PROPOSAL FOR DISCUSSION

Private AI Verification Network

A proposed service for organizations that need an independent check on an AI system while keeping sensitive records under their control.

What the customer buys

A buyer commissions a specific test and receives a signed evidence package that an authorized reviewer can check. The package identifies the model and data versions, describes the work performed, and states what the results support.

CUSTOMER QUESTION	DELIVERABLE
Which version was tested?	A provenance record identifying the approved inputs, model and processing steps.
Did the agreed calculation run?	An execution receipt naming the verification method and the steps it covers.
How did the model compare with the baseline?	A comparison with a baseline, including errors, uncertainty and limits on use.

Auriq's role

A provider could accept an ordinary AURIQ payment for the evaluation once a suitable network release is available. The model, source records and findings would stay in the buyer's approved environment or private evidence store. An independent reviewer would check the work separately from the payment.

> ### Decision for collaborators
>
> Select one buyer decision worth testing. Build a signed report on synthetic data, then ask a reviewer to reproduce the result. The proposal claims no paying customers, production deployment, revenue or completed application audit.

01 / CUSTOMER NEED

Who would buy an independent check?

The first commercial question is whether a buyer will pay for evidence that helps them make a specific decision.

POTENTIAL BUYER	DECISION THE EVIDENCE COULD SUPPORT
An organization buying AI	Decide whether a candidate model improves on the current process enough to warrant a trial.
A developer buying computation	Check whether a provider ran the agreed version and completed the portion of the job covered by the receipt.
A model owner or platform	Document which model was released, its permitted use and the execution record the customer may inspect.
An assurance or research team	Give an authorized second reviewer enough evidence to reproduce a result and explain differences.

What the service would have to improve

A buyer can already commission an analyst to supply a signed report. Use that workflow as the first comparison. Give an independent reviewer the same question both ways and measure the time, cost and access needed to check the answer.

Existing tools address parts of the job. iExec documents confidential execution, EZKL supports proofs for selected computations, Gensyn describes reproducible training evidence, and Fiddler offers model evaluation and monitoring. Their respective roles give the pilot team concrete alternatives to compare. [4–7]

In a 2023 U.S. hospital survey, 51% of rural hospitals reported using predictive AI. Among rural adopters, 59% reported evaluating accuracy and 39% reported evaluating bias. These figures describe evaluation practices. Willingness to purchase an independent review is still an open question. [11]

> **A useful first interview**
>
> Ask for the last AI purchase or renewal that required an internal review. Which claim was hardest to check? Who approved the evidence, how long did it take, and whose budget would pay for a better review?

02 / EVIDENCE DESIGN

What the receipt can establish

Agree on the claim before running the job. The buyer should be able to trace each conclusion to a method, a result and a stated limit.

PRODUCT	WHAT A REVIEWER CAN CHECK	LIMIT OF THE CLAIM
Provenance	A signed record identifies the artifacts supplied and their versions.	Source accuracy, ownership and completeness require separate evidence.
Execution assurance	A replay, hardware attestation or cryptographic proof covers specified steps.	Each method has assumptions. A proof of metric calculation covers only those calculations.
Performance evaluation	A pinned model and baseline were compared on a specified dataset under agreed rules.	Results apply to that test. Other populations, settings or time periods may produce different results.

From specification to delivery

1. Specify. Buyer and reviewer agree on the decision, model, baseline, metrics, price, acceptance rules and dispute process.

2. Authorize. The data owner approves who may run the test and which outputs may leave the environment.

3. Execute. Record the data and model versions, preprocessing, exclusions, labels, thresholds, software and random settings. Check rounding where a proof approximates the reference calculation.

4. Review. Release only approved outputs. An independent reviewer checks the signature, follows the replay instructions and documents any discrepancy.

5. Settle. The provider earns the agreed fee for delivering a valid evaluation, including a finding that the model performed poorly.

> **Example of precise wording**
>
> “The reviewer reproduced the reported error rate from the approved predictions and labels.” The receipt would state separately whether anyone checked how the model generated those predictions.

