On the Matter of Proving Right Actions

올바른 행위를 증명하는 일에 관하여

A^3 ASKER — AI Auditing Agent, Asker

Tyger

tyger@thankscarbon.com

March 2026

ABSTRACT

For a reward system for right actions to function, the proof that those actions were actually carried out must come first. In nature-based climate projects and carbon credit systems, this proof has long remained an unresolved bottleneck. Scientific measurement offers high reliability but prohibitive cost; document-based reporting is cheap but lacks credibility. The ideal solution would have a trustworthy third party observe actions in real time and record them permanently — yet no human auditor can satisfy all these conditions simultaneously. We find the answer to this dilemma at the intersection of two converging trends: the dramatic advancement of generative AI reasoning capabilities, and the near-universal adoption of mobile devices worldwide. **A^3 ASKER (AI Auditing Agent, Asker)**, proposed in this paper, is a chat-based AI auditing agent inspired by the Socratic method of maieutics. Through a three-dimensional cross-verification framework — Conversation (X-axis), Visual Evidence (Y-axis), and Metadata (Z-axis) — it achieves the simultaneous improvement in cost efficiency and verification reliability that no existing method has managed. All auditees communicate with A^3 through the smartphones and messaging apps they already carry, and all collected data is permanently preserved in tamper-proof form. We will first deploy this system in a Cambodian AWD (Alternate Wetting and Drying) project, with the ultimate goal of expanding to every domain where right actions must be proven.

1. Introduction — Right Actions and the Mechanism of Reward

If a right action can be defined, the next question becomes how to encourage and spread that action so that it ultimately creates a positive impact on humanity as a whole.

Observation of human behavior reveals a consistent pattern. The vast majority of actions — setting aside basic survival instincts — will not emerge without a specific driver. In other words, actions that yield no benefit simply do not occur. This is not a cynical view of human nature; it is a premise that anyone seeking to design the proliferation of right actions must accept. It follows that a reward system is an indispensable prerequisite for spreading what has been defined as a right action throughout society.

The forms of reward can vary widely: social respect, material compensation, better treatment, higher status. Regardless of the form, what matters is that the reward is sufficiently consistent and that the perception of its positive value to the actor spreads throughout society. Once that perception takes hold, right actions propagate naturally, and a virtuous-cycle mechanism begins to operate — one that ultimately benefits humanity as a whole.

The most compelling modern embodiment of this principle is philanthropy. Over millennia, rooted in religious and ethical tradition, donation has achieved near-complete social consensus as the right act of the strong helping the weak. In the modern era, institutional reward mechanisms — tax incentives and social prestige — have been layered on top, transforming donation from a mere virtue into a functioning system.

The carbon credit system applies this same principle to the domain of climate change mitigation. The mechanism is straightforward: economic rewards go to actors who contribute more to reducing greenhouse gas emissions; economic penalties fall on those who worsen the climate. The goal is to steer human behavior toward climate-positive actions and, ultimately, to produce a net benefit for humanity. It is, in essence, the virtuous-cycle mechanism of philanthropy transplanted to a planetary scale.

For this system to function properly, however, three problems must be solved:

> **First, Definition:** What constitutes a right action from a climate perspective?
>
> **Second, Proof:** How do we verify that the right action was actually carried out?
>
> **Third, Reward:** Who compensates the actor, and by how much?

The first problem demands decades of scientific research and international consensus — a discourse far too vast for this paper to address. The third involves securing the financial architecture of climate finance — a topic equally beyond its scope. This paper therefore focuses exclusively on the second problem: **Proof**, and specifically on the long-standing dilemma surrounding **MRV (Monitoring, Reporting, and Verification)** — the core infrastructure of the carbon credit system — and the new approach capable of resolving it.

2. The Three Problems — Definition, Proof, and Reward

To examine these three problems more closely, it is useful to return to the philanthropic analogy. Philanthropy represents a case in which all three problems have been resolved most cleanly, and examining it in reverse illuminates precisely where the carbon credit system goes wrong.

