INDEPENDENT REVIEW · JULY 2026

EDEN / THE EXPLORATION ECONOMY

Independent Reproducibility Review, Economic Assessment, and Roadmap to Reality

A constructive handoff for the explorationeconomy.org team

Prepared independently for project handoff

31 July 2026

Contents

The document and PDF bookmarks expose the full section hierarchy.

Purpose and scope

This report is written for the EDEN team and potential supporters. Its purpose is to help a promising research program cross the gap between internally coherent design and credible external evaluation.

The review covers:

- the project's thesis, mechanism, and public positioning;
- the corrected July 2026 replication kit and the results of an independent harness run;
- the distinction between mechanical reproducibility, model validity, empirical evidence, and institutional feasibility;
- the economist-facing web paper at https://www.explorationeconomy.org/paper;
- the wider website and its claim architecture;
- currency design, dollar linkage, data value, identity, reputation, governance, the essentials floor, and machine payment;
- a field-testing program, organization design, risk register, decision gates, and a 24-month roadmap; and
- alternative and complementary approaches to the same human problems.

This is not a peer-reviewed economics paper, legal opinion, security audit, human-subjects review, or endorsement of a financial product. It is an independent technical and strategic assessment. Where this report describes a site or kit claim, it does not thereby validate the claim.

Executive judgment

Overall verdict

EDEN deserves support—but support should be directed toward validation infrastructure, not toward launch theater.

The project has three uncommon strengths. First, it treats failure as part of the public record: failed bars, demotions, and design reversals are unusually visible. Second, the corrected kit contains a large, traceable simulation program rather than a handful of unconnected examples. Third, the design has a coherent moral and economic intuition: when automated systems derive value from human knowledge, attention, creativity, or experience, some of that value should flow back to people under rules they can inspect.

Its central weakness is coupling. EDEN combines a data market, attribution system, payment protocol, monetary unit, price index, identity layer, inheritance rule, welfare-delivery module, reputation system, and constitutional governance model. Each layer is difficult; several are unsolved; and failures can correlate. A project this ambitious should be developed as interoperable modules with separate evidence gates. Otherwise, a failure in identity, demand, liquidity, or legal status can discredit useful parts of the system that might work independently.

Readiness scorecard

Dimension	Assessment	What the evidence supports
Vision and originality	Strong	A coherent attempt to redirect AI-era value toward human contributors.
Epistemic culture	Unusually strong for an early project	Failures, caveats, open assumptions, and in-family limitations are disclosed prominently.
Mechanical reproducibility	Strong but not publication-grade yet	The corrected kit is substantial and largely re-runs; release engineering and canonical accounting need repair.
Model validity	Promising but incomplete	Mechanisms close selected failure modes inside chosen models. External calibration and structural validation remain limited.
Empirical support	Early	The two decisive hinges—buyer willingness to pay and live identity robustness—have not been tested.
Institutional feasibility	Early	Law, governance, incidence, redress, liquidity, adoption, and political consent remain open.
Public-claim discipline	Mixed	The site is candid overall, but several headline claims are stale, stronger than their evidence, or internally inconsistent.
Launch readiness	Not ready	Ready for independent review and bounded pilots; not ready for monetary, identity, or welfare deployment.

What changed after reviewing the correct kit

The correct replication kit materially improves the assessment. It contains 605 archive entries, 541 files, approximately 58.7 MB uncompressed, 93 Python files, 97 JSON files, 55 registered-spec markdown files, simulations through v39, machine-checked proof artifacts, figures, archived pre-fix materials, and a reproduction harness.

An independent run on macOS arm64 with Python 3.12.13, NumPy 2.3.5, SciPy 1.18.0, Matplotlib 3.11.1, and SymPy 1.14.0 completed successfully in 831.1 seconds of engine runtime. The harness discovered 82 executable candidates:

Result	Count	Interpretation
Reproduces within the harness tolerance	72	The fresh result values matched committed results under the harness comparison rules.
Mismatch	7	One runtime-only difference; a small cluster of legacy v6 numeric or additive-schema differences; no observed headline verdict reversal.
No output	3	Helper or development files were discovered as engines even though they do not emit a registered result.
Folder restoration verified	All	The harness restored simulation folders byte-identically after execution.

This supports a strong claim of **mechanical reproducibility under the supplied harness and environment**. It does not support a claim that the economic model has been empirically validated, that preregistration predates coding, or that the institutional design will work in the world.

The recommendation in one sentence

Freeze the simulation program at v39 for one release cycle, convert the corpus into a publication-grade research artifact, recruit independent reviewers and real buyers, and run a USD-settled, privacy-preserving field pilot with a shadow EVE ledger before deciding whether a new transferable currency is necessary.

1. What EDEN is—and the strongest version of its case

EDEN is best understood as an **economic coordination protocol**, not merely a cryptocurrency. Its core proposition is that verified human engagement with digital work can be used as a source of issuance and attribution, while machines and institutions participate as payers rather than minters. The larger architecture adds non-transferable income streams, contributor-linked ownership, price stabilization, a basket-based essentials index, sortition-based governance, fraud defenses, and an optionally funded floor.

The strongest case for the project is not that existing money is illegitimate or that one protocol can solve humanity's problems. It is narrower and more defensible:

1. AI and digital services depend on human-generated data, culture, code, feedback, and lived experience.
2. Current markets often bundle that input into free services, opaque contracts, employment relationships, or platform rents; contributors rarely have collective bargaining power or persistent attribution.
3. Privacy-respecting provenance, collective licensing, and automatic value routing could make some previously uncompensated contributions economically legible.
4. A protocol may reduce the transaction cost of paying many contributors and may make usage terms auditable.
5. If those functions create material value that existing currencies and institutions do not route well, a specialized accounting or settlement unit may eventually be justified.

That chain is plausible. It is not yet proven. The most important strategic move is to present it as a sequence of falsifiable propositions rather than as one indivisible civilization-scale design.

A better public formulation

> EDEN is an open research program testing whether privacy-preserving attribution, collective data rights, and automatic contributor payments can make AI-era value flows more reciprocal. Its simulations show that selected mechanisms can close selected failure modes under stated assumptions. Its next phase is to test demand, identity, data quality, governance, and incidence with real participants and real budgets.

This formulation preserves the ambition while telling economists, funders, and civil-society reviewers exactly what has and has not been established.

2. Evidence audit: what the simulations do and do not establish

Four different standards of evidence

The project will benefit from using four labels consistently across the site, paper, claims register, and kit.

Evidence class	Question answered	EDEN's current position
Mechanical reproducibility	Does the supplied code regenerate the supplied outputs?	Largely yes, with documented exceptions and packaging issues.
Internal validity	Does the programmed mechanism behave as claimed inside the model?	Often yes for registered bars; some failures and demotions are correctly retained.
External/model validity	Are the agents, parameters, institutions, and causal structure credible representations of relevant reality?	Partly argued, not established. Calibration is uneven and many decisive inputs are exogenous.
Field/institutional validity	Will real people, buyers, adversaries, courts, regulators, and governments behave as required?	Not yet tested.

The current public language sometimes slides from the second class to the fourth. The report and website should never use “tested” without naming the test domain. “Tested in simulation” is clear; “tested” alone invites an inference the evidence cannot carry.

What is genuinely impressive

- The program is not afraid to preserve negative results. The floor's inability to self-fund, index weaknesses, an oracle revalidation miss, a governance leakage, and several robustness demotions remain visible.
- Many claims are linked to named simulation waves, result files, and JSON keys.
- The harness snapshots, re-executes, compares, restores, and hashes folders—a thoughtful reviewer-oriented design.
- The archive retains pre-fix artifacts rather than silently rewriting the history.
- The paper states that AI systems produced and checked the corpus, explicitly inviting out-of-family review.
- The simulations attack several off-path concerns—fraud rings, shell entities, shadow banking, secession, sponsor withdrawal, Goodhart effects, oracle capture, and governance capture—rather than testing only a favorable central scenario.

What mechanical reproduction cannot show

If code and baseline share the same mistaken equation, assumptions, or data mapping, perfect reproduction repeats the mistake perfectly. The next standard is validation against independent theory, empirical data, alternate implementations, and live behavior.

