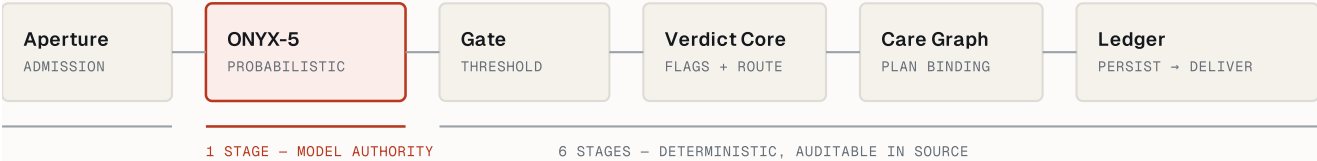


Constrained Vision Assessment for Nail Condition Reporting

Architecture, determinism guarantees and operating limits of the
NailScan Assessment Engine

FIGURE 1 – THE DETERMINISM BOUNDARY



NailScan is an assessment and marketing tool. It is not a medical device, does not diagnose, and is not offered or certified as a clinical instrument. Section 11 states the complete set of limitations and non-claims.

Abstract

NailScan is a production system that produces structured nail-condition assessments from a single consumer photograph. This report describes its architecture, and in particular the decision that separates it from the generative consumer tools occupying the same category: the probabilistic component is confined to one stage of a seven-stage pipeline, restricted to filling five observation fields against a versioned written specification, and gated on a self-reported confidence threshold. Every consequential output — referral recommendation, care plan, urgency of language — is derived from those fields by deterministic code that can be read in a few minutes. We describe the observation schema (ONYX-5), the admission and gating layers, the deterministic derivation layer (Verdict Core), the versioned care library (Care Graph), the validation methodology, and the failure modes encountered in production. We state explicitly, in Section 11, what the system does not do and does not claim.

1 Problem statement

Nail condition is among the most frequently photographed and least frequently assessed surfaces in personal care. Clients photograph their hands constantly and have no vocabulary for what they are looking at. Technicians have the vocabulary and see the hand for the first time when the client is already in the chair, which is the worst possible moment to introduce the subject of damage: it arrives sounding like an upsell because it is delivered at the point of sale.

The reason no serious software has filled this gap is a data problem rather than a demand problem. Conventional supervised computer vision requires a labeled corpus. To detect plate separation reliably, a team must first assemble a large body of photographs exhibiting plate separation, have a qualified human annotate each one, and repeat the exercise whenever the operational definition shifts. The space to be covered is unusually wide: natural nails, gel polish, builder gel, acrylic, poly gel, soft gel extensions, chrome and cat-eye finishes, every stage of grow-out and removal damage, across skin tones and under uncontrolled domestic lighting. For a category of this commercial size, assembling that corpus has never been worth anyone's capital.

What filled the vacuum instead was entertainment: applications that accept a photograph, apply a generative effect, and return a compliment. They produce no record, no localized observation and no reason for anyone to change their behavior. They also establish a damaging prior in the market, which is that software claiming to read nails is a toy.

The opportunity is narrower and more useful than the category's marketing suggests: produce an honest, specific, localized and reproducible reading before the appointment, so that the conversation in the chair is about work rather than persuasion.

2 Design principles

Honest mirror

The system is not permitted to flatter. If the observation fields describe damage, the report describes damage. This is a product decision with commercial consequences — some operators want a tool that compliments every client — and it is load-bearing for everything else in this document. A system that cannot deliver bad news has no informational value, and a client who learns it always says the same thing stops reading it.

Refuse before fabricate

When the evidence does not support a reading, the correct output is no reading. This is harder than it sounds, because a language-capable system asked to describe a blurred photograph will produce fluent, plausible, entirely invented description unless something structurally prevents it. Section 6 describes what prevents it here.