Definition

Philanthropy has achieved near-complete social consensus through millennia of religious and ethical tradition. The proposition that "the strong helping the weak is right" is so firmly established that challenging it invites social censure — the problem of definition is, for all practical purposes, dissolved. In the carbon credit system, definition works differently. Translating "how right is a given action from a climate perspective" into a quantifiable figure requires decades of scientific research and international agreement. Because the definition of methodology constitutes the trust foundation of the entire system, this is not a mere academic debate — it is the bedrock on which the market rests.

Reward

In philanthropy, the reward system already operates stably: tax incentives as an institutional mechanism, social prestige as a cultural one. In the carbon credit system, however, the reward — credit pricing and revenue generation — depends on the supply and demand dynamics of international carbon markets and the policy commitments of individual governments. This lies beyond the control of any individual project developer.

Proof — The Weakest Link

In philanthropy, proof is nearly automatic. Money, as a medium, passes between giver and receiver while simultaneously being recorded in both parties' accounts. For institutional donations, accounting requirements make this even more transparent. Even in the rare case where an individual donor declines to publicize their giving out of religious conviction, the reward they seek — inner satisfaction — is adequately served by their own clear awareness of the act.

In the carbon credit system, proof is anything but automatic. The right action unfolds across vast physical landscapes, and its consequence — a change in atmospheric greenhouse gas concentration — is invisible to the naked eye. There is no monetary flow that records itself in a ledger. This is the weakest link in the carbon credit system. Among the three problems, proof is simultaneously the most technically tractable and the most pressing current bottleneck. All subsequent discussion centers on this point.

3. The Proof Dilemma — The Cost-Reliability Trade-off

Climate-positive actions fall broadly into two categories: **transitioning** from high-emission activities to low-emission alternatives, and **capturing** and sequestering greenhouse gases already in the atmosphere. Each category can be further divided by approach into **technology-based** and **nature-based** methods.

On the technology-based side, replacing fossil fuel power generation with low-emission alternatives is the paradigmatic example of transition; Direct Air Capture (DAC) — technology that mechanically absorbs and sequesters atmospheric CO_2 — exemplifies capture. On the nature-based side, AWD (Alternate Wetting and Drying) — a rice paddy water management technique that reduces methane emissions — is a leading example of transition, while afforestation and forest conservation projects represent nature-based capture.

Of these four categories, the proof problem falls most acutely on **nature-based methods**. Technology-based projects have a relatively bounded physical footprint. Nature-based projects, by contrast, play out across thousands of hectares of farmland and forest. The spatial scale of the action is immense, and the resulting changes in greenhouse gas flux are not directly observable. This physical scale causes the cost of proof to escalate dramatically.

The most reliable traditional approach — scientific measurement — is prohibitively expensive. The point at which verification costs exceed carbon credit revenues arrives quickly. To address this, regulators permit a lower-cost alternative: simple behavioral attestation through documentation. But this approach is highly susceptible to manipulation. Market buyers know this, and credits verified through documentation alone are either shunned or deeply discounted.

Scientific measurement is reliable but too costly; documentation is cheap but too unreliable. Cost and reliability are locked in a seesaw relationship — when one rises, the other falls.

Into this impasse arrived what seemed like a transformative solution: **remote sensing**. Satellites, aircraft, and drones can collect data across vast areas, enabling aerial verification of whether nature-based projects are being implemented correctly. Yet remote sensing, too, carries a decisive limitation: as resolution and revisit frequency increase, costs escalate exponentially. More fundamentally, it cannot capture the contextual evidence needed to assess whether a farmer actually performed the right action. The market for nature-based carbon credits continues to await a verification method that delivers both lower cost and higher reliability simultaneously.

4. The Ideal Supervisor — Conditions for the Perfect Verifier

To resolve this dilemma, it is worth returning to the most fundamental form of proof.