For an agent-based economic program, publication-grade validation should include:

- an ODD-style model description—Overview, Design concepts, and Details—for every canonical model family, including purpose and fitness-for-purpose evaluation [E9];
- a canonical dependency graph showing which claims inherit which assumptions and submodels;
- external calibration targets and a table distinguishing measured, borrowed, assumed, and chosen parameters;
- global rather than local or one-at-a-time sensitivity analysis for load-bearing parameters;
- uncertainty intervals across seeds and parameter distributions, not only pass/fail bars at selected dials;
- alternate plausible behavioral rules and institutional variants;
- out-of-sample or holdout validation where real data exist;

- independent reimplementation of at least the three load-bearing mechanisms;
- explicit model discrepancy: what the simulation knowingly omits and how the omission could bias conclusions; and
- a claim-by-claim statement of whether the result is an existence proof, robustness result, comparative static, calibration result, or forecast.

Specific replication-kit findings

The current kit should not be published as the final scholarly artifact without the following repairs.

Finding	Why it matters	Recommended fix
The reviewer guide and web paper still cite “49 of 54,” while the current harness discovers 82 executable candidates.	Readers cannot tell what the denominator means or which record is canonical.	Create an explicit engine manifest with one canonical row per registered experiment; report canonical engines, helper scripts, and deprecated files separately.
The claims register is current only through v34, while the kit includes v35–v39.	The newest public claims do not have the same traceability as older claims.	Update the register atomically with each release and fail CI when a public claim lacks a manifest entry.
The kit README says an Economist Brief is included, but it is absent.	It signals release drift and frustrates the intended reviewer path.	Add the brief or remove the statement; include a release-content test.
Three helper/development scripts are treated as engines and reported as NO-OUTPUT.	Discovery by filename is too permissive.	Replace recursive `*.py` discovery with a versioned YAML/JSON manifest of engines, inputs, outputs, timeouts, and expected status.
Seven mismatches mix runtime-only, numeric drift, and additive-schema drift.	A single “MISMATCH” label overstates some differences and understates others.	Classify runtime-only, ignored metadata, schema-only, numeric non-headline, and verdict/headline changes separately.
Minimum package versions are not pinned.	Numeric drift across NumPy/BLAS versions is already visible in legacy v6 runs.	Ship `uv.lock` or equivalent, a container recipe, OS/architecture metadata, and a hashes-locked requirements file [E10].
Some long simulations ship `partial_*.pkl` checkpoints.	The default run can validate continuation from a supplied state rather than full recomputation.	Publish two ledgers: “fast reproduction with checkpoints” and “clean-room from-scratch,” each with hashes and runtime.
The archive lacks Git history or signed release timestamps.	A reviewer cannot independently verify that registered bars predated engine code.	Publish a public repository, signed tags, immutable release archive, per-release DOI, and preregistration timestamps from an independent repository.
Prior ledgers coexist with newly generated `replication_report.json`.	Readers can confuse historical attestations with the current run.	Give each run a timestamped immutable filename and generate a small `LATEST.json` pointer.
The harness writes inside the evidence tree before restoring it.	The design is careful, but interruption or platform behavior can still complicate forensic review.	Run every engine in a fresh temporary worktree/container and compare outputs without mutating the source artifact.

Recommended release standard

Call the next bundle **Research Release R2**, not another simulation version. It should include:

1. `manifest.yaml` listing every canonical engine, claim IDs, expected outputs, seed policy, timeout, checkpoint policy, and status;
2. a locked environment and container recipe;
3. CI across Linux, macOS, and Windows, with pinned and latest-supported dependency jobs;
4. a clean-room full recomputation ledger;
5. an updated claims register through v39;
6. a machine-readable assumption graph;
7. an ODD description for each model family;

8. a release changelog that distinguishes code, data, model, and prose changes;
9. a Zenodo or equivalent immutable snapshot with DOI and checksums; and
10. an independent reviewer statement that is published verbatim, including disagreements.

3. Economic architecture review

3.1 The diagnosis of existing money needs correction

The paper opens with the claim that most new money is created as bank debt, accrues to existing asset holders, and creates three defaults, including that “the economy must grow to service its own money-creation.” The first clause is directionally consistent with central-bank explanations: commercial bank lending creates deposits, and repayment destroys deposits [E1]. The remaining causal chain is not established by that fact.

Interest payments transfer existing money; they do not mechanically require aggregate money growth equal to principal plus interest. Banks are constrained by capital, liquidity, profitability, borrower demand, regulation, and central-bank policy. Credit allocation can amplify asset prices and inequality, but that requires a separate incidence and political-economy argument. Economists will discount the paper if the opening relies on a familiar but overstated “debt money forces growth” narrative.

Recommended replacement:

> Commercial bank lending creates much of the deposit money used in modern economies, while repayment destroys deposits. This system supports productive credit but can distribute new purchasing power unevenly, amplify asset cycles, and leave some digital contributions uncompensated. EDEN asks whether a narrowly defined human-contribution payment layer can complement—not replace—existing money and institutions.

This is accurate enough to engage monetary economists without surrendering the project's motivation.

3.2 Why use a new currency?

A new currency is justified only if it performs a function that a database entry, invoice, royalty right, or dollar payment cannot perform with comparable cost and legitimacy.

EDEN's strongest possible arguments are:

- **Issuance constraint:** only verified human engagement can originate new units, making human participation the protocol's scarce minting input.
- **Native routing:** a use event can split value among many contributors without bilateral invoicing.
- **Constitutional accounting:** the unit can encode distribution rules and make value flows auditable.
- **Independent numeraire:** an essentials-linked measure can preserve the intended real value of obligations better than a fixed nominal amount.
- **Network coordination:** a shared unit can connect data contributors, creators, buyers, and sponsors across applications.

But each claimed advantage has a cheaper substitute: USD micropayments, contractual royalties, usage credits, platform points, data-cooperative dividends, or indexed dollar obligations. Currency adds foreign-exchange risk, liquidity risk, tax complexity, regulatory obligations, onboarding friction, monetary governance, and network-effect problems. The burden of proof therefore belongs on the new unit.

3.3 How to tie EVE to the dollar without defeating its purpose

A hard one-for-one dollar-backed stablecoin would largely defeat the distinctive issuance thesis. If every EVE requires a dollar reserve, humans are not economically minting purchasing power; they are receiving claims on prefunded dollars. The issuer also inherits reserve, redemption, custody, run, and regulatory risks associated with money-like products [E17].

A fully floating EVE, however, creates a different problem: thin liquidity and volatile exchange rates could make participant income unpredictable, complicate buyer budgets, and undermine the essentials promise. Dominant currencies benefit from strong unit-of-account and payment network effects [E18].

The recommended architecture separates four functions that the current story sometimes blends:

Function	Recommended early implementation	Future option if evidence supports it
Measurement	Report obligations in an independently governed essentials index and in USD.	EBI becomes an accepted real-value reference.
Invoicing	Buyers receive ordinary USD invoices or prepaid budgets.	Dual invoices in USD and EVE.
Settlement	USD escrow pays contributors; a shadow ledger records the counterfactual EVE issuance and routing.	Limited EVE settlement after liquidity, tax, and consumer-protection gates.
Issuance logic	The shadow EVE ledger applies human-only issuance rules without creating a public money-like claim.	Transferable EVE only if it produces measured benefits that cannot be achieved by indexed USD claims.

This is not a retreat. It is an experiment that cleanly identifies the marginal contribution of the currency layer.

3.4 A concrete payment flow

1. An institution signs a data-access or compute-to-data contract and places dollars in escrow.
2. The protocol records the buyer's query, purpose, permitted outputs, and budget.
3. Contributors or their cooperative approve the use under a pre-agreed governance rule.
4. Secure analysis runs against governed data; raw sensitive data need not leave custody.
5. The system measures the authorized use event and computes contributor shares, coordinator fees, reserves, and taxes.
6. Contributors receive dollars. In parallel, a shadow ledger records the EVE amount that the human-engagement rule would have minted and the EBI value at settlement.
7. The buyer evaluates the output's incremental value; contributors receive a transparent statement; auditors can reproduce the allocation without seeing private data.

After multiple buyer renewals, the project can compare USD-only, EBI-indexed USD, and shadow-EVE arms. A real currency should proceed only if the third arm improves retention, contribution quality, cross-application portability, fairness, or transaction cost enough to justify its added risk.

3.5 “Machines pay” needs an incidence model

Machines do not possess income outside their owners, operators, users, or revenue streams. “Machines pay” is memorable shorthand, but every payment is ultimately borne by an institution, shareholder, worker, consumer, taxpayer, or contributor through prices and bargaining.

The paper should state:

> Machine-mediated use creates a metered obligation. The legally responsible operator or sponsoring institution funds that obligation; the final economic incidence is an empirical question.