The determinism boundary

A probabilistic component is appropriate for perception and inappropriate for advice. NailScan draws an explicit line between the two: the perceptual layer fills observation fields, and nothing else in the system consults the perceptual layer again. Advice is computed from the fields. The boundary is architectural rather than a matter of prompt discipline, so it cannot be eroded by a copy change.

Specification as data, not weights

The rubric that governs what counts as ridging, what counts as lifting and what must never be said is prose, held in context at inference time. It is versioned, change-logged and stamped onto every record. It can be revised in an afternoon by the practitioners who disagree with it, without retraining, data collection or a model release. This is the property that makes the product tractable at all, and Section 5 treats it in detail.

Publish the limits

A vendor that lists no limitations is describing a demonstration rather than a deployed system. Section 11 is not a legal appendix; it is part of the specification, and the sales material carries the same list.

3 System architecture

NailScan executes seven stages. Exactly one is probabilistic.

STAGE	COMPONENT	CLASS	FUNCTION
1	Aperture	Deterministic	Image admission control. Resolution, focus, exposure and framing are checked before any inference is purchased.
2	ONYX-5 extraction	Probabilistic	The perceptual layer reads the photograph against the versioned rubric and fills five observation fields.

STAGE	COMPONENT	CLASS	FUNCTION
3	Confidence gate	Deterministic	Compares the self-reported confidence against the operating threshold. Halts the pipeline below it.
4	Verdict Core — flags	Deterministic	Derives boolean flags from the observation fields by rule.
5	Verdict Core — routing	Deterministic	Resolves whether the report recommends professional review. Source in Appendix A.
6	Care Graph	Deterministic	Classifies wear type and binds a versioned, practitioner-authored care sequence.
7	Ledger	Deterministic	Persists the complete record, then attempts delivery and integration writes.

The ordering of stage 7 is deliberate and is the most frequently skipped design decision in systems of this shape. Persistence precedes delivery. A failure in an email provider, a messaging gateway or a customer relationship system can delay a notification; it cannot destroy an assessment. Integration writes that fail are queued and re-attempted on a bounded schedule with explicit limits on attempt count and record age.

4 The observation schema

ONYX-5 defines five fields. The perceptual layer cannot add a sixth, cannot omit one and cannot answer in prose. A response failing validation is rejected at the tool boundary and retried within a bounded budget; it is never passed through to rendering.

FIELD	CAPTURES	CONSTRAINT
surface	Plate texture: longitudinal ridging, pitting, lamellar peeling, roughness, residue from prior enhancement wear.	Visual description only. No etiology, no cause attribution.
color	Bed and plate coloration read through the plate: uniformity, staining, yellowing, discoloration, banding, region of onset.	Sets <code>discoloration_flagged</code> when outside the range defined as normal in the rubric.
structure	Thickness, breakage, splitting, thinning, lifting, and separation of the plate from the bed.	Sets <code>onycholysis_visible</code> on visible separation.
cuticle	Eponychium and proximal fold: dryness, build-up, trauma indicators, evidence of prior over-cutting.	Observational. Feeds the Care Graph; never reaches the referral rule.
confidence	The perceptual layer's assessment of whether the image supports a reading at all.	Float in [0,1]. Below the operating threshold the assessment is discarded, not returned.