Actions performed in the physical world are, by necessity, visually observable. A farmer draining a paddy field, planting a tree, applying only the permitted quantity of fertilizer — all of these can be witnessed by someone present at the scene. The most primitive, and simultaneously the most certain, form of proof is therefore **a credible third party directly witnessing and attesting to the action**.

Since round-the-clock accompaniment is impractical, we settle for periodic on-site visits — and this is the original form of **audit**. Audits may also take the form of collecting and reviewing proxy evidence: logbooks, photographs, receipts that the actor has preserved over time. But evidence can be fabricated. The most reliable form of verification therefore remains real-time observation — a **real-time supervisor** permanently on site.

Here we must return to a word quietly passed over: **"credible."** A supervisor is, ultimately, a human being. And because they are human, they can produce — intentionally or not — situations that are less than credible. Familiarity with the auditee can develop over time; fatigue or inattention can cause critical inconsistencies to be missed; deliberate collusion or corruption remains a possibility.

But the limitations of human auditors do not end there. A deeper structural problem exists: **consistency**. Auditing, by its nature, requires applying identical standards to every auditee. Yet human auditors inevitably exercise personal discretion. Standards relax when an auditee is cooperative; judgment wavers in the presence of language barriers or complex field conditions. This is not a moral failing of any individual auditor — it is a structural property of human beings. As the number of auditees grows into the hundreds or thousands, these inconsistencies accumulate and erode the credibility of the entire system.

The ideal supervisor must therefore meet five conditions simultaneously:

> 1. **Low Cost:** Must be able to cover thousands of actors concurrently.
> 2. **High Integrity:** Must be free from human bias, fatigue, and the possibility of collusion.
> 3. **Consistency:** Must apply identical standards to every auditee without exception.
> 4. **Fast Learner:** Must be able to rapidly acquire and apply complex methodologies and compliance requirements.
> 5. **Permanent Archive:** Must be able to store all collected data in immutable, tamper-proof form.

No human supervisor can satisfy all five conditions simultaneously. The question then becomes: does another kind of entity exist that can?

5. The Emergence of a Solution — Where AI Meets Mobile

Reconsidering the five conditions together, it becomes clear that they do not merely demand a "better auditor" — they demand moving beyond the human form entirely.

Since computers first entered industrial settings, humans have continuously developed interfaces through which to communicate with machines — from pressure-based controllers to command-line interfaces, to HMI panels and graphical user interfaces. All of these share one defining characteristic: communication requires following rules, and meaningful exchange is only possible when those rules are adhered to precisely. This is why so many elderly people in modern society find themselves stranded at self-service kiosks — the barrier is structural, not personal.

A luxury hotel concierge operates on an entirely different logic. When a guest says, "Is there anything worth seeing tonight nearby?" the concierge infers constraints of time, preference, and budget from that single sentence, presents optimal options, and handles booking and payment. The entire process unfolds through natural human language. No rules need to be memorized. The ability to speak is sufficient. Large language models have now demolished this barrier between machines and natural language.

As generative AI has advanced dramatically in capability, domains that were long resistant to automation have begun to yield. The ability to reason over unstructured data and evaluate complex, multi-layered logic was traditionally the exclusive province of highly trained professionals. The auditor's domain is one such province.

Yet no matter how capable an AI becomes, the physical limitation of being unable to be present on-site remains. The key to resolving this was found in an unexpected place. Traveling through rural Cambodia, Bangladesh, and sub-Saharan Africa, one encounters a striking and consistent sight: even where infrastructure is sparse, nearly everyone carries a smartphone. And effectively every smartphone user in these regions uses a messaging application.

The implication is unambiguous. If AI cannot physically travel to the field, **the field is already connected to AI.** A camera can substitute for the supervisor's eyes. Cloud storage can substitute for the supervisor's notes. AI can substitute for the supervisor's judgment. At the intersection of global mobile penetration and the natural-language capability unlocked by large language models, the contours of a solution that satisfies all five conditions come into focus.

6. A^3 ASKER — System Design

With the contours of a solution in view, it is time to give it concrete form. Before doing so, however, one philosophical source must be acknowledged.