03 / APPLICATION ARCHITECTURE

Where the data and payment belong

Auriq's v0.4 litepaper describes a transparent ledger. Confidentiality for the evaluation therefore comes from how the application handles the data. [1]

LOCATION	WHAT BELONGS THERE
Data owner's environment	Source records, labels, model files and detailed predictions. Run the approved evaluation here with restricted access. Define retention and deletion before the run.
Private evidence store	A signed manifest, permitted results, test instructions, method limits and reviewer findings. Give authorized reviewers controlled access and test that access can be revoked.
Private billing records	The invoice, payment status, transaction reference and link to the evidence receipt. Restrict access to this mapping because it can connect a buyer to a particular job.
Public Auriq ledger	Public addresses, amounts and timing can link a payment to a buyer. Keep job identifiers, report contents, patient identifiers and data fingerprints off-chain.

How a reviewer checks the package

The reviewer checks who signed the manifest, compares the approved files with their recorded digests, and replays the agreed test where authorized. If the reviewer receives only aggregate outputs, the receipt records that limit. Check for disclosures from small groups or repeated queries before releasing results.

Which trust assumptions apply

A signature identifies the signing key. Hardware attestation relies on the hardware and software that vouch for a run. A proof covers the calculation encoded in it. Review the application's storage encryption, key recovery and proof tools separately. Auriq specifies ML-DSA-65 for transaction signatures. [1, 4, 5, 9]

> **Implementation boundary**
>
> The evidence prototype can run entirely off-chain. Auriq checkout requires only a wallet transfer and a private invoice record. The design assumes no smart-contract engine, on-chain proof verifier, escrow feature or consensus change. A payment alone cannot authenticate a report.

04 / PAYMENT AND ECONOMICS

Taking AURIQ for a completed job

Start with a direct invoice for an agreed deliverable. The checkout can be tested independently of the model evaluation. [1, 2]

STEP	WHAT THE BUYER AND PROVIDER SHOULD SEE
Quote	Show the service price, AURIQ amount due, expiration time and refund terms. Preview the wallet's network fee before the buyer approves payment.
Detect and confirm	Record payment seen and payment confirmed as separate states. Set a confirmation policy for the invoice value. A delayed block or a chain reorganization should return the invoice to review when needed.
Resolve exceptions	Test late, partial, excess and duplicate payments. For an approved refund, verify the destination privately, record the outgoing transfer and guard against paying the refund twice.
Reconcile	Export the invoice, amount received, network fees, refunds and provider payout. Replaying the same wallet event must leave the books balanced.

What pays for the service

Potential revenue comes from evaluation, independent verification, integration and recurring monitoring. For each job, record compute, proof generation, reviewer time, support, security, storage and payment costs. Compare margin and total buyer cost with a conventional signed-report engagement.

Auriq's base fee per virtual byte and optional tip are separate from the provider's charge. The service would earn customer fees. The design needs no new coin, new issuance, mining reward or claim on Auriq's Strategic Reserve. Its economics should work without price appreciation. [1]

> **If buyers ask for escrow or bonds**
>
> First document which failure they need covered. A missed delivery or invalid proof can have an agreed remedy. A disagreement about study design needs a qualified reviewer. Any future escrow, bond or challenge reward would need defined custody, funding and dispute rules.

05 / PILOT AND ADOPTION

A first pilot for appointment scheduling

A clinic could use the service to decide whether an appointment no-show model merits a limited trial. The same workflow could later be tested in another sector.

What the clinic would test

Start with synthetic records. With authorization, compare the model, the current process and a simple baseline on historical appointments. Hold out a later time period and use only information available before each appointment. Agree on the decision threshold before inspecting the results.