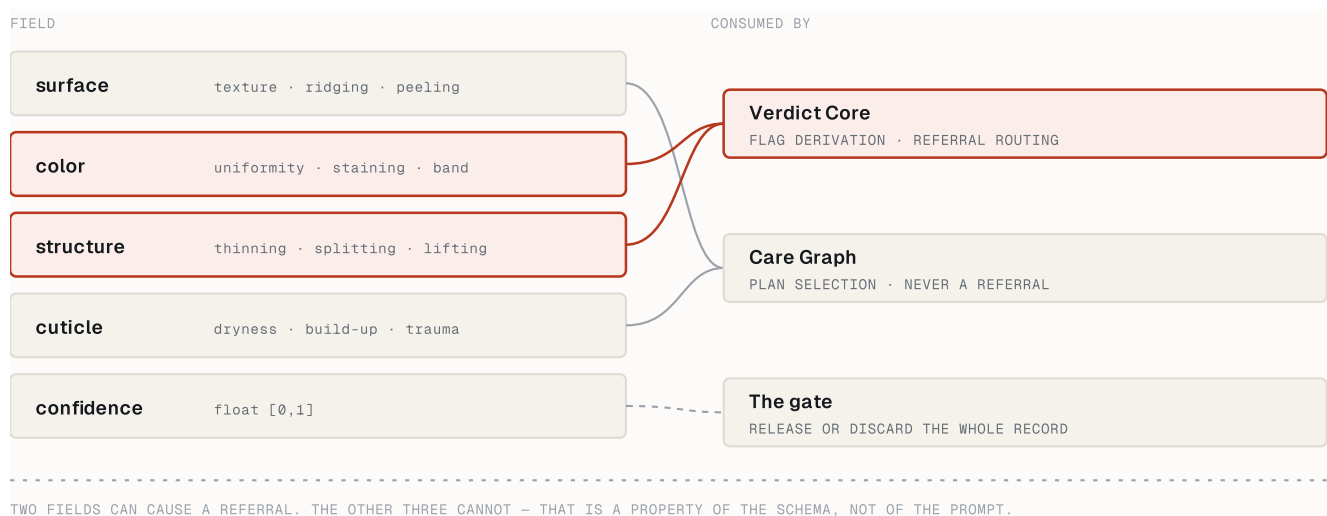


Figure 2 – Field consumption. Which observations are allowed to change what the client is told.

Why five, and why these

The schema is small because every field must be independently useful downstream. A field earns its place if it either drives a rule, selects a care path, or is worth a sentence in a report a technician would not be embarrassed by. Fields that merely add texture to the prose were removed. The practical test applied during design was whether a front-desk operator could read the field on a contact record and open a telephone call with it.

The schema is also the reason the output is a database row rather than a paragraph. Structured fields can be mapped into customer-relationship custom fields, filtered, counted, segmented and used as automation triggers. Free-text assessment output can do none of these things, which is why products that emit prose cannot be wired into an acquisition system without a second extraction step that reintroduces the same reliability problem one layer down.

5 Constrained extraction

The perceptual layer is a multimodal vision system: it accepts an image and a written specification in a single context and reasons over both jointly. Four properties of that arrangement matter operationally.

The rubric is read, not learned

ONYX-5 is prose held in context. It contains field definitions, boundary cases, disambiguation instructions, worked descriptions of edge presentations, and an explicit list of statements the layer must never produce — condition names, diagnoses, reassurance, and any claim about causation. Because the specification is data, revising it is an editorial act rather than an engineering program.

Output conformance is enforced outside the model

Schema conformance is validated at the tool boundary. A malformed or incomplete response is not repaired, reinterpreted or partially accepted; it is rejected and retried under a bounded budget. Where the budget is exhausted, the assessment fails closed and the client is asked to resubmit. There is no fallback path that produces a report from an unvalidated response.

Sampling is configured for minimum variance

Extraction runs at low sampling temperature. The layer is being asked to describe rather than to compose, and the configuration reflects that.

Every record carries its rubric version

Reports are stamped with the rubric version under which they were produced. This is what makes a historical report explicable: when the definition of lifting changed in a later revision, it remains possible to state precisely which reports were produced under which definition, and to re-render any of them exactly as the client saw them.

6 Admission control and confidence gating

Two independent mechanisms prevent the system from answering when it should not.