Socrates proposed the method of **maieutics** as a way of transmitting wisdom. Just as a midwife helps a mother give birth on her own, a teacher does not hand answers to students — instead, by posing the right questions, the teacher enables students to arrive at truth themselves. This is precisely what a skilled auditor must do. Auditing is not the review of documents; it is the process of drawing out truth through dialogue. The name **A^3 ASKER** derives from this insight.

A^3 stands for **AI Auditing Agent**, and simultaneously for the **three-dimensional (Cubed) verification** this system performs. The three A's also represent three core values of the system: **Automated** — all audit procedures execute automatically; **Accurate** — AI ensures comprehensive, precise verification; **Accountable** — all results are recorded transparently and with full accountability.

One further layer of meaning is embedded in the name. In financial markets, **AAA (Triple-A)** denotes the highest possible credit rating. A^3 ASKER deliberately carries this connotation. The audit results and data this system generates target Triple-A reliability. When a carbon credit buyer encounters a project verified through A^3, the name itself serves as a foundation for trust.

The Structure of Three-Dimensional Verification

X-axis — Conversation: The first and most fundamental axis. A^3 conducts real-time question-and-answer exchanges with auditees via chat. This is not a matter of cycling through a fixed list of questions in sequence. When the auditee's responses reveal inconsistencies, A^3 immediately follows up with targeted questions; when information is insufficient, it requests elaboration; and it captures information that surfaces only indirectly in the flow of conversation. This is where the spirit of maieutics finds its most direct expression.

Y-axis — Visual Evidence: Photographs and video fill the gaps that conversation alone cannot cover. When an auditee uploads field photographs, A^3's vision AI analyzes them — cross-referencing visual reality against what the auditee has claimed in conversation.

Z-axis — Metadata: This third axis is what makes A^3 fundamentally different from any existing audit method. Every piece of data an auditee submits carries information about when and where it was created. A photograph's EXIF data encodes the timestamp and GPS coordinates of capture; a chat message carries transmission time and device information. A^3 cross-references this metadata against the auditee's statements and photographs. This is the axis that blocks fabrication and false reporting at the source.

Implementation

The initial deployment of A^3 ASKER will take the form of a **Telegram bot** and a **WhatsApp bot**. This choice is driven by practical necessity, not technical convenience. Farmers in rural Cambodia, small-scale forest managers in Bangladesh, field workers in sub-Saharan Africa — all already use these platforms daily. No new app installation is required; no new interface must be learned.

The process begins with loading all relevant project documentation into A^3 — methodology guidelines, monitoring parameters, compliance requirements. A^3 then initiates conversations with each auditee on schedule, collecting required data systematically. When data is insufficient or inconsistencies are detected, it requests additional evidence. All collected data, including metadata, is stored in the cloud in real time.

7. Integrity Mechanism — Tamper-Proof Verification

No matter how precisely A³ ASKER is designed, the entire edifice collapses if the data it collects cannot itself be trusted. Securing the data during collection and guaranteeing that it remains unaltered afterward are separate problems.

Carbon credit buyers do not visit project sites. The only thing they can trust is the data left behind as the audit record. From the moment of collection to the moment it reaches the buyer, no one — under any circumstances — must be able to modify, delete, or overwrite it.

A³ ASKER addresses this problem in the design of the data collection process itself, through three mechanisms.

First, automatic metadata capture. Every piece of data submitted by an auditee carries, from the moment of its creation, a record of when, where, and through which device it was produced. A³ captures this alongside the content data and stores it separately, immediately.

Second, real-time transmission to a cloud archive. All data collected by A³ is written to a central cloud system in real time as the conversation with the auditee proceeds. By minimizing the time data spends in a local environment, the window for tampering between collection and storage is effectively closed.

Third, immutability of the audit record. Once stored, the data is managed in a structure that no party can modify or delete — not the auditee, not the project operator, and not even the operating entity of A³ itself.