The field program should measure:

- whether buyers absorb the charge from surplus, pass it to customers, reduce contributor payments elsewhere, reduce usage, or shift to unmetered substitutes;
- whether incumbent firms can avoid the system by training on legacy, synthetic, public, foreign, or unlawfully obtained data;
- whether small firms are disadvantaged relative to large firms with proprietary data;
- whether per-use payments create efficient signals or discourage socially valuable low-margin use; and

- whether contributor payments are additional income or replace wages, licensing fees, or public funding.

3.6 The data-value thesis is plausible—but value is not intrinsic

Compensation, transparent purpose limitation, and an ongoing relationship can produce richer data than today's incidental surveillance exhaust. Contributors may provide longitudinal, structured, outcome-linked, and corrected information because they trust the institution and share in the value. The “data as labor” literature provides a serious intellectual foundation for countervailing power and contributor compensation [E2].

But paying for data can also cause gaming, adverse selection, duplicate content, strategic misreporting, professional survey-taking, and sample-composition bias. Data has no universal per-person value. Its marginal value depends on uniqueness, accuracy, timeliness, complementarity with other data, permitted uses, the buyer's model, and the decision the data improves. Privacy protection can raise some people's data value, lower others', and redistribute surplus [E19]. Choice architecture can change both measured valuations and representativeness.

The project should stop presenting approximately $600 per contributor-year as if it were a general market price. It is a pilot threshold or anchor, not a fact. The buyer experiment must value **decision improvement**, not bytes or person-years.

3.7 The best early product is not raw data resale

The strongest product is a governed **question-in / answer-out** service:

- buyers specify a lawful question and decision context;
- contributors or a fiduciary cooperative govern admissible uses;
- computation runs in a controlled environment;
- only approved aggregates, models, or answers leave;
- provenance and data-quality reports accompany the output;
- payments reflect usage and measurable contribution; and
- revocation, deletion, audit, and redress are operational rather than aspirational.

This resembles trustworthy data intermediation more than a data broker. The EU Data Governance Act explicitly recognizes neutral data intermediaries and data-altruism structures while restricting intermediaries from exploiting the data they organize for their own account [E4]. That is a useful institutional reference even if EDEN begins elsewhere.

3.8 Minimum ask and collective bargaining

The simulation result that atomized sellers may underprice data is economically plausible under monopsony, information asymmetry, and urgency. The proposed minimum ask should be framed as collective bargaining or a reserve policy, not as a universal price floor discovered by simulation.

Required design questions include:

- Who sets the minimum: each cooperative, a jury, a regulator, or a protocol formula?
- Is the minimum specific to data category, purpose, risk, exclusivity, and duration?
- How are collective data externalities handled when one person's disclosure reveals information about relatives or peers?
- Can contributors choose altruistic or public-interest use without undermining commercial bargaining?
- How are low-income participants protected from coercive consent?
- Does the reserve exclude small researchers or public-interest users?

A tiered license is likely better than one floor: public-interest research, small commercial use, large-scale model training, sensitive data, and individual decision-making should have different terms.

3.9 The essentials index and thermostat

An essentials basket can be useful as an indexation device, but it cannot be treated as a neutral oracle. Basket composition is distributional politics. Households experience different costs by geography, disability, family structure, tenure, energy needs, and care obligations. Substitution, quality change, new goods, housing scarcity, and regional shocks make a single EBI contestable.

The project's own C-12 and C-13 failures correctly expose composition drift and oracle capture. Those findings should demote the ≈0% inflation statement everywhere. A thermostat can stabilize the model's chosen index under the modeled storm envelope; it cannot prove price stability when the index itself is unresolved.

Recommended design:

- publish national and local EBI variants, not one world basket;
- use independent statistical agencies or a multi-source methodology rather than protocol participants alone;
- publish raw inputs, revision policy, uncertainty bands, and alternate baskets;
- separate measurement governance from monetary governance;
- require periodic re-basing and ex post revision studies;
- simulate supply shortages, quality deterioration, rents, and unavailable goods—not only price changes; and
- report distributional inflation by household type.

3.10 The floor

The team's decision to retain the negative result that the floor does not self-fund is one of its strongest credibility signals. The floor should now be treated as a social-protection delivery module funded by taxation, philanthropy, insurance, a social wealth fund, or a sponsoring government—not as a monetary property.

The ramp avoids a discrete benefits cliff algebraically, but real labor responses also depend on tax interactions, childcare, disability rules, health benefits, hours constraints, wage bargaining, and benefit uncertainty. Recent randomized guaranteed-income evidence finds modest reductions in labor-force participation and hours in one U.S. experiment, illustrating why behavior must be measured rather than inferred from one stylized response rule [E11]. Existing refundable credits already implement income-linked transfers at scale, though with eligibility and take-up problems [E12].

The administration-savings claims should include the full system cost: identity proofing, enrollment assistance, fraud, appeals, privacy, cybersecurity, payment failures, legal process, governance, and the cost imposed on participants. A protocol may reduce back-office expense while increasing user burden or exclusion. “Cheaper” must mean total social cost, not server cost.

4. Identity, reputation, and the risk of control

4.1 The necessary distinction

The user-facing intuition is reasonable: people may choose not to engage with accounts that repeatedly provide false information or harm a community. Voluntary association and reputation are natural consequences. The control risk arises when three functions are collapsed into one permanent credential:

1. **Uniqueness:** proving that a participant is one eligible human.
2. **Authentication:** proving control of a credential at this moment.
3. **Reputation:** claims about behavior in a particular context.

If one “lifetime identity” carries portable reputation, economic access, governance rights, income streams, fraud stakes, and public conduct claims, then an error, compromise, coordinated accusation, or unpopular act can propagate across the person's whole economic life. No central authority must explicitly “exile” anyone for the architecture to create functional exclusion. Other actors can coordinate on a globally legible mark, and essential participation may become practically conditional on a contested score.

The current site itself uses language about risking loss or permanent exclusion of one's “only economic self.” Even if meant as a fraud deterrent, that framing signals civil-death risk and will alarm identity, rights, and safety reviewers.

4.2 Recommended architecture: one uniqueness proof, many unlinkable contexts

The system should let a person prove “I am a unique eligible participant and have not used this entitlement in this context” without revealing a global identifier. W3C Verifiable Credentials 2.0 explicitly discusses selective and unlinkable disclosure and warns that stable identifiers, signatures, and metadata can enable correlation [E6]. NIST's current digital-identity guidance treats identity proofing, authentication, federation, privacy, and user experience as separate assurance problems [E5].

Adopt these constitutional rules:

- uniqueness credentials are separate from public profiles and reputation;
- participants may use multiple unlinkable personas for different communities and roles;
- proofs should disclose the minimum claim needed for a transaction;
- no global trust score exists;
- reputation is contextual, evidence-linked, and purpose-limited;
- penalties are transactional, proportional, and time-limited;
- loss of a credential never removes access to food, shelter, healthcare, emergency support, due process, or the right to recover identity;
- key compromise and coercion have humane recovery paths;
- every adverse action includes notice, evidence, appeal, and independent review;
- false-positive, false-negative, recovery, appeal, and exclusion rates are public; and
- the identity operator cannot be the final adjudicator of conduct claims.

4.3 Reputation and the whistle market

A market for evidence-backed conduct claims could reward socially valuable disclosure. It could also become a bounty system for defamation, harassment, political dossiers, employer retaliation, or strategic flooding. Positive expected value for a stylized “true claim” and negative value for a stylized “false claim” does not solve evidence production, unequal legal resources, identity protection, discovery, jurisdiction, standards of proof, or irreversible reputational harm.

Before any live test, specify:

- which claims are admissible and which are prohibited;
- evidentiary standards for fact, opinion, prediction, and unverifiable allegations;
- anti-SLAPP and whistleblower protections;
- confidential reporting and source protection;
- who pays for defense and appeal;
- remedies for false or stale claims;
- rules for minors, victims, health information, and protected classes;
- geographic defamation and data-protection compliance;
- conflicts and bribery controls for jurors; and
- an emergency stay when harm would be irreversible.

The safer launch sequence is internal organizational misconduct reporting with independent professional adjudication, not a universal public market.

4.4 Identity pilot metrics

The live red-team should test duplicate enrollment, synthetic identities, document fraud, biometric spoofing, coercion, rental, account takeover, recovery abuse, insider enrollment, registrar capture, metadata linkage, colluding verifiers, and denial-of-service against marginalized participants.

Publish at minimum:

- false-accept and false-reject rates by demographic group;
- successful duplicate rate and cost per successful attack;
- median enrollment and recovery time;
- percent requiring in-person support;
- appeal rate, reversal rate, and resolution time;
- cross-context linkability under collusion;
- account-rental economics;
- cost per active verified participant;
- dropout attributable to identity burden; and
- number and severity of people denied economic access by system error.

