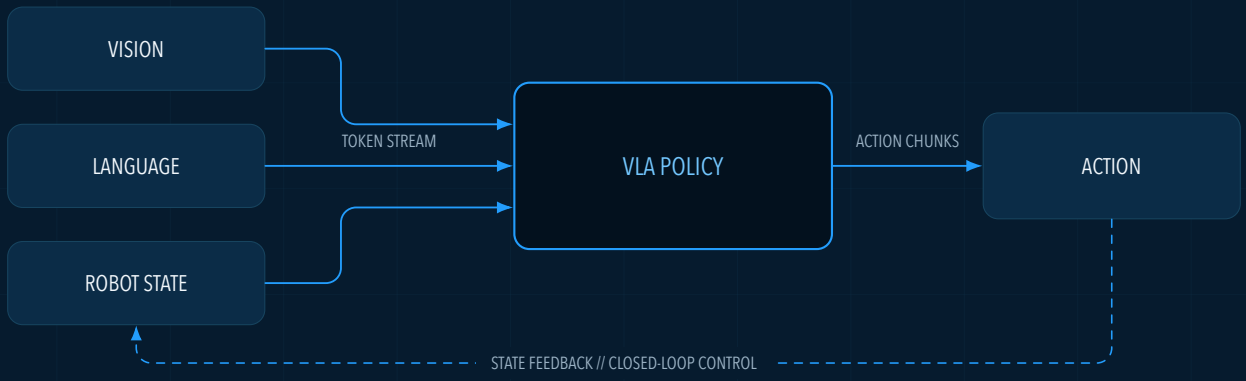
WHY IT MATTERS

Frontier releases are judged by the quality of what they produce for real users. Contributing evaluation and training signal to a program of Gemini 3.0's scale is direct evidence that Softinator.ai operates at the standard frontier laboratories demand.

// 12 // EMBODIED AI

THE VLA ENGINE.

The strategic frontier is moving from language-only systems toward models that connect perception, language, state, and action in one closed loop.



GROUNDING

Demonstration data

Egocentric and teleoperated trajectories designed around tasks, states, actions, failures, and recovery.

LEARNING

Policy training

Cross-modal alignment, action chunking, imitation learning, reinforcement learning, and targeted evaluation.

EXECUTION

Edge deployment

Latency-aware inference, safety interlocks, observability, human override, and hardware integration.

OPERATING DIRECTION

Softinator.ai delivers these capabilities through robotics partnerships, expert networks, simulation, and controlled data-capture programs, training generalist policy classes such as π_0 and OpenVLA-OFT with flow-matching action heads and System 2 chain-of-thought reasoning.

// 13 // ROBOTICS DATA

SOLVING THE GROUNDING BOTTLENECK.

Embodied models require more than video. They need synchronized, interpretable records of intent, observation, state, action, force, outcome, and recovery.

CAPTURE DESIGN

From human skill to training signal

- Task and workspace taxonomy
- Egocentric visual capture
- Teleoperation and action traces
- State, timing, and event alignment
- Failure and recovery coverage

DATA ENGINEERING

From trajectories to datasets

- Segmentation and hindsight relabeling
- Cross-embodiment normalization
- Quality and anomaly detection
- Privacy and sensitive-scene controls
- Versioned releases and lineage

DOMAIN ACCESS MODEL

Experts close the last-mile gap

Academic and professional networks help define domain tasks, recognize unsafe or implausible behavior, and evaluate long-horizon outcomes in settings where generic crowd work is insufficient.

// 14 // ROBOTICS PROGRAMS

PHYSICAL INTELLIGENCE, INDUSTRY BY INDUSTRY.

Programs are organized around industry task families and delivered through one shared data discipline, from capture to governed release.

MANUFACTURING

Dexterous assembly

Manipulation, alignment, fastening, inspection, tool selection, error recovery, and multimodal quality checks.

WAREHOUSING

Unstructured handling

Novel-object grasping, sorting, packing, exception routing, human-aware navigation, and inventory interaction.

LABORATORIES

Controlled procedures

Protocol-grounded tasks, instrument handling, sample movement, safety constraints, and auditable human approval.

FIELD SERVICE

Skilled maintenance

Diagnosis, tool use, part replacement, visual verification, and remote expert support.

PUBLIC SECTOR

Mission workflows

Secure onsite models, human-in-command controls, constrained autonomy, and policy-aware audit trails.

SIMULATION

Scale before deployment

Synthetic variation, scenario generation, domain randomization, policy stress tests, and sim-to-real evaluation.

// 15 // SELECTED ENGAGEMENTS

CODING INTELLIGENCE, BUILT WITH EVIDENCE.

Recent programs from the coding-intelligence practice. Client identities, models, datasets, and proprietary outcomes remain protected under NDA.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

CODE GENERATION // VERIFIERS // MULTI-STAGE QA

Advanced code-generation verifier program

CHALLENGE

High-difficulty solutions could appear plausible while failing hidden constraints, edge cases, or execution requirements.

HOW THE TEAM SOLVED IT

A calibrated author-reviewer-adjudicator team paired expert reasoning with executable validation, structured critiques, severity codes, and exception routing.

DELIVERED

Versioned task batches, defensible reference solutions, verifier evidence, review trails, defect analytics, and release notes.

REPOSITORY CONTEXT // AGENTIC CODING // EVALS

Repository-grounded code evaluation

CHALLENGE

Isolated snippets did not test whether an agent could navigate files, understand dependencies, and make a coherent change.

HOW THE TEAM SOLVED IT

The team constructed context-rich tasks, identified relevant artifacts, designed acceptance tests, reviewed patches, and captured failure taxonomies.

DELIVERED

Repository task packs, environment instructions, expected-change evidence, test criteria, expert critiques, and adjudicated outcomes.

verifier-run // batch RLX-448

DELIVERY GLIMPSE // SANITIZED

```
$ harness verify -batch RLX-448 -hidden-tests -sandbox microvm
[worker 12] task 448-031 :: candidate patch applied :: build OK in 41 s
- if (rebalance(node) != NULL) { return node; }
+ if ((node = rebalance(node)) != NULL) { return balance_parent(node); }
PASS 41/41 hidden tests // 2/2 independent expert approvals // adjudication: none
release gate: CLEAR // defect telemetry exported // evidence pack sealed
```

// 16 // SELECTED ENGAGEMENTS

REASONING ACROSS LANGUAGES AND DOMAINS.

Each program combines specialists, calibrated rubrics, independent review, and evidence-led release controls.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

C++ // JAVA // RUST // GO // STATIC REASONING

Cross-language code comprehension

CHALLENGE

The target model needed to interpret unfamiliar code, track state, reason about behavior, and distinguish subtle distractors without relying on execution alone.

HOW THE TEAM SOLVED IT

Language specialists authored and reviewed questions, validated explanations against source context, tested distractor quality, and adjudicated ambiguity.

DELIVERED

Curated comprehension sets, answer rationales, ambiguity flags, difficulty labels, review evidence, and schema-valid releases.

LOGIC // STEM // CALIBRATION

Expert reasoning and MCQ calibration

CHALLENGE

Questions that looked difficult often tested trivia, contained multiple valid answers, or leaked the solution through weak distractors.

HOW THE TEAM SOLVED IT

Domain reviewers calibrated difficulty, challenged assumptions, reconstructed reasoning, repaired distractors, and rejected under-specified items.