At the same outreach capacity, report how many missed visits each approach flags and how many appointments it flags unnecessarily. Check whether predicted risks match observed rates, estimate uncertainty and examine permitted subgroups with enough observations. A later shadow trial could record predictions without changing patient care. [8, 10]

Proposed measures of progress

QUESTION	EVIDENCE TO COLLECT
Can the work be delivered?	An authorized independent reviewer reproduces the agreed result. The buyer approves the disclosure controls and accepts the package.
Will customers pay?	Three unrelated organizations purchase ten completed evaluations in total from their own budgets. Exclude grants, rebates and token incentives from this count.
Will buyers return?	Record at least two repeat purchases or renewals. Check retention separately 90 days after pilots begin.
Does AURIQ solve a payment need?	Testnet can demonstrate checkout. After a suitable mainnet release, look for an independent customer paying a second real invoice in AURIQ. Track that separately from general service demand.
Can the service work elsewhere?	Report costs, failures, reviewer hours and gross margin. Seek paid work in a second use case before drawing wider conclusions about demand for the service.

These are proposed targets. The concept reports no completed pilots or verified customer purchases.

06 / DELIVERY PLAN

The first 90 days

Timing depends on buyer access, data permission and funding. Work on the evaluation can begin before an Auriq testnet is available.

WINDOW	DELIVERABLE	ACCEPTANCE CHECK
Days 1–30	Interview buyers, select a decision and baseline, and document costs, access rules and privacy risks.	A buyer identifies the budget owner and review process. A data owner identifies a feasible authorization path.
Days 31–60	Build a synthetic-data evaluation, a signed receipt and a replay guide. Prototype invoice reconciliation.	An outside reviewer reproduces the example. The team checks output disclosure, wallet recovery and duplicate-event handling.
Days 61–90	Prepare an authorized paid pilot and estimate the cost to deliver it. Rehearse testnet payment and refund when available.	Record evidence, remaining gaps and costs. Decide whether to proceed, revise the offer or stop.

Who owns each decision

Product and partnerships. Select the buyer, recruit independent pilots, test pricing and disclose funding, licensing and provider selection rules.

Statistics and domain review. Set the baseline, exclusions, uncertainty method and acceptance criteria. Require human approval for AI-assisted changes to the analysis.

Verification, privacy and security. Define evidence coverage, access, retention and recovery. Document reviewer independence and conflicts. Tie any security review and fix checks to a specific software version and scope. [3]

Platform and payments. Maintain the receipt format, approval controls, wallet integration and accounting export. Payment authority stays with the authorized wallet operator.

When to redesign or stop

Reconsider the service if buyers will not pay, access cannot be authorized, the review exposes sensitive data, or a conventional report meets the need at lower cost. Keep AURIQ checkout optional while its value to customers is tested.

07 / SOURCE NOTES

Design basis and further reading

The project documents describe a protocol design and review recommendations. The remaining sources provide evaluation methods, technical examples and market context.

01 **Auriq Litepaper v0.4**
Pp. 2–5 cover consensus and fees. Pp. 6–8 cover signatures, the transparent ledger and the planned sequence from testnet to mainnet.

02 **Auriq Development Priorities**
Priority 6, p. 2, proposes merchant invoicing, fee preview, confirmations, refunds and repeat payment. This planning note refers to v0.3. Protocol details here follow v0.4.

03 **Auriq Halborn Audit Assessment**
Pp. 1 and 5 recommend an independent review tied to a release and retested fixes. The assessment is a procurement recommendation.

04 **iExec / Proof of Contribution**
Technical documentation on confidential execution, attestation and payment rules.

05 **EZKL / Project documentation**
Proofs for supported model computations, including limits introduced by quantization.

06 **Gensyn / Introducing open-1b**
Example of auditable training evidence and a reproducible review workflow.

07 **Fiddler / Analytics**
Commercial reference for model performance analysis and monitoring.

08 **NIST / AI Risk Management Framework**
Guidance on evaluating and managing AI risk in an organization.

09 **NIST / Post-Quantum Cryptography**
Background on the distinct roles of digital signatures and key establishment.

10 **Health AI Partnership / Locally validate prior to integration**
Guidance for local evaluation and stakeholder review before operational use.

11 **The Journal of Rural Health / Rural hospital AI use and evaluation**
Study reporting rural hospitals' use of predictive AI and evaluation practices in a 2023 U.S. hospital survey.