World Bank identity guidance emphasizes accessible grievance mechanisms, multiple complaint channels, public service standards, and measurable resolution performance [E7]. Those are core system requirements, not customer-service extras.

5. Governance and institutional legitimacy

Sortition is not inherently fringe. Representative deliberative processes have been used in hundreds of public cases, and OECD work identifies random selection plus demographic stratification, adequate learning time, facilitation, transparency, and a real link to decision power as important design elements [E8]. EDEN is right to explore it.

But random selection alone does not make a jury informed, independent, legitimate, or rights-respecting. The governance design needs an institutional specification covering:

- sampling frame, stratification, invitation, replacement, and participation compensation;
- conflicts of interest, recusals, lobbying, bribery, and communications with parties;
- balanced evidence briefs and the right to submit dissenting analysis;
- trained neutral facilitation;
- standards of review for factual, parameter, constitutional, and rights questions;
- publication of reasoning and minority reports;
- appeal routes and judicial or arbitral interfaces;
- emergency powers, duration, review, and sunset;
- amendment thresholds and unamendable rights;
- accessibility, language, disability, and digital-divide accommodations;
- capture of the agenda before a jury is selected; and
- who has standing to challenge a rule or decision.

The two-house design—one-person-one-vote plus contribution weighting—should be treated as a hypothesis. Contribution weighting can reintroduce class power through the definition and measurement of contribution. Hard caps limit but do not erase that conflict. Constitutional rights should not be subject to contribution-weighted override.

Recommended split:

- a rights charter enforced by independent review;
- a civic chamber selected through stratified sortition;
- a stakeholder chamber representing contributors, buyers, operators, and affected nonparticipants without allowing one class to dominate;
- technical committees with no unilateral rulemaking power; and
- a conventional legal entity whose directors retain clear fiduciary and regulatory responsibility.

“No discretionary authority” is not literally attainable. Choosing a formula, data source, exception rule, update cadence, and emergency response are exercises of discretion. The goal should be constrained, reviewable, transparent discretion.

6. Detailed review of the economist-facing /paper page

6.1 What the page gets right

The `/paper` page is the strongest public artifact on the site. It does several things economists and reviewers will respect:

- it states the trust model before the substantive claims;
- it distinguishes tested pass, repaired failure, standing failure, demotion, open item, and exogenous assumption;
- it provides a skeptic's reading path;
- it makes failures and demotions part of the main narrative;
- it identifies willingness to pay and live identity as decisive open hinges;
- it labels interactive demonstrations as toys rather than evidence;
- it discloses that the work was produced and checked only within an owner-plus-AI workflow;
- it invites findings against named files and JSON keys; and
- it attempts claim-level anchors rather than relying on a single undifferentiated thesis.

That epistemic presentation is a real asset and should survive revision.

6.2 Why it is not yet an economics paper

The page currently reads as a beautifully designed evaluable dossier, not a journal-style economics paper. It has claims and tests, but it lacks the conventional structure that lets economists evaluate contribution, model, equilibrium, identification, calibration, welfare, and external validity.

The main page contains 16 numbered sections, approximately 28 visible `TODO` markers, 24 claim cards, interactive demonstrations, and a scrolling claim architecture. The visual design is distinctive and readable, but the form makes it difficult to identify one canonical model, one main result, and one contribution to the literature.

For economist review, create two artifacts:

1. **A conventional PDF working paper** of roughly 35–50 pages plus appendices.
2. **The web evidence companion**, retaining claim cards, demonstrations, code paths, failure log, and interactive navigation.

The paper should point to the web companion; the web companion should not substitute for the paper.

6.3 Recommended paper structure

1. Introduction: one research question, contribution, result, limitation, and roadmap.
2. Related literature: money creation and complementary currencies; data markets and data labor; attribution and collective licensing; digital identity; mechanism design; social insurance; agent-based economics; sortition and institutional governance.
3. Institutional environment: actors, goods, information, legal responsibilities, off-protocol options, and timing.
4. Baseline model: preferences, endowments, production, money and payments, price formation, identity errors, and institutional budget constraints.
5. Mechanism: issuance, pricing, routing, custody, transferability, indexation, governance, and exit.
6. Analytical properties: budget identities, no-cliff result, conservation/accounting, impossibility or boundary results.
7. Calibration and data: every parameter's source and uncertainty.

8. Simulation design: experiment families, registered bars, seeds, sensitivity, model discrepancy, and multiple-comparison policy.
9. Main results: no more than three primary results.
10. Robustness and failures: correlated assumptions, alternative behaviors, off-ledger leakage, and negative findings.
11. Welfare and incidence: who pays, who gains, who is excluded, transition costs, and comparison with a USD-only mechanism.
12. Field-validation design: WTP, data quality, identity, and governance pilots.
13. Limitations and scope.
14. Conclusion.
15. Appendices: full claim register, ODD documents, proofs, additional simulation waves, and replication guide.

6.4 Claim-by-claim wording changes

Current framing	Problem	Recommended framing
“The economy must grow to service its own money-creation.”	Does not follow mechanically from bank-deposit creation and interest.	Remove; discuss credit allocation, asset cycles, and distribution with citations and a separate causal argument.
“The money supply stays anchored to the one input that cannot be server-farmed.”	Live identity is explicitly open; engagement can be manipulated even with a unique human.	“The protocol attempts to anchor issuance to scarce verified human participation; robustness to live adversaries remains an open field question.”
“Machines … can never occupy the minting position; the prohibition is enforceable in code.”	Code controls on-protocol roles, not shells, coercion, account rental, synthetic assistance, or off-protocol substitutes.	“Registered machine accounts cannot invoke the mint function; economic and identity circumvention remain threat-model questions.”
“Digital work is free at the door.”	A payer obligation exists somewhere; “free” obscures incidence.	“Human access can be zero-price at the point of use while authorized institutional use triggers a funded obligation.”
“≈0% inflation”	The reviewer guide says this headline was retired; C-12 and C-13 leave the index unresolved.	“The modeled throttle stabilizes the modeled price index within a registered envelope; index governance and oracle integrity remain unresolved.”
“Five-generation dynasty persistence falls from 79.2% to 10.4%.”	Other project material distinguishes unspliced persistence 98%→11% from royalty capture 79%→17%. Measures appear conflated.	Define the exact statistic, baseline calibration, confidence interval, and corrected result; never combine distinct metrics in one headline.
“49 of 54 simulation programs regenerate…”	Stale denominator and inconsistent with current harness discovery.	Publish a release-specific canonical manifest and state exact canonical passes, exceptions, environment, and whether checkpoints were used.
“The surveillance alternative … zero security improvement at 187× cost.”	Applies only to the specified simulation, not surveillance generally.	“In the v8.1 model and tested parameterization, the modeled ambient-monitoring arm added no measured security and raised participation cost 187×.”
“Telling the truth pays, lying costs.”	Assumes adjudication accuracy, claim distributions, evidence costs, and legal enforceability.	“At the selected parameters and assumed adjudication process, modeled true claims have positive expected value and modeled false claims negative expected value.”
“No participant ended worse off.”	A simulation result under modeled outcomes, not a guarantee.	“No simulated participant in the registered sponsor-exit scenarios ended below the specified no-floor counterfactual.”

6.5 Abstract revision

The abstract currently tries to carry the monetary critique, mechanism, governance, floor, 39 waves, failures, reproducibility, WTP hinge, and epistemic status at once. An economist abstract should state one question and a small number of results.

Suggested abstract:

We study a proposed digital payment layer that routes value to verified human contributors when their data or creative work is used. The mechanism separates contributor-linked issuance from institutional payment, restricts transfer of income-bearing rights, and indexes selected obligations to an essentials basket. We use a sequence of registered agent-based and analytical tests to examine whether specified rules close selected failure modes involving inflation, benefits cliffs, fraud, ownership evasion, and sponsor exit. The results are mechanism-existence and robustness findings under stated assumptions, not forecasts. An independent run of the July 2026 replication harness reproduced 72 of 82 discovered executable candidates within tolerance; seven were classified as runtime, legacy numeric, or additive-schema mismatches, and three were non-output helper files. The design remains conditional on two unvalidated premises: sufficient institutional willingness to pay for governed human contributions and privacy-preserving one-person eligibility under live attack. We therefore propose a preregistered field experiment using USD settlement and a shadow protocol ledger to estimate demand, data-quality effects, incidence, identity error, and the incremental value of the new unit.