Aperture

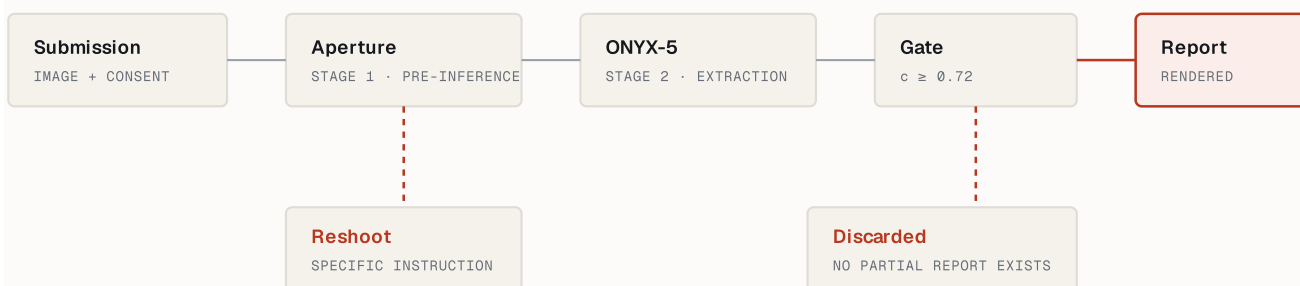
Stage 1 evaluates resolution, focus, exposure and framing before any inference is purchased. Rejection here is cheap, immediate and specific: the client is told what to change. A significant fraction of unusable submissions never reach the perceptual layer at all, which improves both unit economics and the quality of the confidence distribution downstream.

The confidence gate

Stage 3 compares the confidence field against an operating threshold, currently 0.72 in the reference deployment. Below it, the pipeline halts. No partial report is rendered, no fields are shown, and no record of an assessment is delivered to the client — the observation fields the layer produced are retained internally for analysis and discarded from the client path.

The threshold is a product decision rather than a technical constant. Raising it increases report integrity and reduces completion rate; lowering it does the reverse. Operators are shown the trade-off explicitly and the value is configurable per deployment. What is not configurable is the existence of the gate.

SUBMISSION PATH — TWO INDEPENDENT WAYS TO STOP



A BLURRED PHOTOGRAPH CAN STILL PRODUCE FOUR FLUENT FIELDS. THE GATE IS OUTSIDE THE MODEL BECAUSE THE MODEL CANNOT BE THE THING THAT DECLINES.

Figure 3 — Admission control and confidence gating. The two places the pipeline is permitted to answer nothing.

On the value of refusal. The perceptual layer, handed a blurred photograph, is entirely capable of producing four fields of fluent and plausible description. Nothing in the language of that output would signal to a reader that it was unsupported. The gate exists because the failure mode of this class of system is not visible error but confident invention, and the only reliable control is a mechanism outside the model that declines to release the result.

7 Deterministic derivation

Stages 4 through 6 consume the observation fields and produce everything the client is actually told. No stage in this group makes an inference call.

Flag derivation

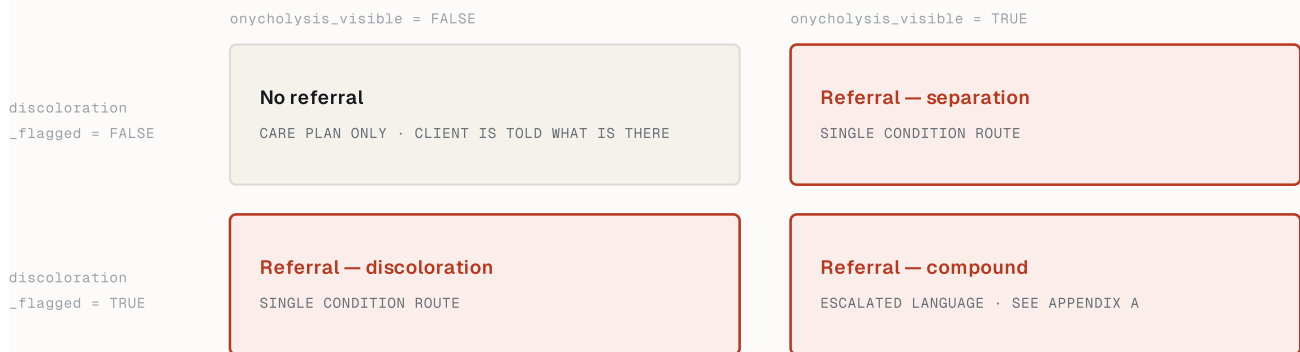
Boolean flags are computed by rule from the observation fields: `discoloration_flagged` from color outside the defined normal range, `onycholysis_visible` from visible plate separation, and `lifting` and `color_change` as the constituents of the compound case.