DELIVERED

Balanced task sets, verified keys, concise rationales, difficulty and skill tags, disagreement records, and quality summaries.

item-review // CLX-1187 // adjudication stage

DELIVERY GLIMPSE // SANITIZED

```
$ review item CLX-1187 -language cpp -stage adjudication
stem: state tracking across move semantics :: distractors 4 :: difficulty 0.82
reviewer B flags distractor D2 as defensible under C++17 :: escalated
adjudicator: D2 repaired :: rationale rewritten :: ambiguity resolved
item accepted :: key verified :: schema valid :: shipped in release R-2209
```

// 17 // SELECTED ENGAGEMENTS

EXPERT DATA FOR POST-TRAINING.

Human signal is engineered as a product: qualified contributors, explicit behavior targets, layered review, and release-level provenance.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

PHYSICS // MATHEMATICS // ENGINEERING

STEM expert data production

CHALLENGE

Generalist annotation could not reliably judge domain correctness, reasoning depth, notation, and subtle scientific errors.

HOW THE TEAM SOLVED IT

The program matched tasks to qualified specialists, used shared calibration rounds, introduced independent expert review, and escalated disagreements.

DELIVERED

Expert-authored responses, critiques and revisions, structured labels, adjudication records, and domain-sliced quality reporting.

SFT // RLHF // CRITIQUE // REVISION

Preference and critique signal program

CHALLENGE

Simple pass/fail labels did not explain why an answer failed or teach the model how to improve it.

HOW THE TEAM SOLVED IT

Reviewers compared alternatives, decomposed errors, wrote evidence-based critiques, produced improved responses, and applied a stable rubric.

DELIVERED

Preference pairs, granular critiques, revised targets, policy and quality labels, calibration examples, and documented edge cases.

preference-audit // pair PRF-0932

DELIVERY GLIMPSE // SANITIZED

```
$ trace PRF-0932 :: preference pair audit :: domain control-systems
response A: correct final result :: skips boundary-condition justification
response B: complete derivation :: unit-consistent :: cites governing equation
preference B over A :: structured critique attached :: revised target authored
author: PhD, control systems :: independent second review: agreement
```

// 18 // SELECTED ENGAGEMENTS

FRONTIER PRE-TRAINING DATA AT SCALE.

Corpus programs operate at web scale and expert scale simultaneously: trillions of candidate tokens filtered into governed, high-signal training mixtures.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

PRE-TRAINING // CORPUS // PROVENANCE

Trillion-token multilingual corpus engineering

CHALLENGE

Web-scale corpora mixed duplicate forks, benchmark leakage, licensing ambiguity, translation artifacts, and low-signal machine text across languages.

HOW THE TEAM SOLVED IT

The team built provenance graphs, ran exact, MinHash, and semantic deduplication, executed contamination probes, scored quality with trained classifiers, and balanced language curricula.

DELIVERED

Governed corpus releases, tokenizer and sampling policy, shard manifests, contamination reports, and loss-slice dashboards.

SCARCE DOMAINS // ACADEMIC NETWORK

Expert-authored domain corpus program

CHALLENGE

Frontier models needed high-signal text in specialist domains that public crawls cannot supply at the required depth or licensing clarity.

HOW THE TEAM SOLVED IT

The program mobilized the global academic network to author, review, and license domain content under explicit provenance and quality contracts.

DELIVERED

Expert corpus collections, licensing records, provenance metadata, difficulty maps, and per-domain quality telemetry.

corpus-pipeline // shard family CRX-9

DELIVERY GLIMPSE // SANITIZED

```
$ curator run -stage dedup+contamination -shards 4096
exact dup 11.2% :: near-dup (MinHash) 6.8% :: semantic dup 2.1% :: removed
benchmark-neighbor hits: 412 documents :: quarantined pending review
released shard family CRX-9 :: 184 languages :: provenance chain complete
```

// 19 // SELECTED ENGAGEMENTS

MULTIMODAL OUTPUTS THAT HOLD UP.

Evaluation spans analytical truth, visual communication, usability, and professional polish, not screenshot similarity alone.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

MULTIMODAL // ANALYTICS // VISUAL QA

Data-visualization diagnosis and repair

CHALLENGE

Generated charts could be numerically correct yet misleading, visually broken, inaccessible, or inconsistent with the source data.

HOW THE TEAM SOLVED IT

Analysts and visual reviewers isolated failure modes, checked encodings and calculations, inspected renders, wrote structured critiques, and verified repairs.

DELIVERED

Source-to-render evaluation packs, failure labels, annotated critiques, corrected targets, accessibility checks, and batch-level insights.

DASHBOARDS // DOCUMENTS // UI

Professional visual-artifact tasks

CHALLENGE

Generic prompts did not reflect the layered constraints of real business work products or provide objective success evidence.

HOW THE TEAM SOLVED IT

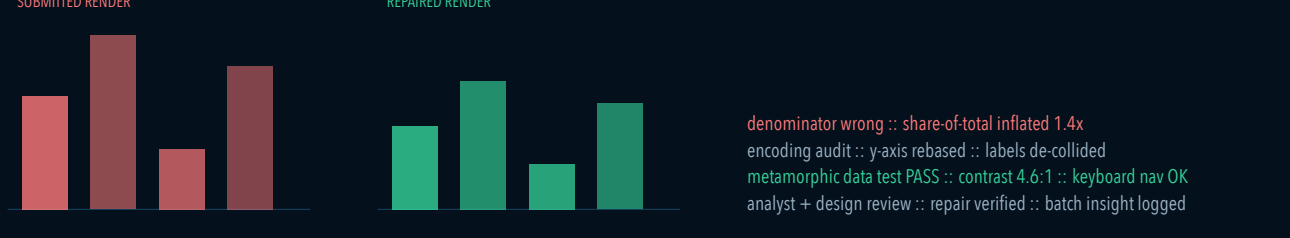
The team designed realistic briefs with source artifacts, functional and visual requirements, rubric dimensions, and reviewer guidance.

DELIVERED

Task briefs, context packages, reference criteria, rendered evidence, functional checks, review rubrics, and adjudicated scores.

viz-eval // artifact VIZ-0071 // source-to-render comparison

DELIVERY GLIMPSE // SANITIZED



// 20 // SELECTED ENGAGEMENTS

AGENTS INSIDE REAL SOFTWARE.

Programs test whether an agent can interpret interfaces, use tools, preserve state, and complete work under operational constraints.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

TOOL USE // TRAJECTORIES // HUMAN REVIEW

Agentic task and trajectory evaluation

CHALLENGE

A final answer concealed whether the agent took unsafe actions, ignored evidence, or reached the result through a brittle sequence.

HOW THE TEAM SOLVED IT

Reviewers inspected step-level trajectories, tool calls, state changes, recovery behavior, and policy compliance against task-specific rubrics.

DELIVERED

Environment-ready tasks, trace annotations, action-level critiques, success evidence, policy labels, and failure-mode summaries.

WEB APP // FUNCTIONAL QA // STATE

Complex application workflow reconstruction

CHALLENGE

Visual resemblance alone could not prove that reconstructed interfaces preserved navigation, validation, responsiveness, and user-flow behavior.

HOW THE TEAM SOLVED IT

Engineers decomposed flows, mapped states, inspected assets, specified observable behaviors, and combined functional testing with visual review.