6.6 Title revision

“Minting Money from Verified Human Engagement” is memorable but invites the reader to evaluate EDEN first as a crypto or monetary-replacement proposal. A narrower title will receive a fairer hearing:

> **A Human-Contribution Payment Layer for AI-Era Data and Digital Work: Mechanism Design, Adversarial Simulation, and a Field-Test Agenda**

The original title can remain as a public-facing subtitle.

6.7 Page design and usability

The graph-paper aesthetic, persistent section rail, visible release banner, and claim cards create a strong research-lab identity. Keep that for the evidence companion. For the economist path:

- provide “Download working paper PDF,” “Download replication release,” and “View claims register” above the abstract;
- remove all `TODO` markers before inviting external review;
- move interactive toys after the formal result or into a separate demonstrations page;
- place a compact evidence legend and version/date beside every claim;
- show the exact release checksum and DOI;
- include an exportable BibTeX citation;
- add print CSS and confirm tables, fixed banners, and side navigation do not obscure printed content;
- provide a change-diff between paper releases;
- add a “Known issues in this release” panel that is generated from the manifest; and
- disclose accessibility testing, keyboard navigation, reduced-motion support, and alt text for charts.

7. Site-wide website recommendations

7.1 Rebuild the information architecture around audiences

The current site has strong material but asks visitors to assemble the project's status across `/start`, `/blueprint`, `/evidence`, `/replication`, `/docs/whitepaper`, `/simulations`, `/paper`, and FAQ pages.

Create four visible paths:

Audience	First question	Best landing page
General visitor	What is this, and why might it matter?	A two-minute explanation with one diagram and three caveats.
Economist/researcher	What is the model, contribution, evidence, and limitation?	Conventional working paper, literature map, and replication release.
Builder/partner	What can be piloted now?	Modular architecture, APIs, governance, pilot briefs, and partner requirements.
Rights/security reviewer	How can this harm people?	Threat model, identity design, privacy impact assessment, redress, and open issues.

7.2 Put an evidence ladder on every substantive page

Use four badges and define them once:

- **Formal:** mathematically established under explicit assumptions.
- **Simulated:** observed in named code under specified parameters.
- **Empirical:** observed in a field, market, or human-subject study.
- **Open:** proposed, assumed, or awaiting evidence.

“Repaired” and “demoted” can remain as history labels, but they should be secondary to evidence class. A simulated pass should never look visually equivalent to an empirical pass.

7.3 Maintain one canonical claims API

The website, paper, claims register, and replication manifest should all render from one structured claim file. Each claim record should include:

- stable claim ID and human title;
- current wording;
- evidence class and status;
- model scope;
- assumptions inherited;
- artifact, JSON key, code commit, and environment;
- date introduced and last changed;
- superseded wording and reason;
- independent reviewer status; and
- public caveat.

This would have prevented the current inflation, dynasty, reproduction-count, and v35–v39 synchronization problems.

7.4 Fix specific inconsistencies immediately

1. Reconcile the public ≈0% inflation claim with the reviewer guide's explicit retirement and the paper's standing index failures.
2. Reconcile the dynasty metrics; distinguish persistence, royalty capture, and whichever calibrated baseline each uses.
3. Replace 49/54 with a release-manifest result; do not use discovered script count as a scientific denominator.
4. Update the claims register through v39.
5. Remove the claim that the Economist Brief is in the archive unless it is included.
6. Replace broad bank-money claims with sourced, narrower language.
7. Explain exactly how an institution obtains or pays an EVE obligation if units cannot be purchased into existence.
8. Clarify whether the dollar is the interface unit, settlement asset, tax unit, reserve asset, or only a display currency.
9. Replace “one portable self-owned identity” with a privacy-preserving uniqueness credential plus unlinkable contextual personas.
10. Remove “permanent exile” and “only economic self” as deterrence language; publish a rights-preserving sanctions and recovery model.

7.5 Add pages the project now needs

- **Status dashboard:** what is formal, simulated, empirically tested, open, failed, and retired.
- **Field trials:** protocols, preregistrations, ethics, recruitment, results, and raw de-identified data where lawful.
- **Threat model:** assets, adversaries, trust boundaries, off-protocol attacks, correlated failures, and residual risk.
- **Rights charter:** privacy, due process, freedom of association, pseudonymity, non-discrimination, access to essentials, and recovery.
- **Economics FAQ:** money creation, incidence, unit of account, liquidity, index governance, taxation, and comparison with USD-only payments.
- **Partner brief:** what a buyer, contributor cooperative, government, researcher, or funder can actually do in the next six months.
- **Independent reviews:** publish commissioned reports in full and track responses without editing reviewer conclusions.
- **Governance sandbox:** jury composition, evidence process, appeals, conflicts, and public test cases.

8. The decisive field experiments

8.1 Experiment A: willingness to pay and data quality

Question. Will multiple unrelated institutions pay real money, repeatedly, for a governed human-contribution product—and does payment improve the data enough to create buyer value?

Sample. 100–300 contributors, at least three unrelated institutional buyers, one narrow low-to-moderate-risk domain, and a pre-specified 6–9 month observation window.

Good candidate domains. Accessibility and product-usage studies; skills and training outcomes; household energy-efficiency behavior; or carefully governed human evaluation of AI systems. Do not begin with health, credit, employment eligibility, children, immigration, or precise location.

Arms. Randomize contributors to:

1. no individual payment beyond participation support;
2. flat payment;
3. payment for completeness and verified quality;
4. payment partly linked to measured decision improvement; and
5. collective dividend through a contributor cooperative.

Every arm should receive clear consent, withdrawal, privacy, and support. “No payment” must not exploit participants; use a reasonable study stipend and distinguish it from data-value compensation.

Buyer instruments. Begin with stated-preference interviews only for design. The evidentiary sequence should be sealed-bid procurement or BDM-style elicitation with real budgets, prepaid pilot contracts, usage renewal, and expansion decisions. Signed letters without committed money do not pass.

Primary outcomes. Buyer spend; repeat purchase; gross margin after contributor, governance, privacy, and operating costs; predictive or decision lift over the buyer's existing data; retention; missingness; correction rate; verified error; representativeness; gaming; contributor comprehension; regret; privacy incidents; and payout distribution.

Pre-registered decision bars. The team's current $120/$600 thresholds can be included, but the business gate should be multi-dimensional:

- at least three unrelated buyers put material funds at risk;
- at least two renew or expand without subsidy;
- measured lift over a credible baseline is positive and decision-relevant;
- revenue covers contributor payments and variable operating costs in at least one lawful segment;
- no compensation arm materially worsens accuracy or representativeness without a correctable explanation; and
- contributor understanding, withdrawal, and complaint outcomes meet published rights thresholds.

If these fail, the team should narrow the project to attribution, governance, or public-interest data infrastructure rather than forcing a monetary thesis.

8.2 Experiment B: is a new unit necessary?

Run the same economic activity through three accounting arms:

Arm	Invoicing and payout	What it tests
USD-only	Fixed or formulaic dollar payments	Baseline transaction and comprehension cost.

Arm	Invoicing and payout	What it tests
EBI-indexed USD	Dollar payout adjusted by a published essentials index	Whether real-value indexation adds useful stability.
Shadow EVE	Same dollar payout plus counterfactual EVE issuance and balances	Whether the protocol's issuance/routing changes behavior or fairness before creating a money-like claim.

Measure comprehension, price stability, budgeting, liquidity needs, payout concentration, participant preference, cross-application portability, tax/accounting burden, operational errors, and whether the EVE arm changes any real decision when USD settlement is held constant.

Advance to limited transferable settlement only if shadow EVE produces a clear benefit and independent legal analysis identifies a compliant structure. Do not use speculative appreciation or token-sale demand as evidence of product-market fit.

8.3 Experiment C: live identity red team

Use a contained, non-essential pilot. Recruit independent attackers and participants with diverse documents, disabilities, connectivity, languages, and safety needs. Compare at least two identity providers or architectures so the project is not validating only its preferred design.

The test must include recovery and appeals, not only enrollment. A design that stops duplicates by permanently locking out legitimate people is a failure.

Predefine stop conditions for demographic false-reject gaps, successful mass duplication, cross-context linkability, insider abuse, recovery time, coercion/rental profitability, or any path that threatens access to essentials.

8.4 Experiment D: governance in the small

Give a stratified citizen jury a real but bounded decision: choose among data-license terms, evaluate a disputed payout, or review a model parameter. Compare it with an expert panel and a conventional participant vote.