Referral routing

A single function, reproduced verbatim in Appendix A, resolves whether the report recommends professional review. It returns on the first matching condition, so a single recommendation is issued rather than an enumeration of reasons — enumerating every finding to a lay reader is alarming rather than informative. The complete flag set is written to the record regardless, for the operator's staff to read.

The perceptual layer neither participates in this decision nor is able to suppress it. This matters commercially as well as clinically: the operator cannot ask for the referral logic to be softened for their salon, because it is not a tone setting.

VERDICT CORE — REFERRAL.JS v1.8 · COMPLETE TRUTH TABLE



FOUR INPUTS, FOUR OUTPUTS, NO INFERENCE CALL. THE FUNCTION IS PRINTED IN FULL IN APPENDIX A AND IS THE SAME SOURCE THAT RUNS.

Figure 4 — Referral derivation. The entire decision surface, printed rather than described.

Care Graph selection

Wear type — natural, gel, extension, or damaged-recovering — is resolved from the observation fields, and selects a pre-authored care sequence from a versioned library. The sequences are written and reviewed by practitioners, not generated at request time. Identical inputs therefore return an identical plan, and a plan can be revised centrally without touching any other part of the system.

8 Reproducibility

This section states the property precisely, because the category's marketing has made the word unreliable.

What is guaranteed

For a fixed rubric version and care library version, the derived outputs — flags, referral decision, wear type, care plan binding and rendered report structure — are a pure function of the five observation fields. This is guaranteed by construction rather than by testing: stages 4 through 6 are ordinary code with no non-deterministic input. Two records with identical observation fields cannot produce different advice.

What is not guaranteed, and is instead engineered

The observation fields themselves are produced by a probabilistic system. We do not claim bit-identical extraction across arbitrary conditions or across backbone versions, and any vendor who does should be asked how. What we do is bound the variance: low sampling temperature, a fixed and versioned rubric, schema-enforced output shape, and a confidence gate that removes the low-quality inputs where variance concentrates. In the reference deployment, repeated extraction on a fixed image is stable at the level of the fields that drive rules.

Why the distinction is the whole argument

An unconstrained generative tool re-derives its advice on every submission, so two runs on one photograph can disagree about whether the client should see a professional. Under this architecture that specific failure is impossible: advice is not re-derived, it is computed. The remaining variance is confined to descriptive language in fields that are themselves recorded, versioned and auditable.

9 Validation methodology

Model output never compared against ground truth is a guess with good formatting. Validation in the reference deployment is performed in person.

In-chair comparison

Clients who scan are offered an in-studio assessment at no charge. A technician walks the report nail by nail under USB microscope magnification and records, per field, whether the scan agreed with what was visible under the lens.

Disagreement classification

Disagreements are classified into three causes, and only one of them is a model problem.

1. **Specification defect.** The rubric was ambiguous about the case. This is the most common cause and is remedied by editing the rubric, not by changing anything else.
2. **Image limitation.** The photograph could not have supported the reading. Where this recurs, the remedy is Aperture or the threshold, not the perceptual layer.
3. **Extraction error.** The specification was clear, the image was adequate, and the field was wrong. These are tracked separately and are the only category that speaks to backbone selection.

Longitudinal tracking

Because every record is retained without expiry, an assessment can be compared against the same client's condition at a later interval with consent. Retention photography converts the scanner from a marketing instrument into a record of whether the care plan worked.