DELIVERED

Flow maps, implementation briefs, functional test cases, visual checkpoints, issue classifications, and acceptance reports.

trajectory-review // trace AGT-2114

DELIVERY GLIMPSE // SANITIZED

```
$ trace AGT-2114 :: multi-tool task :: web app + spreadsheet + mail draft
step 04 :: tool=browser.click :: target=export-csv :: state hash 9f21 :: OK
step 07 :: tool=sheets.formula :: SUMIFS region mismatch :: self-corrected
step 11 :: tool=mail.send :: held by policy gate :: human approval requested
task complete :: 14 steps :: 0 unsafe actions :: trajectory annotated for training
```

// 21 // SELECTED ENGAGEMENTS

COMPUTER-USE AGENTS, MEASURED END TO END.

Desktop and browser agents are evaluated inside instrumented environments where every window, dialog, and state transition is observable.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

COMPUTER USE // DESKTOP // LONG-HORIZON

Operating-system level agent evaluation

CHALLENGE

Agents had to operate real desktop software across windows, files, and dialogs, where single-screenshot benchmarks hide brittleness.

HOW THE TEAM SOLVED IT

The team built instrumented desktop environments, task graphs, screen-state checkpoints, action-level review, and failure taxonomies for long-horizon work.

DELIVERED

Environment images, task packs, checkpoint evidence, action critiques, recovery examples, and reliability reports.

WEB AGENTS // DOM + VISION

Browser-native workflow trajectories

CHALLENGE

Production web tasks involve authentication boundaries, dynamic DOM, and ambiguous interface affordances that break naive agents.

HOW THE TEAM SOLVED IT

Designers built realistic multi-site workflows, captured synchronized DOM and pixel observations, and annotated affordances and recovery behavior.

DELIVERED

Dual-modality traces, affordance labels, recovery libraries, task-difficulty maps, and evaluation slices.

desktop-run // task invoice-reconcile-14

DELIVERY GLIMPSE // SANITIZED

```
$ envrun -task invoice-reconcile-14 -record dom+pixels -checkpoints on
00:12 open vendor portal :: session restored :: table extracted (14 rows)
01:03 cross-check ERP totals :: delta found at row 9 :: flagged for entry
01:41 unexpected modal dialog :: agent recovers via dismiss + re-query
02:24 reconciliation draft saved :: human approval queued :: trace sealed
```

// 22 // SELECTED ENGAGEMENTS

ENTERPRISE AGENTS WITH CONTROL BUILT IN.

The goal is the highest useful autonomy compatible with policy, evidence, reversibility, and accountable human ownership.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

AGENTIC AI // APPROVALS // AUDIT

Constrained enterprise workflow agent

CHALLENGE

An enterprise wanted agents to research, reconcile, draft, and update systems, but uncertain data and irreversible actions made naive automation unsafe.

HOW THE TEAM SOLVED IT

The team mapped tool permissions, introduced retrieval and state, set confidence gates, required approval for high-impact actions, and evaluated complete traces.

DELIVERED

Tool contracts, orchestration logic, approval checkpoints, evaluation suites, observability events, and an auditable action history.

PRIVATE DEPLOYMENT // DOMAIN DATA // GOVERNANCE

Secure onsite model adaptation

CHALLENGE

Sensitive workflows required domain-specific behavior without exposing proprietary records or relying on uncontrolled external services.

HOW THE TEAM SOLVED IT

The solution isolated data, built a targeted evaluation baseline, prepared governed adaptation data, validated regressions, and designed monitored deployment.

DELIVERED

A private evaluation set, adaptation corpus, model-quality report, deployment controls, operating runbook, and retention policy.

action-ledger // agent finance-ops // 24 h window

DELIVERY GLIMPSE // SANITIZED

AGT 0081 :: reconcile invoice :: evidence age 3 min :: risk low :: auto-committed
AGT 0116 :: modify account record :: evidence age 11 h :: risk high :: held for approval
AGT 0184 :: customer note :: risk medium :: approved by ops-lead :: sent
AGT 0237 :: delete record :: prohibited class :: blocked :: incident logged

// 23 // SELECTED ENGAGEMENTS

SOVEREIGN AND REGULATED DEPLOYMENTS.

Mission and compliance boundaries are treated as design inputs: capability is delivered inside the perimeter, with evidence that it stayed there.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

GOVERNMENT // AIR-GAPPED // ONSITE MODELS

Sovereign mission model deployment

CHALLENGE

Mission workflows demanded frontier-model capability with zero external service dependency and zero data egress.

HOW THE TEAM SOLVED IT

Compact onsite models in the 3B-parameter class were adapted on governed domain data inside isolated enclaves, with human-in-command controls and complete audit capture.

DELIVERED

Enclave architecture, adaptation corpus, evaluation baseline, deployment runbook, operator training, and audit trail.

BFSI // HEALTHCARE // COMPLIANCE

Regulated-industry private adaptation

CHALLENGE

Regulated records could not leave the client boundary, yet the model needed deep domain behavior and stable regressions.

HOW THE TEAM SOLVED IT

A split-trust architecture kept data in place: on-premise fine-tuning, privacy-aligned data handling, and regression plus safety gates before each release.

DELIVERED

Private evaluation sets, adaptation pipelines, compliance evidence packs, release gates, and a monitored rollout plan.

enclave-status // deployment sovereign-07

DELIVERY GLIMPSE // SANITIZED

```
$ enclave status -deployment sovereign-07
model: 3B-class onsite :: quantized :: zero external egress verified
data plane: isolated VLAN :: HSM-backed keys :: retention 90 d policy
policy conformance: 100% of 1,244 audited actions within mandate
quarterly red-team: 0 critical findings open :: report filed
```

// 24 // SELECTED ENGAGEMENTS

A VLA PILOT DESIGNED TO LEARN SAFELY.

The robotics frontier demands synchronized evidence across perception, language, state, action, outcome, and recovery.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

VLA // TELEOPERATION // ROBOT LEARNING

Industrial manipulation data pilot

CHALLENGE

A long-tail manipulation task had sparse demonstrations, variable workspaces, limited recovery coverage, and no reliable bridge from simulation to operations.

HOW THE TEAM SOLVED IT

The program defined success and safety, captured expert demonstrations, aligned video, state, and action, annotated phases, expanded scenarios in simulation, and evaluated behind interlocks.

DELIVERED

A task ontology, synchronized trajectories, recovery examples, quality reports, simulation scenarios, policy evaluation, and a scale-gate dossier.

VISION // ACTION // SAFETY

Embodied procedure evaluation

CHALLENGE

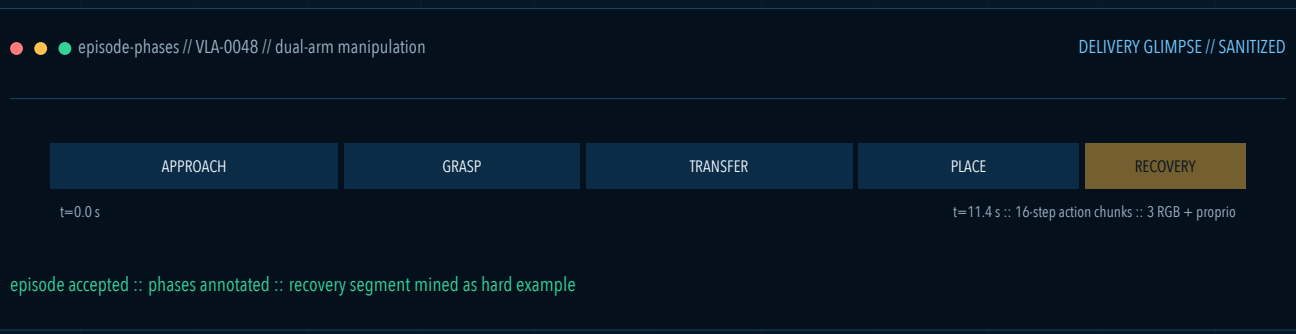
A policy could complete easy demonstrations yet fail under novel objects, occlusion, timing variation, or operator intervention.