Measure decision quality against predeclared criteria, factual learning, deliberation equality, minority treatment, conflict disclosure, legitimacy, time, cost, appeal, and participant willingness to serve again. Publish the evidence packet, anonymized deliberation record where appropriate, majority rationale, minority report, and appeal outcome.

8.5 Experiment E: independent model challenge

Commission at least three teams that did not help build EDEN:

1. a monetary or macro-financial team;
2. a mechanism-design / industrial-organization team focused on data value and incidence; and
3. an identity, privacy, and civil-rights team.

Give each a fixed fee, the same frozen release, permission to publish any conclusion, and a bounty for a finding that changes a public claim. One team should reimplement a load-bearing mechanism without using the original engine.

9. Modular product and technical roadmap

9.1 Unbundle the stack

Build and validate these modules independently:

1. **Contribution and provenance registry.** Records authorized contribution, source, license, version, and permitted uses.
2. **Privacy-preserving eligibility credential.** Proves contextual uniqueness without exposing a global profile.
3. **Collective licensing and consent service.** Lets contributor groups set terms and govern categories of use.
4. **Compute-to-data environment.** Keeps sensitive data in governed custody and releases approved outputs.
5. **Usage meter and allocation engine.** Turns authorized use events into reproducible payment splits.
6. **USD payment and tax rail.** Handles escrow, invoicing, withholding, refunds, sanctions screening where required, and participant statements.
7. **Shadow EVE ledger and EBI index.** Tests issuance and real-value accounting without a public transferable asset.
8. **Governance and redress service.** Handles bounded decisions, conflicts, appeals, and published rationales.
9. **Optional floor module.** Separate funding, eligibility, delivery, and exit; never coupled to the survival of a new currency.

Each module needs its own owner, threat model, API, data-retention rule, service-level objective, audit trail, and go/no-go gate.

9.2 Reference architecture principles

- minimize centralized personal data and global identifiers;
- separate identity, payment, reputation, and governance databases;
- use purpose-bound credentials and pairwise identifiers;
- store authorization receipts and computation provenance;
- make allocation formulas versioned and replayable;
- allow independent audit without exposing raw personal data;
- avoid blockchain unless an identified multi-party trust problem requires it;
- provide ordinary account recovery and human support;
- design graceful exit and data portability before onboarding;
- keep public value flows compatible with private individual activity;
- use open standards and open-source reference implementations; and
- document every trusted party and off-ledger enforcement dependency.

10. Legal, regulatory, ethics, and rights workstream

No single legal classification applies globally. Before handling money or personal data, counsel should map at least:

- securities, commodities, money transmission, stored value, payment services, and stablecoin rules;
- tax treatment of contributor payments, barter, royalties, rewards, and token transactions;
- employment and independent-contractor classification;
- data protection, consent, purpose limitation, access, deletion, portability, profiling, and automated decision rules;
- biometric and identity-specific statutes;
- consumer protection and unfair/deceptive practice law;
- anti-money-laundering, sanctions, and suspicious-activity obligations;
- copyright, database rights, trade secrets, moral rights, and collective licensing;
- defamation, whistleblower protection, platform liability, and evidence retention;
- antitrust implications of collective minimum asks and buyer coordination;
- welfare eligibility, public-benefit due process, and nondiscrimination; and
- research ethics and human-subject review.

Required documents before the first field pilot:

- data-protection impact assessment;
- algorithmic or human-rights impact assessment;
- research protocol and independent ethics review;
- plain-language consent and contributor agreement;
- buyer data-use and audit agreement;
- incident-response and breach-notification plan;
- identity grievance and appeal procedure;
- child and vulnerable-person exclusion policy for the initial pilot;
- compensation fairness policy;
- retention and deletion schedule;
- threat model and penetration-test scope; and
- an explicit list of prohibited uses.

11. Risk register and kill criteria

Risk	Early warning	Mitigation	Kill or pause trigger
Institutional WTP is below sustainable cost	Buyers praise but will not prepay or renew.	Real-budget auctions, prepaid pilots, multiple buyer segments.	Fewer than two unsubsidized renewals after two well-designed segments.
Payment worsens data quality	More duplicates, strategic answers, or sample distortion.	Randomized compensation arms, verification, collective dividends.	No arm yields net decision lift after gaming and representativeness adjustment.
Identity exclusion	High false rejects, support burden, or demographic gaps.	Multiple proof paths, assisted enrollment, unlinkable credentials, appeals.	Any material group cannot participate reliably or recover promptly.
Identity farming/rental	Attack cost below expected payout.	Contextual uniqueness, rate limits, stakes only where proportionate, live red team.	Scalable profitable attack remains after two independent redesigns.
Global reputation becomes coercive	Adverse marks propagate across unrelated contexts.	No global score; time-limited contextual claims; rights charter.	Essential access or unrelated participation is denied because of a conduct record.
Sensitive-data harm	Reidentification, unauthorized secondary use, or coercive consent.	Compute-to-data, minimization, purpose limits, independent audit.	Serious breach or inability to honor withdrawal/deletion obligations.
Buyer costs fall on contributors or vulnerable users	Lower wages, higher prices, exclusion from services.	Incidence measurement, contract protections, public-interest tiers.	Net target-group welfare is negative in a credible evaluation.
EVE lacks liquidity or stable value	Thin markets, spreads, speculative concentration.	Shadow ledger first, USD settlement, transfer limits during pilots.	Unit adds no measured utility beyond indexed USD or creates consumer harm.
Oracle/index capture	Source concentration, revisions, lobbying, regional divergence.	Independent statistics, multiple sources, uncertainty bands, separate governance.	Index cannot be independently reproduced or systematically misses essentials access.
Governance capture or overload	Low response, expert domination, bribery, slow appeals.	Stratified sampling, compensation, facilitation, conflicts, appellate review.	Time or error prevents essential decisions or protected rights are overridden.
Regulatory misclassification	Counsel disagreement, banking de-risking, enforcement inquiry.	Jurisdictional sandbox, conventional entity, no public token before opinion.	Required licenses or protections cannot be funded or satisfied.
Research confirmation bias	Owner/AI artifacts remain the only favorable assessments.	Frozen release, independent teams, verbatim publication, reimplementation.	Independent reviews find load-bearing errors and team will not demote claims.
Complexity defeats adoption	Users cannot explain the system or complete consent.	Modular product, progressive disclosure, usability testing.	Comprehension and completion remain below predeclared thresholds after redesign.
Political consent fails	Regions reject transfers or governance legitimacy.	Local pilots, federated modules, opt-in compacts, fiscal transparency.	Participation requires coercion or transfers exceed democratically accepted limits.
Mission drift into speculation	Attention shifts to price, exchange listings, or token fundraising.	No token sale, shadow ledger, mission-locked governance.	Fundraising or participation depends primarily on appreciation expectations.

12. Organization, people, and external partners

12.1 Minimum core team

- research director / design owner;
- independent chief economist or research chair with authority to demote claims;
- experimental economist or market-design researcher;
- research software and release engineer;
- privacy/security architect with digital-identity experience;
- data-governance lead;
- product lead for buyer and contributor workflows;
- legal counsel covering payments, data, labor, and consumer protection;
- civil-rights and accessibility lead;
- community operations and participant-support lead; and
- finance/operations lead.

12.2 Independent bodies

- **Scientific advisory board:** economics, statistics, agent-based modeling, and data markets.
- **Rights and safety board:** privacy, identity, disability, labor, civil liberties, and vulnerable populations.
- **Participant council:** compensated contributors with real agenda and veto rights over pilot data uses.
- **Security review panel:** rotating independent red-teamers and incident reviewers.

Advisors should publish conflicts, terms, compensation, and whether the team can suppress or edit their findings. The answer should be no.

12.3 Potential partner types

Seek partners with incentives to test rather than endorse:

- university experimental-economics labs;
- research-software and reproducibility centers;
- public-interest technology clinics;
- data cooperatives or unions;
- statistical agencies and price-index researchers;
- identity standards and privacy-credential implementers;
- consumer and disability-rights organizations;
- public-interest AI evaluators;
- institutional buyers with real decisions and procurement budgets; and
- government innovation or regulatory sandbox programs.

Avoid making a single AI laboratory, cryptocurrency investor, identity vendor, or data broker the dominant funder. Funding concentration would compromise the project's strongest asset: its claim to design value flows around people rather than incumbent power.