On agreement figures. Per-field agreement rates are deployment-specific, sensitive to operator technique and to the rubric version in force, and are not published here as marketing. They are made available under NDA and are walked through on technical evaluation calls. We regard a headline accuracy percentage divorced from its rubric version and its image admission policy as approximately meaningless, and we decline to produce one.

10 Deployment observations

NailScan was built for a working nail studio and operated daily before it was offered to anyone else. The figures below are from that reference deployment and are not projections.

- 294 scan records processed and retained end to end.
- 287 records reconciled against customer-relationship contacts by identity and a ± 10 minute timestamp window during a historical backfill.
- Approximately 250 captured leads worked by telephone directly from the scan findings.
- An internal in-chair mode operated alongside the public scanner, with lead capture bypassed, analytics suppressed and no contact record created.

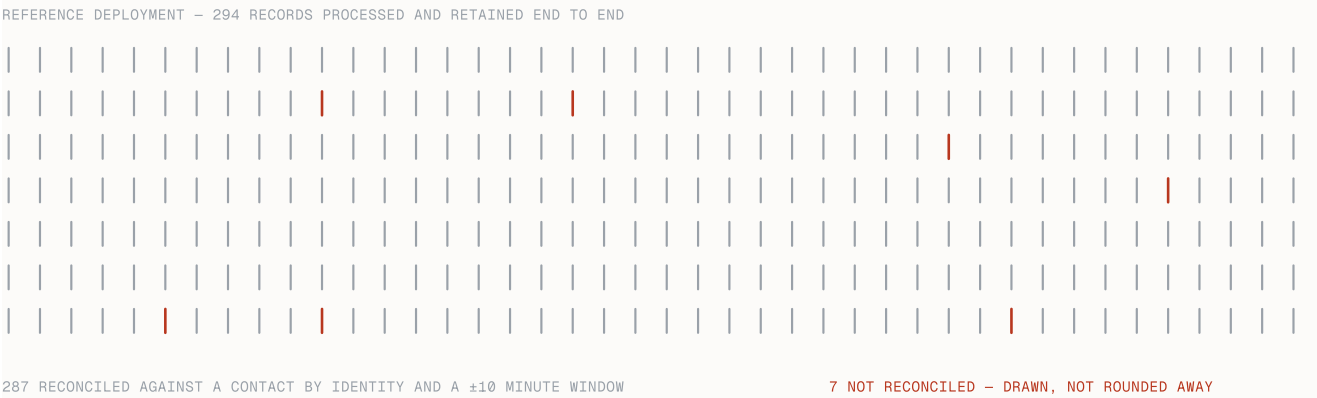


Figure 5 – The reference corpus. Every record the system has produced, including the seven it cannot account for.

Production failures and the controls they produced

Four failures are documented here because each one produced a structural control, and because a system of this shape fails silently by default.

FAILURE	EFFECT	CONTROL NOW IN PLACE
A 180-day storage expiry inherited from a default configuration	Completed reports were being deleted without notice or log entry	Expiry removed; affected records migrated in bulk; retention policy made explicit, documented and covered by test
Integration field mappings deployed unwired	Contacts arrived in the operator's system with the findings fields empty, so staff calling leads had no context	Post-deployment assertions verify every mapping before a deployment is considered complete
An automation left in draft state while paid traffic ran	Captured leads received no follow-up for the duration of a campaign	Workflow state is verified as a pre-condition of any campaign going live
A hardcoded test identity in a code path reachable in production	A phantom contact was created in the live customer-relationship system	Test paths are environment-gated; staging cannot write to production data

The common property is that none of these failed loudly. Each produced a system that continued to serve traffic and return reports while quietly losing value. This is characteristic of pipelines whose terminal writes land in third-party systems, and it is the reason the controls above are assertions and pre-conditions rather than dashboards.