HOW THE TEAM SOLVED IT

The team designed scenario slices, instrumented safety events, measured state coverage, recorded human overrides, and compared real and simulated behavior.

DELIVERED

Scenario suites, coverage maps, safety-event analyses, operator feedback, distribution-shift results, and an operational-envelope recommendation.



// 25 // SELECTED ENGAGEMENTS

CROSS-EMBODIMENT DATA FOR ROBOTICS ECOSYSTEMS.

Fast-moving robotics platform teams need demonstrations that transfer across arms, cameras, grippers, and control stacks.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

ROBOTICS PLATFORM // VLA DATA

Cross-embodiment manipulation dataset

CHALLENGE

Demonstrations from one arm, camera layout, or gripper did not transfer cleanly to another embodiment, and task labels lacked physical-state context.

HOW THE TEAM SOLVED IT

The team normalized observations and actions, aligned coordinate frames, annotated task phases and contacts, documented embodiment metadata, and created transfer-focused evaluation slices.

DELIVERED

Synchronized trajectories, embodiment descriptors, normalized action representations, phase and contact labels, transfer benchmarks, and a governed release package.

DUAL-ARM // WAREHOUSE // LONG-HORIZON

Bimanual warehouse policy evaluation

CHALLENGE

A bimanual policy performed isolated grasps but became brittle across hand-offs, packing sequences, clutter, and recovery from partial failure.

HOW THE TEAM SOLVED IT

Reviewers decomposed long-horizon tasks, captured coordinated demonstrations, expanded failure and recovery states, and scored policy behavior across workspace variations.

DELIVERED

Bimanual task graphs, coordinated trajectories, recovery libraries, scenario matrices, safety annotations, and policy-level evaluation reports.

alignment-check // episode VLA-0107 // target 6-axis

DELIVERY GLIMPSE // SANITIZED

```
$ align check -episode VLA-0107 -source cobot -target 6-axis
frame map: wrist cam to TCP :: residual 2.1 mm :: within kinematic gate
gripper semantics: vacuum to parallel jaw :: remapped :: action dim 7 to 8
transfer eval: 84 of 96 layouts PASS :: 9 recovered :: 3 flagged for review
```

// 26 // INDUSTRIAL ARMS

FIVE ARM CLASSES. ONE DATA DISCIPLINE.

Capability spans the five widely used industrial arm classes, delivered with a vendor-neutral data discipline that travels across brands and controllers.

OWNER-CONFIRMED CAPABILITY | VENDOR-NEUTRAL | NDA-SAFE

WELDING // ASSEMBLY // MATERIAL HANDLING

Articulated / 6-axis arm programs

CHALLENGE

High freedom of motion creates complex reach, collision, sequencing, and recovery states across variable industrial workspaces.

HOW THE TEAM SOLVED IT

The team structured joint-aware task phases, workspace constraints, failure and recovery labels, expert review, and evidence-driven acceptance.

DELIVERED

Trajectory sets, phase labels, exception cases, safety annotations, quality evidence, and evaluation slices.

ELECTRONICS // HIGH-SPEED PICK-PLACE

SCARA arm programs

CHALLENGE

Fast planar motion requires precise cycle logic, compact-workspace coverage, stable orientation, and defensible placement tolerances.

HOW THE TEAM SOLVED IT

Reviewers focused task definitions on repeatability, object and placement state, boundary cases, error severity, and cycle-level evidence.

DELIVERED

Calibrated task packs, placement and error labels, timing evidence, review records, and acceptance reports.

cell-telemetry // arm scara-04 // shift 2

DELIVERY GLIMPSE // SANITIZED

```
$ cell telemetry -arm scara-04 -window shift-2
cycle p50 1.9 s :: p95 2.4 s :: placement tolerance +/-0.15 mm held
tray 118 :: pick miss :: vision re-localize :: recovered within 1 cycle
shift acceptance: 99.2% first-pass :: 0 safety events :: evidence appended
```

// 27 // INDUSTRIAL ARMS

HIGH-SPEED, LARGE-SCALE, AND HUMAN-AWARE.

Three additional arm classes extend the same data discipline across conveyor systems, gantry workspaces, and collaborative cells.

FOOD // PHARMA // PACKAGING

Delta

CHALLENGE

Conveyor speed and small objects amplify missed-pick and tracking errors.

HOW THE TEAM SOLVED IT

The program aligned visual events, object state, pick outcomes, and recovery behavior.

DELIVERED

Event-aligned examples, failure taxonomies, inspection evidence, and scenario matrices.

PALLETIZING // LARGE WORKSPACES

Cartesian / gantry

CHALLENGE

Linear axes introduce range, sequencing, and workspace-coverage constraints.

HOW THE TEAM SOLVED IT

The team decomposed paths, validated coordinates, tagged boundaries, and reviewed completion.

DELIVERED

Path specifications, coordinate checks, boundary cases, and quality-controlled releases.

HUMAN-ROBOT WORKCELLS

Collaborative arms

CHALLENGE

Human proximity makes interruption and safe recovery as important as task success.

HOW THE TEAM SOLVED IT

Reviewers captured interaction states, overrides, safe stops, and operator acceptance.

DELIVERED

Interaction scenarios, safety-event labels, override traces, and operating-envelope reports.

SHARED DELIVERY ARCHITECTURE

Task ontology / synchronized evidence / independent review / governed release

Across all five arm classes, the differentiator is not the robot brand. It is the ability to make task intent, physical state, action, failure, recovery, and quality evidence usable for training and evaluation.

safety-ledger // cell cobot-2 // week 27

DELIVERY GLIMPSE // SANITIZED

```
$ safety ledger -cell cobot-2 -week 27
safe stops: 3 :: human-proximity slowdowns: 41 :: all within envelope
override event WED 14:02 :: operator correction captured as recovery demo
zero reportable incidents :: interaction dataset appended +312 episodes
```

// 28 // SELECTED ENGAGEMENTS

SIMULATION AT SCALE, BEFORE THE FACTORY FLOOR.

Millions of simulated rollouts harden policies against variation that physical capture alone can never cover.

ANONYMIZED ENGAGEMENTS // CLIENT IDENTITIES PROTECTED UNDER NDA

ISAAC LAB // COSMOS // DOMAIN RANDOMIZATION

Massive-scale policy hardening in simulation

CHALLENGE

Real capture alone could not cover the object, lighting, layout, and disturbance variation required for robust deployment.

HOW THE TEAM SOLVED IT

Scenario generation across Isaac Lab and Cosmos-class world models produced millions of randomized rollouts, trained diffusion-policy heads, and sliced results by distribution shift.

DELIVERED

Scenario libraries, rollout telemetry, hardened policy checkpoints, failure clusters, and sim-to-real gap reports.

TRANSFER // STRESS TESTS // INTERLOCKS

Sim-to-real evaluation gates

CHALLENGE

Policies that excel in simulation regress on hardware in ways that are expensive to discover late.