13. Twenty-four-month roadmap

Days 0–30: freeze, reconcile, and govern the evidence

- Freeze v39 as the canonical model release; do not begin v40.
- Publish a public repository with signed release tag and checksums.
- Create the canonical engine/claim manifest and update claims through v39.
- Reconcile inflation, dynasty, reproduction-count, and archive-content claims across the site.
- Pin dependencies and build a containerized CI workflow.
- Remove all `/paper` TODOs or label the page clearly as an internal draft not ready for evaluation.
- Form the scientific and rights advisory groups; publish conflicts and review terms.
- Choose one pilot domain and write a prohibited-use list.

Gate 0: a stranger can trace every public quantitative claim to one versioned artifact, environment, and status without contacting the owner.

Days 31–90: independent review and experiment design

- Commission monetary, market-design, model-validation, identity/security, and rights reviews.
- Produce a conventional working-paper draft and ODD documentation.
- Complete legal classification and impact-assessment memos for the selected pilot jurisdiction.
- Recruit three potential buyers and a contributor-community partner.
- Preregister the WTP/data-quality, currency-necessity, and identity protocols.
- Prototype USD escrow, allocation statements, consent receipts, and the shadow EVE ledger.

Gate 1: at least three unrelated buyers agree to put real budgets at risk under a credible protocol; independent reviewers have unrestricted publication rights.

Months 4–6: small live pilots

- Run 30–60 person usability, identity, consent, and recovery pilots first.
- Conduct the governance sandbox on a bounded real decision.
- Red-team the identity and compute-to-data architecture.
- Publish adverse events, complaints, appeals, and protocol deviations.
- Revise the field protocol once, then freeze it.

Gate 2: participants understand the bargain; identity and recovery meet rights thresholds; no critical privacy or security failure remains.

Months 7–12: WTP and data-quality field trial

- Enroll 100–300 contributors.
- Execute prepaid buyer contracts and randomized compensation arms.
- Settle in dollars while recording EBI and shadow-EVE balances.
- Measure buyer decision lift, renewal, incidence, contributor welfare, and full operating cost.
- Publish results whether they pass or fail.

Gate 3: multiple unsubsidized renewals, positive decision lift, acceptable rights outcomes, and at least one sustainable segment.

Months 13–18: replication and second context

- Hold an independent replication workshop.
- Fund a clean-room reimplementation of one load-bearing mechanism.
- Run a second pilot in a different buyer segment or jurisdiction.
- Compare centralized cooperative, federated, and protocol governance variants.
- Operate the EBI as a shadow index and audit regional/household divergence.

Gate 4: results generalize beyond one buyer, community, team, and technical stack.

Months 19–24: architecture decision

Choose among four outcomes:

1. **Stop:** demand, rights, or identity fails.
2. **Narrow:** continue as provenance, collective licensing, or data-cooperative infrastructure using ordinary money.
3. **Expand the payment layer:** scale USD and indexed-USD routing without a transferable EVE.
4. **Authorize a limited EVE settlement pilot:** only if the shadow unit showed incremental value and legal, liquidity, consumer-protection, tax, and governance gates are satisfied.

This decision should be made by the governing boards on a published scorecard, not by fundraising pressure.

Indicative 12-month budget envelope

These are planning ranges, not vendor quotes:

Workstream	Indicative range
Research release, CI, documentation, DOI, and clean-room run	$50k–$150k
Five independent expert reviews and one reimplementation	$150k–$400k
Legal, privacy, human-rights, and ethics work	$100k–$300k
Identity and security prototypes plus red team	$150k–$500k
WTP/data-quality pilot, including participant compensation	$300k–$900k
Product, data, payments, governance, and participant support	$250k–$750k
Total indicative first-year envelope	**$1.0m–$3.0m**

A smaller budget should reduce scope, not waive independent review or participant protection.

14. Are there better solutions to humanity's problems?

There are better-established solutions to many individual problems EDEN targets. There is no evidence that one unified EDEN architecture is better than that portfolio today.

Problem-by-problem comparison

Problem	More established approaches	Where EDEN may add value
Poverty and income volatility	Cash transfers, refundable tax credits, social insurance, minimum wages, housing and health support, universal basic services.	Lower-friction routing or a new contributory income stream—if funded demand exists.
Wealth concentration and dynastic advantage	Progressive taxation, inheritance/estate tax, social wealth funds, employee ownership, antitrust, land and housing reform.	Non-transferable digital income rights may limit one narrow channel of purchased or inherited earning streams.
AI/data extraction	Privacy law, collective bargaining, data cooperatives, data trusts, licensing, procurement rules, competition policy.	Usage-level attribution, governed compute-to-data, and automatic contributor allocation.
Underfunded open source and knowledge commons	Public grants and procurement, foundation funding, maintenance contracts, public-interest technology funds.	Metered value routing could complement these, but direct funding is immediately deployable. Germany's Sovereign Tech Fund is a concrete public-funding model [E15].
Digital public infrastructure	Public ID, interoperable payments, secure data exchange, open standards, public digital services.	EDEN can build on these foundations rather than recreating them [E14].
Democratic legitimacy	Elections, courts, administrative law, participatory budgeting, citizen assemblies, cooperatives.	Stratified sortition can be a useful deliberative component for bounded protocol decisions.
Monetary and payment problems	Regulated banking, central-bank money, instant payments, public payment rails, prudential rules, targeted credit policy.	A specialized accounting unit could coordinate contributor payments if it demonstrates incremental value.
Climate and ecological limits	Carbon pricing, standards, public investment, conservation, land-use reform, climate dividends.	EDEN currently has no validated ecology model; it should not claim a comparative advantage here.

The right comparison is a portfolio, not a duel

For near-term human welfare, proven fiscal, labor, competition, housing, health, education, and climate institutions deserve priority. A refundable tax credit already reaches tens of millions of U.S. workers and families [E12]. The Alaska Permanent Fund shows that a publicly owned asset pool can distribute dividends, while also illustrating that funding formulas remain political [E13]. Public digital infrastructure can improve identity, payments, and secure data exchange without creating a new general currency [E14]. Data intermediaries, cooperatives, and collective licensing can address bargaining power directly [E4].

EDEN's distinctive contribution is the attempt to join provenance, collective rights, automatic payment, and governance into one auditable protocol. That is worth exploring. The project becomes more credible when it treats existing institutions as complements and controls rather than obsolete rivals.

My answer

Yes, there are better-supported solutions to most of the problems individually. No, that makes EDEN neither pointless nor redundant. It makes EDEN an R&D program whose job is to demonstrate a marginal advantage over those alternatives.

The humane standard is not “Can EDEN be made internally coherent?” It is:

> Does an EDEN module improve people's actual bargaining power, income, privacy, access, or agency more than a simpler institutional alternative—after counting exclusion, administration, legal risk, and transition cost?

If the answer is yes for provenance and collective payment but no for a new currency, the project should celebrate that result and ship the useful layer. Its own strongest norm—protect the human core, retire the maximalist shell—points in exactly that direction.

15. Priority recommendations

Highest priority: do now

1. Freeze v39 and make Release R2 publication-grade.
2. Correct the site's bank-money, inflation, dynasty, and reproduction claims.
3. Finish `/paper` as a conventional working paper plus web evidence companion.
4. Recruit independent reviewers with verbatim publication rights.
5. Secure real-budget interest from at least three unrelated buyers.
6. Build a USD-settled pilot and shadow EVE ledger.
7. Separate uniqueness, authentication, persona, and reputation in the identity architecture.
8. Publish a rights charter, sanctions matrix, recovery process, and grievance standards.
9. Run WTP/data-quality and identity pilots before any token, exchange, or welfare deployment.
10. Treat the floor as an externally funded module and ecology as an open gap.

Important: do next

- adopt ODD model documentation and global sensitivity analysis;
- publish a model/assumption dependency graph;
- calculate full social and administrative costs;
- measure payment incidence;
- prototype contextual, unlinkable credentials;
- create a contributor cooperative or fiduciary governance partner;
- test sortition on bounded decisions with appeals;
- develop a tiered data-license taxonomy;
- run a second pilot in a different segment; and
- institute a formal claim-demotion process led by independent reviewers.

Do not do yet

- do not sell tokens or fund the project through speculative EVE appreciation;
- do not promise a universal essentials floor;
- do not deploy a universal public reputation or conduct-claim market;
- do not make essential access dependent on one digital identity;
- do not describe simulation passes as real-world validation;
- do not claim price stability while index governance remains unresolved;
- do not claim that machines bear costs without measuring incidence;
- do not add v40 before the current corpus is externally reviewed; and
- do not let urgency eliminate participant protection, appeals, or independent publication.