11 Limitations and non-claims

The following are stated as specification, not as disclaimer.

1. **NailScan does not diagnose.** No condition is named, no infection identified and no pathology asserted. The system reports what is visible and, where the rules require, recommends professional review.

- 2. **NailScan is not a medical device** and is not offered, marketed or certified as one. Deploying operators are supplied with compliant language and are contractually asked to use it.
- 3. **It cannot see what a photograph cannot show.** Subsurface conditions, anything beneath opaque polish or an enhancement, and anything outside the frame are out of scope. The report states this rather than inferring around it.
- 4. **It does not replace a technician.** It is a triage and acquisition layer that delivers a better-informed client into a chair, where a trained human performs the work.
- 5. **It refuses.** Below the confidence threshold the output is a resubmission request. Operators who require a result on every submission are not served by this system.
- 6. **It does not flatter.** Where the fields describe damage, the report describes damage.
- 7. **Extraction is not claimed to be bit-deterministic.** See Section 8.2. The derivation layer is; the perceptual layer is variance-bounded.
- 8. **Agreement rates are not published.** See Section 9. They are deployment-specific and are released under NDA.

12 **Data handling**

Clients are photographing their own hands and supplying contact details. The following are the operating commitments.

DOMAIN	COMMITMENT
Ownership	Scan records, contacts and findings belong to the deploying operator. Full structured export is available on request at any time, including after termination.
Images	Photographs are processed to produce that client's assessment. They are not sold, not licensed onward and not contributed to third-party model training. Image retention is configurable per deployment, including a mode that discards the photograph once structured fields are extracted.
Consent	Marketing consent is captured explicitly at the point of capture and stored by reference on the scan record, so the question of what was agreed and when is answerable from the record.
Access control	Operator tooling sits behind an identity layer with an allow list and one-time codes.
Deletion	Client deletion requests remove the record and the contact linkage. Deletion is never a side effect of a storage policy.
Portability	The scanner runs on the operator's domain and the integration writes into a system the operator controls. Discontinuing the relationship does not remove the asset.

13 Conclusion

The useful claim about applied vision systems in consumer categories is not that a model can perceive something. It is that a system can be built around a model such that the model's known failure modes cannot reach the user. NailScan is an instance of that argument in a category where the alternative products are entertainment: perception is confined to one bounded stage, its output is shaped by a versioned specification rather than by learned labels, its release is gated on self-assessed confidence, and every consequential statement made to a client is computed by code that a buyer can read and argue with.

The commercial consequence is incidental to the design and the reason operators buy it: a system that can be trusted to describe a client's nails honestly is a system a stranger will exchange their contact details to use.

Appendix A — Verdict Core referral function

```
// verdict-core/referral.js · v1.8
// Executes after ONYX-5 extraction and the confidence gate.
// Reads observation fields only. No inference call.

function deriveReferral(obs) {
  if (obs.discoloration_flagged)      return REFER;          // R1
  if (obs.onycholysis_visible)       return REFER;          // R2
  if (obs.lifting && obs.color_change) return REFER;          // R3 compound
  return NOT_INDICATED;
}
```

R3 encodes a practitioner judgement that neither lifting nor a color change is independently sufficient to warrant referral language, while their co-occurrence is. The function returns on first match; the complete flag set is written to the record irrespective of which rule fired.

Appendix B — Component glossary

COMPONENT	ROLE
NSE	NailScan Assessment Engine. The complete seven-stage pipeline.
Aperture	Image admission control. Pre-inference validation of resolution, focus, exposure and framing.
ONYX-5	The five-field observation schema and the versioned prose rubric that governs how it is filled.
Verdict Core	Deterministic derivation layer. Flag computation and referral routing.
Care Graph	Versioned, practitioner-authored care sequence library, bound by wear type.

COMPONENT	ROLE
Ledger	Durable record store. Write-before-deliver, no expiry, bounded retry on integration writes.