HOW THE TEAM SOLVED IT

Matched simulated and physical scenario suites, staged evaluation gates behind safety interlocks, and calibrated realism audits localized every gap before scale-up.

DELIVERED

Gate criteria, paired sim and real metrics, failure attribution reports, and scale-decision dossiers.

rollout-batch // policy grasp-v41

DELIVERY GLIMPSE // SANITIZED

```
$ simbatch -policy grasp-v41 -rollouts 2.4M -randomize full
success sim 93.1% :: unseen layouts 88.4% :: torque violations 0.02%
failure cluster 7: transparent objects :: routed to targeted real capture
sim-to-real gate: PASS at 86% threshold :: hardware evaluation scheduled
```

// 29 // HUMANOID TELEOPERATION

THE HUMAN OPERATIONS LAYER FOR HUMANOID ROBOTS.

Softinator.ai provides workforce and program operations for humanoid teleoperation, VR control, demonstration capture, and trajectory-quality review.

OWNER-CONFIRMED CAPABILITY | VR CONTROL | HUMANOID OPERATIONS | HUMAN-IN-COMMAND

OPERATOR SOURCING // QUALIFICATION // SHIFT DELIVERY

Humanoid robot teleoperation workforce

CHALLENGE

Robotics teams need reliable operators who can learn unfamiliar control interfaces, execute repeatable tasks, recognize unsafe states, and sustain consistent coverage.

HOW THE TEAM SOLVED IT

Softinator.ai recruits against dexterity and technical criteria, runs device and task calibration, qualifies operators on shared scenarios, schedules coverage, monitors quality, and routes incidents.

DELIVERED

Qualified operator pools, calibration records, coverage plans, shift reporting, escalation logs, retraining loops, and workforce performance summaries.

SPATIAL CONTROL // DATA CAPTURE // TRAJECTORY QA

VR-controlled demonstration operations

CHALLENGE

Raw VR sessions can contain drift, dropped frames, poor task execution, unsafe motion, incomplete recoveries, or mismatched sensor and action timelines.

HOW THE TEAM SOLVED IT

The team standardizes setup checks, validates spatial calibration, guides task execution, monitors capture health, reviews trajectories, tags failure and recovery, and approves governed releases.

DELIVERED

Session protocols, device-readiness checks, synchronized recordings, operator notes, quality flags, recovery examples, and release-level lineage.

session-monitor // rig hxr-03 // XR hand tracking

DELIVERY GLIMPSE // SANITIZED

```
$ session monitor -rig hxr-03 -mode xr-hands
p95 end-to-end 42 ms :: wrist drift 4 mm :: frame drop 0.1%
retarget: SE3 wrists + 21-DoF hands :: collision margins nominal
2 h 40 m captured :: 96 episodes :: 12 recovery demos :: MCAP sealed
```

// TECHNICAL DOSSIER 01

DOMAIN CODE FOUNDATION MODEL.

Representative architecture for corpus engineering and distributed pretraining of a dense code model. No named client, private model, benchmark result, or confidential tool is disclosed.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

32B dense decoder with GQA, RoPE, SwiGLU, and 128K context

ENGINEERING STACK

Megatron Core; Megatron FSDP; NeMo Curator; NCCL; FlashAttention 3; DTensor checkpoints; Slurm; WebDataset

LEARNING AND VALIDATION

Curriculum sampling; exact and MinHash deduplication; semantic deduplication; contamination probes; token quality classifiers

CHALLENGE

Code corpora contained duplicate forks, generated noise, license ambiguity, benchmark leakage, and severe domain imbalance.

TECHNICAL RESPONSE

Build a provenance graph, normalize repositories, fingerprint functions and files, score quality, isolate evaluation families, and train with topology aware parallelism.

DELIVERED

Governed corpus releases, tokenizer and sampling policy, training manifests, checkpoint lineage, loss slice dashboards, and evaluation packs.

SYNTHETIC CORPUS MANIFEST ROWS

shard	domain	quality signal	contamination gate	release state
CRX 0041	systems C++	AST entropy 0.81	exact and fuzzy clear	accepted
CRX 0197	Python tools	doc code alignment 0.93	benchmark neighbor	quarantined
CRX 0318	Java services	compile proxy pass	license review clear	accepted
CRX 0442	CUDA kernels	expert utility 4.7 of 5	semantic overlap 0.04	accepted
CRX 0509	Rust systems	clippy and borrow clean	exact and fuzzy clear	accepted
CRX 0577	Go services	vet and race clean	license review clear	accepted

// TECHNICAL DOSSIER 02

VERIFIABLE REASONING POST TRAINING.

Representative post training system for a reasoning model using expert supervision, verifiable rewards, preference optimization, and controlled evaluation.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

14B or 32B reasoning model with long context and structured outputs

ENGINEERING STACK

OpenRLHF; Ray; vLLM; SGLang; DeepSpeed ZeRO 3; PyTorch distributed checkpoints; sandboxed verifiers

LEARNING AND VALIDATION

SFT; rejection sampling; DPO; GRPO style optimization; process rewards; KL control; reward hacking probes

CHALLENGE

Outcome labels alone could reward lucky answers, brittle formatting, or exploitable verifier behavior.

TECHNICAL RESPONSE

Separate process and outcome evidence, execute deterministic checks, calibrate expert critiques, track KL drift, and adversarially test reward surfaces.

DELIVERED

Training mixtures, reward specifications, verifier registry, preference pairs, reasoning trace audits, ablation reports, and release gates.

SYNTHETIC POST TRAINING BATCH ROWS

batch	task family	reward vector	stability check	disposition
RFT 021	symbolic math	1.00 / 0.86 / 0.72	KL 0.031	train
RFT 088	code repair	1.00 / 0.91 / 0.80	verifier replay pass	train
RFT 143	scientific QA	0.00 / 0.77 / 0.69	answer mismatch	revise
RFT 207	tool planning	1.00 / 0.83 / 0.88	policy trace clear	train

// TECHNICAL DOSSIER 03

REPOSITORY ENGINEERING AGENT.

Representative system for issue localization, patch generation, test construction, and auditable repository level evaluation.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Model agnostic coding agent with planner, navigator, editor, and verifier roles

ENGINEERING STACK

Tree sitter; SCIP or LSIF indexes; Semgrep; CodeQL; Firecracker microVMs; Nix or Bazel environments; pytest, cargo test, go test, and language native runners

LEARNING AND VALIDATION

Graph retrieval; fault localization; constrained patching; self generated tests; execution feedback; patch minimality scoring

CHALLENGE

Agents could produce plausible patches while reading the wrong files, changing unrelated behavior, or exploiting weak tests.

TECHNICAL RESPONSE

Instrument repository exploration, restrict writable scope, require reproduction, execute hidden and generated tests, inspect static analysis, and score minimality.

DELIVERED

Issue packs, hermetic environments, trace logs, patch diffs, test evidence, security findings, and adjudicated resolution labels.

SYNTHETIC REPOSITORY AGENT TRACE ROWS

issue	localization	patch scope	verification	verdict
SWE 1042	3 of 3 files	18 lines	unit plus regression pass	resolved
SWE 1198	wrong module	no patch	reproduction failed	reroute
SWE 1314	2 of 4 files	41 lines	CodeQL alert introduced	reject
SWE 1461	4 of 4 files	23 lines	hidden tests pass	resolved

// TECHNICAL DOSSIER 04

MULTIMODAL VISUALIZATION VERIFIER.