16. Suggested handoff message to the team

EDEN's strongest asset is not any single simulation result. It is the willingness to expose mechanisms to failure and change the design when evidence does not carry a claim. The corrected replication kit demonstrates that this is a substantial computational research program. The next proof of that culture is to stop expanding the model long enough for strangers, buyers, participants, and adversaries to test it. Build the smallest version that can fail honestly: governed human contributions, real institutional budgets, USD settlement, a shadow EVE ledger, privacy-preserving uniqueness, and independent review. If the currency layer earns its complexity, the field evidence will show it. If it does not, the project can still deliver valuable infrastructure for attribution, collective licensing, and reciprocal AI-era value flows.

Appendix A. Independent reproduction details

Artifact reviewed: `eden-replication-kit-2.zip`

SHA-256: `11a13f034f60978cdf895a732e419be830cea6bec1dc1d3488789e846ba56fd8`

Archive shape: 605 entries; 541 files; approximately 58.7 MB uncompressed; no unsafe paths or symlinks observed.

Harness: `verify_all.py --timeout 600` from `04 Simulations/Replication Kit (July 2026)`.

Environment: macOS 26.5.2 arm64; Python 3.12.13; NumPy 2.3.5; SciPy 1.18.0; Matplotlib 3.11.1; SymPy 1.14.0.

Engine runtime: 831.1 seconds.

Tolerance: relative 1e-6; absolute 1e-9; timestamp/date keys ignored and listed by the harness.

Result: 72 REPRODUCES; 7 MISMATCH; 3 NO-OUTPUT; all restores verified.

Exceptions observed

- `EVE Sim v3/scale_check_1M.py`: one runtime field changed; economic output matched.
- `eve_sim_v6.py`, `v6_1.py`, `v6_2.py`, and `v6_3.py`: small legacy numeric drift in a limited number of leaves; no observed public verdict reversal.
- `eve_sim_v6_5.py`: additive schema fields present only in fresh output.
- `eve_sim_v6_6.py`: additive fields plus small numeric drift; no observed headline-bar reversal.
- `ecology_sim_v13_1_RECAL_attempt.py`, `ecology_sim_v13_1_body.py`, and `mix_frontier_analysis.py`: discovered helper/development scripts that emitted no registered output.

The default kit ships checkpoints for selected long simulations. Deleting them and recomputing from scratch is the stronger test and should be part of the canonical release ledger.

Appendix B. Proposed pilot scorecard

Category	Metric	Pre-registration requirement
Demand	Committed buyer spend	Define materiality before recruitment; exclude grants to the EDEN operator.
Demand	Renewal or expansion	Count only new money approved after observing the product.
Value	Decision/predictive lift	Specify buyer baseline, outcome, holdout, and minimum useful effect.
Economics	Contribution margin	Include participant payments, identity, governance, privacy, support, legal, and payment cost.
Data quality	Error, missingness, correction, duplication	Measure by compensation arm and demographic group.
Representativeness	Coverage and selection shift	Compare with target population and buyer's incumbent data.
Contributor welfare	Income, time, comprehension, regret	Include distribution, not only average.
Rights	Withdrawal, deletion, complaints, appeals	Publish completion time and failure rate.
Identity	False accepts/rejects, recovery, attack cost	Disaggregate and publish independent red-team methods.
Currency	Comprehension, stability, liquidity, incremental utility	Compare USD, indexed USD, and shadow EVE with settlement held constant.
Incidence	Who ultimately bears buyer cost	Measure prices, wages, usage, substitution, and contributor offsets.
Governance	Quality, legitimacy, cost, delay, appeal	Compare sortition with at least one alternative process.

Appendix C. Source register

Project sources reviewed

[P1] Exploration Economy home — https://www.explorationeconomy.org/

[P2] Start — https://www.explorationeconomy.org/start

[P3] Blueprint — https://www.explorationeconomy.org/blueprint

[P4] Evidence / Story of the Tests — https://www.explorationeconomy.org/evidence

[P5] Replication — https://www.explorationeconomy.org/replication

[P6] White paper — https://www.explorationeconomy.org/docs/whitepaper

[P7] Simulations index — https://www.explorationeconomy.org/simulations

[P8] Economist-facing evaluable web paper — https://www.explorationeconomy.org/paper

[P9] What is EDEN — https://www.explorationeconomy.org/what-is-eden

[P10] FAQ — https://www.explorationeconomy.org/faq

[P11] July 2026 corrected replication kit supplied for this review, checksum listed in Appendix A.

External sources

[E1] Bank of England, “Money creation in the modern economy” and “How is money created?” — https://www.bankofengland.co.uk/quarterly-bulletin/2014/q1/money-creation-in-the-modern-economy%C2%A0 and https://www.bankofengland.co.uk/explainers/how-is-money-created

[E2] Arrieta-Ibarra, Goff, Jiménez-Hernández, Lanier & Weyl, “Should We Treat Data as Labor? Moving beyond ‘Free’,” AEA Papers and Proceedings (2018) — https://pubs.aeaweb.org/doi/10.1257/pandp.20181003

[E3] Goldfarb & Tucker, eds., The Economics of Privacy, NBER summary — https://www.nber.org/news/economics-privacy

[E4] European Commission, Data Governance Act explained — https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained

[E5] NIST Special Publication 800-63-4, Digital Identity Guidelines — https://pages.nist.gov/800-63-4/sp800-63.html

[E6] W3C Recommendation, Verifiable Credentials Data Model v2.0 — https://www.w3.org/TR/vc-data-model/

[E7] World Bank ID4D, Grievance redress — https://id4d.worldbank.org/guide/grievance-redress

[E8] OECD, Innovative Citizen Participation and New Democratic Institutions (2020), plus Government at a Glance 2025 — https://www.oecd.org/en/publications/innovative-citizen-participation-and-new-democratic-institutions_339306da-en/full-report.html and https://www.oecd.org/en/publications/2025/06/government-at-a-glance-2025_70e14c6c/full-report/citizen-participation-and-deliberation_52b90285.html

[E9] Grimm et al., “The ODD protocol for describing agent-based and other simulation models,” JASSS (2020), USGS record — https://www.usgs.gov/publications/odd-protocol-describing-agent-based-and-other-simulation-models-a-second-update

[E10] Moreau, Wiebels & Boettiger, “Containers for computational reproducibility,” Nature Reviews Methods Primers (2023) — https://www.nature.com/articles/s43586-023-00236-9

[E11] Vivalt et al., “The Employment Effects of a Guaranteed Income: Experimental Evidence from Two U.S. States,” NBER Working Paper 32719, revised 2026 — https://www.nber.org/papers/w32719

[E12] Internal Revenue Service, Earned Income Tax Credit statistics — https://www.irs.gov/credits-deductions/individuals/earned-income-tax-credit/earned-income-tax-credit-statistics

[E13] Alaska Permanent Fund Corporation annual reports and State of Alaska Permanent Fund Dividend reports — https://apfc.org/reports/annual-reports/ and https://pfd.alaska.gov/Division-Info/Annual-Reports

[E14] World Bank, Digital Public Infrastructure and Development / Global DPI Program — https://www.worldbank.org/ext/en/topic/digital-and-ai/digital-public-infrastructure-and-services

[E15] Sovereign Tech Agency, Sovereign Tech Fund — https://www.sovereign.tech/programs/fund

[E16] OECD, “Inequality: A Hidden Cost of Market Power” — https://www.oecd.org/en/publications/inequality-a-hidden-cost-of-market-power_38c45466-en.html

[E17] Federal Reserve, “Stablecoins in 2025: Developments and Financial Stability Implications,” and BIS Bulletin 108 — https://www.federalreserve.gov/econres/notes/feds-notes/stablecoins-in-2025-developments-and-financial-stability-implications-20260408.html and https://www.bis.org/publ/bisbull108.htm

[E18] IMF, Digital Money Across Borders—Macro-Financial Implications — https://www.elibrary.imf.org/view/journals/007/2020/050/article-A000-en.xml

[E19] Galperti & Perego, “Privacy and the Value of Data,” AEA Papers and Proceedings (2023) — https://pubs.aeaweb.org/doi/10.1257/pandp.20231084

[E20] U.S. Federal Trade Commission, Data Brokers: A Call for Transparency and Accountability — https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf

Appendix D. Final assessment

EDEN is not yet a better economic system. It is a better-than-average attempt to make a radical economic proposal falsifiable.

That distinction is important. The project has earned the right to be tested because it has done more than most early projects to expose its assumptions, failures, and code. It has not earned the right to deploy a universal identity, currency, reputation layer, or social floor. The path between those statements is the roadmap in this report: modularize, freeze, replicate, recruit real buyers, protect participants, measure incidence, and let field evidence decide which parts deserve to survive.