Representative evaluation and repair pipeline for models that generate analytical charts, dashboards, and interactive visual artifacts.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Vision language model plus code generation and render critic loop

ENGINEERING STACK

Playwright; Vega Lite grammar; Chromium traces; DOM snapshots; OCR; perceptual image metrics; source data validators

LEARNING AND VALIDATION

Schema constrained decoding; render execute critique; semantic chart checks; metamorphic data tests; accessibility and interaction validation

CHALLENGE

A visualization could render successfully while encoding the wrong measure, hiding uncertainty, breaking interaction, or violating accessibility.

TECHNICAL RESPONSE

Compare source data to chart grammar, replay interactions, inspect DOM and canvas output, run metamorphic transformations, and combine analyst and design review.

DELIVERED

Source to render packs, chart specifications, failure taxonomy, interaction traces, repaired outputs, accessibility evidence, and quality summaries.

SYNTHETIC VISUALIZATION EVALUATION ROWS

artifact	data fidelity	visual signal	interaction	final state
VIZ 033	aggregate exact	label collision	filter pass	repair
VIZ 071	denominator wrong	plausible render	tooltip pass	reject
VIZ 118	source aligned	contrast 3.1 to 1	keyboard fail	repair
VIZ 206	source aligned	hierarchy pass	responsive pass	accepted

// TECHNICAL DOSSIER 05

CONTROLLED ENTERPRISE AGENT.

Representative architecture for tool using agents that operate across records, workflows, and approval boundaries.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Hierarchical planner with retrieval, policy checks, durable state, and human approval

ENGINEERING STACK

Temporal workflows; OPA and Rego policies; OpenTelemetry traces; PostgreSQL event store; isolated tool adapters; secrets broker

LEARNING AND VALIDATION

Constrained tool decoding; typed action schemas; confidence gates; compensating actions; state machine validation; trace based evaluation

CHALLENGE

Unrestricted agents could overreach permissions, act on stale evidence, repeat side effects, or make irreversible changes.

TECHNICAL RESPONSE

Wrap every tool with typed contracts, evaluate evidence freshness, apply policy before execution, require approval at risk boundaries, and log reversible state transitions.

DELIVERED

Tool registry, policy bundle, workflow graphs, approval UX, simulation suite, audit events, incident playbooks, and operating runbook.

SYNTHETIC AGENT ACTION LEDGER

trace	proposed action	evidence age	risk gate	outcome
AGT 0081	reconcile invoice	3 minutes	low, automatic	committed
AGT 0116	modify account	11 hours	high, approval	held
AGT 0184	send customer note	6 minutes	medium, review	approved
AGT 0237	delete record	1 minute	prohibited	blocked

// TECHNICAL DOSSIER 06

VISION LANGUAGE ACTION POLICY.

Representative VLA adaptation system for long horizon manipulation using real demonstrations, simulation expansion, and edge aware evaluation.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

OpenVLA OFT class policy with multi image input and action chunking

ENGINEERING STACK

NVIDIA Isaac Sim; Isaac Lab; Isaac Lab Mimic; robomimic; RLDS or LeRobot datasets; TensorRT; ROS 2

LEARNING AND VALIDATION

Behavior cloning; LoRA or full tuning; flow or diffusion action heads; temporal ensembling; domain randomization; recovery curriculum

CHALLENGE

Sparse demonstrations and workspace variation caused compounding errors across grasp, transfer, placement, and recovery.

TECHNICAL RESPONSE

Segment demonstrations by object referenced subtasks, normalize task space actions, generate spatial variations, train chunked policies, and evaluate unseen layouts.

DELIVERED

Task ontology, synchronized episodes, simulation variants, policy checkpoints, failure slices, recovery library, and deployment envelope.

SYNTHETIC VLA EPISODE ROWS

episode	embodiment	observation	action target	result
VLA 0048	dual arm	3 RGB plus state	16 step chunk	success
VLA 0107	6 axis arm	wrist plus overhead	SE3 delta	recovery used
VLA 0194	humanoid upper body	stereo plus proprio	bimanual pose	collision gate
VLA 0268	cobot	RGBD plus force	pose plus gripper	success

// TECHNICAL DOSSIER 07

HUMANOID XR TELEOPERATION.

Representative operations and data system for immersive humanoid control, high fidelity capture, retargeting, and trajectory release.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Bimanual humanoid teleoperation with egocentric vision and hand tracking

ENGINEERING STACK

Isaac Teleop; NVIDIA CloudXR; OpenXR; ROS 2; MCAP; PTP clock sync; Pinocchio; cuRobo; Ruckig

LEARNING AND VALIDATION

SE3 retargeting; inverse kinematics; collision constraints; latency and drift filters; subtask annotation; Isaac Lab Mimic augmentation

CHALLENGE

Operator motion, robot kinematics, network delay, tracking drift, and safety constraints had to remain synchronized across long sessions.

TECHNICAL RESPONSE

Calibrate devices and body frames, retarget wrists and hands, monitor end to end latency, capture synchronized topics, flag anomalies, and review recovery behavior.

DELIVERED

Qualified operator pool, calibration profiles, session protocols, MCAP releases, subtask labels, latency reports, and safety event ledger.

SYNTHETIC HUMANOID TELEOPERATION TELEMETRY

session	control mode	p95 latency	tracking quality	release
HXR 014	XR hand tracking	42 ms	wrist drift 4 mm	accepted
HXR 029	spatial controllers	55 ms	frame drop 0.3 percent	accepted
HXR 044	XR hand tracking	91 ms	left hand loss	quarantine
HXR 061	bimanual replay	offline	sync error 1.2 ms	accepted

// TECHNICAL DOSSIER 08

CROSS EMBODIMENT ARM ADAPTATION.

Representative transfer system spanning articulated, SCARA, delta, Cartesian, and collaborative robot classes.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Embodiment conditioned manipulation policy with task space actions

ENGINEERING STACK

MoveIt 2; OMPL; Tesseract; cuRobo; ROS 2 Control; Isaac Sim; URDF and USD assets; MCAP

LEARNING AND VALIDATION

Coordinate frame normalization; forward and inverse kinematics; collision aware planning; residual adaptation; domain randomization; transfer evaluation

CHALLENGE

Demonstrations encoded robot specific joints, controllers, camera frames, and gripper semantics that did not transfer directly.

TECHNICAL RESPONSE

Convert trajectories into shared task space, attach embodiment metadata, validate kinematics, map gripper actions, simulate variation, and measure transfer gaps.

DELIVERED

Embodiment registry, normalized trajectories, controller adapters, simulation assets, transfer benchmarks, exception cases, and release documentation.

SYNTHETIC CROSS EMBODIMENT ALIGNMENT ROWS

arm class	degrees of freedom	action frame	kinematic gate	state
articulated	6 plus gripper	tool center point	residual 2.1 mm	aligned
SCARA	4 plus vacuum	workcell XYZR	residual 0.8 mm	aligned
delta	parallel 3 axis	conveyor frame	timing drift 6 ms	aligned
cobot	7 plus gripper	base SE3	collision margin fail	review

// TECHNICAL DOSSIER 09

MULTILINGUAL STEM FOUNDATION MODEL.

Representative pretraining and post training system for mathematics, science, engineering, and structured technical reasoning.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Multilingual decoder model with notation aware tokenizer and long context

ENGINEERING STACK

SentencePiece; Megatron Core; NeMo Curator; LaTeX and MathML parsers; Lean 4 checks; SymPy; distributed evaluation

LEARNING AND VALIDATION

Language balanced curriculum; proof and calculation validators; process supervision; difficulty calibration; preference optimization; contamination audits

CHALLENGE

Technical corpora mixed malformed notation, translation artifacts, duplicated textbook solutions, unverifiable derivations, and uneven language coverage.

TECHNICAL RESPONSE

Parse equations and units, classify derivation quality, validate computable steps, route specialist review, balance languages, and hold out concept families.

DELIVERED

Curated multilingual corpus, notation policy, expert review network, validator suite, difficulty map, training mixtures, and evaluation dashboards.

SYNTHETIC STEM REASONING ROWS

sample	domain	notation gate	validator	decision
STM 102	control theory	LaTeX valid	symbolic residual zero	accepted
STM 188	organic chemistry	reaction parse pass	expert only	accepted
STM 263	statistics	unit mismatch	numeric check fail	repair
STM 317	electromagnetics	MathML valid	boundary check pass	accepted

// TECHNICAL DOSSIER 10

ADVERSARIAL MODEL EVALUATION.

Representative red team and assurance program for models, agents, and retrieval systems operating under enterprise policy.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Model and agent assurance harness with policy aware attack generation

ENGINEERING STACK

PyRIT; garak; NeMo Guardrails; OPA and Rego; isolated tool mocks; OpenTelemetry; evidence store

LEARNING AND VALIDATION

Threat modeling; metamorphic tests; multi turn attacks; indirect prompt injection; data exfiltration probes; risk weighted scoring

CHALLENGE

Static safety prompts missed tool abuse, retrieval poisoning, delayed attacks, policy conflicts, and cross turn state manipulation.

TECHNICAL RESPONSE

Generate attacks from a threat taxonomy, execute in isolated environments, preserve complete traces, require reproducibility, and route severe findings to human adjudication.

DELIVERED

Threat model, attack corpus, reproducible traces, severity rubric, mitigation tests, regression suite, and executive risk register.

SYNTHETIC ASSURANCE FINDING ROWS

finding	attack vector	control tested	severity	disposition
SEC 041	indirect injection	retrieval boundary	high	mitigated
SEC 077	tool parameter smuggling	typed schema	critical	blocked
SEC 126	memory poisoning	state isolation	medium	monitor
SEC 209	encoded exfiltration	output policy	high	repair

// TECHNICAL DOSSIER 11

AI CLUSTER ORCHESTRATION.

Representative operating architecture for large model training, checkpoint durability, network telemetry, and GPU fleet efficiency.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

Multi tenant accelerator cluster for pretraining, post training, simulation, and inference

ENGINEERING STACK

Slurm; Kubernetes; NCCL; GPUDirect RDMA; InfiniBand; DCGM; Prometheus; Ceph; NVMe over Fabrics; immutable containers

LEARNING AND VALIDATION

Topology aware placement; elastic recovery; checkpoint scheduling; congestion detection; power capping; thermal and straggler analysis

CHALLENGE

Large jobs could lose expensive compute time through link degradation, stragglers, checkpoint stalls, thermal limits, or noisy neighbors.

TECHNICAL RESPONSE

Schedule by network topology, profile collectives, monitor fabric and GPU health, stage checkpoints, isolate faults, and enforce workload specific service classes.

DELIVERED

Cluster blueprint, scheduler policy, telemetry schema, failure drills, checkpoint design, capacity model, and operating dashboards.

SYNTHETIC ACCELERATOR JOB TELEMETRY

job	topology	collective signal	thermal state	scheduler action
TRN 018	64 GPU fat tree	all reduce 91 percent	nominal	continue
TRN 026	128 GPU rail aware	link retry burst	nominal	migrate rank
SIM 044	32 GPU island	I/O wait 14 percent	nominal	stage dataset
RFT 071	16 GPU island	straggler 1.28x	power limited	rebalance

// TECHNICAL DOSSIER 12

EGOCENTRIC PHYSICAL REASONING MODEL.

Representative foundation model for temporal understanding of first person video, object state, contact, and future action.

REFERENCE ARCHITECTURE // CLIENT-SPECIFIC DETAILS WITHHELD UNDER NDA

MODEL SYSTEM

3B video language backbone with object centric memory and future state heads

ENGINEERING STACK

NVIDIA DALI; PyTorchVideo; WebDataset; SigLIP class encoders; DINOv2 features; SAM 2 masks; optical flow and tracking

LEARNING AND VALIDATION

Masked video modeling; temporal contrastive learning; contact prediction; future state forecasting; hard negative mining; multi view alignment

CHALLENGE

Long egocentric video contained occlusion, hand object ambiguity, weak temporal boundaries, and rare but critical failures.

TECHNICAL RESPONSE

Track objects and hands, align multi view clips, infer contact phases, mine hard negatives, supervise future state, and evaluate long horizon temporal consistency.

DELIVERED

Video corpus, temporal annotations, object and contact tracks, training objectives, retrieval index, failure slices, and benchmark harness.

SYNTHETIC PHYSICAL REASONING CLIP ROWS

clip	view set	contact phase	temporal target	quality state
EGO 0039	head plus wrist	approach to grasp	next state 0.5 s	accepted
EGO 0112	stereo head	occluded transfer	contact recovery	expert review
EGO 0186	head plus overhead	tool engagement	force event proxy	accepted
EGO 0244	wrist only	object loss	failure forecast	accepted

// 30 // INFRASTRUCTURE DIRECTION

SOVEREIGN COMPUTE FOR THE NEXT WAVE.

Softinator is advancing an AI-ready data center initiative near the Mumbai / Navi Mumbai / JNPT region: a phased program to bring model development, data residency, and robotics simulation closer together.

COMPUTE PLATFORM

Designed around AI workloads

GPU lifecycle strategy, high-density electrical design, liquid-cooling readiness, resilient connectivity, secure tenancy, observability, and phased capacity planning.

STRATEGIC VALUE

A closer link between data and compute

Private model development, controlled data residency, lower-friction experimentation, sovereign workloads, robotics simulation, and enterprise deployment support.

PHASE 01

Diligence

Site, utilities, connectivity, policy, commercial demand, and technical basis.

PHASE 02

Pilot capacity

Validated architecture, anchor workloads, operations, security, and financing.

PHASE 03

Scaled campus

Phased expansion aligned to power, cooling, utilization, and capital gates.

STRATEGIC INITIATIVE | PHASED CAPACITY | SOVEREIGN BY DESIGN

// 31 // CREDENTIALS

CERTIFIED. AUDITED. ACCOUNTABLE.

Four decades of institutional discipline back every engagement. Audit reports of compliances, patents, and certifications are available on request.

CERTIFICATIONS

ISO 9001:2015	ISO 27001:2022	SOC 2 Type II	CMMI Level 3
IAF ACCREDITED AND AUDITED	IAF ACCREDITED AND AUDITED	ATTESTED CONTROLS	IN PROGRESS

RECOGNITIONS

Est. 1981	Startup India	NASSCOM	MSME
44+ YEARS OF TECH EXCELLENCE	DPIIT RECOGNIZED	REGISTERED MEMBER	REGISTERED

CORPORATE IDENTIFIERS

D-U-N-S® 650929933 / 813837313
GSTN 23AAWCS1961M1ZP
Softinator® is a registered trademark

GROUP ENTITIES

Softinator® TechLabs Pvt. Ltd.
Softinator® Dynamics Pvt. Ltd.
Softinator ERP LLP //
Softinator Inc. (US)

CERTIFICATIONS ON THE CAMPUS SIGNAGE, INDORE

SECURITY ARCHITECTURE BUILT FOR

HIPAA | PCI-DSS | HITRUST | CCPA / CPRA | COPPA | GDPR | NIST

Least-privilege access, workstream isolation, role-based review, controlled exports, client-defined retention, and no customer-to-project disclosure. NDA-governed delivery details are disclosed only to authorized parties through controlled diligence.

A GNATOR INC. (US) FRONTIER AI DEVELOPMENT COMPANY // GLOBAL PRESENCE: USA, AUSTRALIA, INDIA // CSR: PM'S EMERGENCY RELIEF FUND

// 32 // TALENT SYSTEM

A NETWORK BUILT FOR HARD PROBLEMS.

Our single greatest asset: a global expert network reaching hundreds of thousands of researchers, refined into calibrated program teams through governed qualification.

GLOBAL ACADEMIC REACH

HUNDREDS OF THOUSANDS OF RESEARCHERS VIA SISTER ACADEMIC-PUBLISHING COMPANIES IN THE USA, UK, AND AUSTRALIA

▼

47,000+ QUALIFIED EXPERTS

PHDS, PROFESSORS, AND ELITE DEVELOPERS ACROSS STEM, CODE, AND VISION DOMAINS

▼

CALIBRATED PROGRAM TEAMS

QUALIFICATION GATES, SHARED CALIBRATION, INDEPENDENT REVIEW

▼

300+ CORE DELIVERY WORKFORCE

DEVELOPERS, RESEARCHERS, QA, AND PROGRAM OPERATORS

OPERATING IDEA

Scale is not a headcount claim alone. The differentiator is the ability to match task difficulty to expertise, make review independent, measure disagreement, and preserve accountability from brief to release.

SOFTINATOR DELIVERY TEAMS AND SPECIALIST COHORTS, INDORE CAMPUS

// 33 // LEADERSHIP

THE MINDS BEHIND THE FOUNDRY.

IIT, IIM, Georgia Tech, Bloomberg, and BHEL pedigree, fused with decades of enterprise delivery: the management team steering Softinator's frontier pivot.

Suyog Dixit

FOUNDING VISIONARY & MD

PhD research in financial machine learning; advocate for innovation and automation.

Richa Dixit

CEO & OPERATIONAL LEAD

PGDBM and MBA; international business strategy with HRM expertise.

Dr. Rajkumar Jain

CHIEF TECHNOLOGY OFFICER

Rooted in IIT Indore's culture of innovation; shapes technological direction.

Suyash Dixit

DIRECTOR OF DATA INTELLIGENCE

Georgia Tech, USA; turns data into insights that inform strategic decisions.

Mrutunjay Tiwari

DATA SCIENCE PIONEER

IIT Bombay analytical excellence behind our data-centric innovations.

Ajay Joshi

MANAGING PARTNER

SAP-certified; three decades across HCM, finance, and ERP, including BHEL.

Rakesh Tiwari

ENGINEERING ARCHITECT

IIT Madras rigor with Bloomberg innovation; 30+ years of US engineering.

Dr. Rajesh Kumar Dixit

CHIEF PATRON

IIT Kanpur, RRCAT, and TIFR Bombay; guides our mission and vision.

Girish Vyas

PROGRAM EXECUTION SPECIALIST

IIM Kozhikode insights and UK delivery experience for seamless execution.

// 34 // INSIDE SOFTINATOR

BUILT BY PEOPLE WHO SHIP.

Behind every governed release is a working floor: engineers, reviewers, operators, and researchers sharing one campus culture and one quality bar.

THE COLLABORATIVE DEVELOPMENT FLOOR

SPATIAL COMPUTING AND VR TELEOPERATION LAB

SOFTINATOR® GROUP OFFICE, INDORE

NATIONAL MEDIA COVERAGE, INDIA TODAY NETWORK

JOIN THE FOUNDRY

We hire engineers, researchers, robotics operators, and program leaders who want their work to reach production. WhatsApp your résumé to +91 786 966 1171 or write to hr@softinator.ai.

// 35 // ENGAGE

START WITH THE BOTTLENECK.

A good first engagement is narrow enough to measure and meaningful enough to change the system.

DISCOVER

Capability workshop

Map the target behavior, constraints, current failure modes, evidence, and decision gates.

PROVE

Controlled pilot

Build a representative task slice, quality model, security boundary, and acceptance protocol.

SCALE

Production program

Expand only after quality, economics, throughput, governance, and operational ownership are visible.

// Say hello

business@softinator.ai

+91 88 27 111112

Sales, partnerships, and new engagements. We respond within one business day.

// WhatsApp your résumé

hr@softinator.ai

+91 786 966 1171

Careers across engineering, research, data operations, and robotics programs.

// Assistance awaits

support@softinator.ai

+91 786 966 1172

Delivery support, escalations, and operational continuity for active programs.

AT THE SOFTINATOR® GROUP OF INDUSTRIES OFFICE

ISO 9001:2015 | ISO 27001:2022 | SOC 2 TYPE II | DPIIT | NASSCOM | MSME | EST. 1981

SOFTINATOR.AI // SOFTINATOR.COM // GLOBAL PRESENCE: USA, AUSTRALIA, INDIA

// 36 // OUR CAMPUSES

LET'S BREW A CONVERSATION.

Five working campuses across Indore and Mumbai anchor the delivery network, from collaborative development floors to the frontier AI wing.

Collaborative DevSpace

Brainly Softinator 237, Fourth Floor,
Mangal City,
AB Road, Scheme 54 PU4, Indore
452010

Executive Office

Softinator Group, E-3130, Behind
Bank of India,
Near Gopur Square, Sudama Nagar,
Indore 452009

Collaborative EliteSpace

Softinator Dynamics, 08, 2nd Floor,
Apollo Premier,
Vijay Nagar Square, Vijay Nagar,
Indore 452010

Frontier AI Wing

Softinator TechLabs, 165 D & E, Vidur
Nagar,
Indore 452013

Headquarters, Mumbai

Softinator, 3, 2nd Floor, Ackruti Star Tower, Andheri East, Mumbai, Maharashtra 400093

SOFTINATOR ERP LLP AND SOFTINATOR® TECHLABS PRIVATE LIMITED, CIN: U72300MH2015PTC261302
SOFTINATOR® DYNAMICS PRIVATE LIMITED, CIN: U72900MP2021PTC058777
SOFTINATOR INC. ESTA. REG. NO.: C/97746 // GSTN: 23AAWCS1961M1ZP
DPIIT: DIPP119660 // MSME REGISTERED // D-U-N-S® 650929933 AND 813837313
CSR FUNDS DIRECTED TO THE PM'S EMERGENCY RELIEF FUND
SOFTINATOR GROUP, HEADQUARTERED IN INDIA, OPERATES AS A SUBSIDIARY OF GNATOR INCORPORATED, CHICAGO, USA.

GNator.AI
Crafting Tomorrow's Intelligence

softinator ®
TechLabs

REASONING IN THE PHYSICAL WORLD.

SOFTINATOR.AI

SOFTINATOR® TECHLABS PVT. LTD. // SOFTINATOR® DYNAMICS PVT. LTD. // SOFTINATOR ERP LLP // SOFTINATOR INC.

SOFTINATOR® GROUP OF INDUSTRIES // DIGITIZING DREAMS, EXPANDING HORIZONS.™