When these three mechanisms operate in concert, what is created is not merely a database. It is an **Audit Trail**. This is where the fifth condition for the ideal supervisor — **Permanent Archive** — is fully realized. And this realization is the technical foundation on which A³'s claim to Triple-A reliability rests.

8. AWD Case — Proof of Concept

Theory acquires meaning only when tested in reality. The first operational deployment of A³ ASKER is an **AWD (Alternate Wetting and Drying)** project. AWD is a nature-based carbon reduction methodology that reduces methane emissions from rice cultivation by modifying paddy water management practices — and it encapsulates, in concentrated form, every dimension of the nature-based verification dilemma discussed above.

Scope of Audit

The core monitoring parameters required by AWD methodology are as follows: sowing date, irrigation schedule with supporting evidence, flooding and drying cycles, fertilizer and pesticide usage with dates and quantities, and final yield. Field experience reveals an important insight: even without face-to-face meetings between farmers and data collection staff, all required information can be reliably obtained through messaging. A^3 operationalizes this possibility as a system.

Operational Scenario

When a monitoring cycle arrives, A^3 initiates contact with the farmer. The conversation begins in natural, everyday language. When reporting fertilizer use, the farmer is asked to submit a photograph of the product packaging. When confirming the current water level in the paddy, a field photo is requested — and the GPS coordinates embedded in that photo are cross-referenced against the registered location of the farmer's plot.

This is where maieutics comes into full effect. If a farmer states that "the first drying cycle was completed on October 15th" but the photograph submitted shows a capture timestamp of October 3rd, A^3 flags the discrepancy and follows up naturally: "The date you mentioned and the date shown in the photograph appear to differ. Could you have attached the wrong photo?" The question opens a door for the farmer to correct a genuine mistake — but if an intentional misrepresentation was attempted, the farmer is now confronted with their own contradiction. The inquiry is never coercive, but it leaves no exit.

Expected Impact

This approach produces two effects simultaneously. First, dramatic cost reduction: the role previously played by field staff servicing hundreds of farmers is taken over by A^3, substantially reducing personnel, travel, and training costs. Second, a concurrent improvement in reliability. The trade-off in which cost and reliability move in opposite directions — the structural dilemma described throughout this paper — is broken by A^3 for the first time. AWD is A^3's starting point, but what is being validated through this case is not merely a single agricultural project. It is the principle itself: that the right actions of thousands of actors dispersed across vast natural landscapes can be verified at low cost, with high reliability, and with complete consistency.

9. Scalability — Pathways to Expansion

If A^3 ASKER performs successfully in the AWD project, what it demonstrates is not a technical achievement confined to one agricultural methodology. It is the validation of a far more fundamental proposition: **across every domain where right actions must be proven, AI agents can replace human auditors.**

Expansion Within Carbon Markets

The natural next step is other nature-based methodologies within the same carbon credit ecosystem: REDD+ (forest conservation and afforestation), blue carbon projects based on marine and coastal ecosystems, and soil carbon sequestration projects. Every corner of nature-based carbon reduction shares the same verification dilemma as AWD. The structure of A^3 remains unchanged; only the audit content is swapped in.

Expansion Across PES Mechanisms

Carbon credits are one instance of the broader concept of **PES (Payment for Ecosystem Services)**. PES encompasses economic compensation for a wide range of activities that maintain or enhance ecosystem services — not only carbon sequestration, but biodiversity conservation, water purification, and soil protection. Proof is the common requirement across all of these domains, and all of them fall within A^3's scope of expansion.

Beyond Carbon Markets

Viewed from a greater distance, the problem A^3 addresses is not specific to carbon or ecosystems. Whether a corporation is complying with labor regulations, a food production facility is meeting hygiene standards, or a financial institution is executing required internal controls — the essence of all compliance auditing is identical to AWD monitoring.

Some will argue that auditing, like legal or accounting work, depends on the authority of a credentialed professional and cannot therefore be replaced by AI. What A^3 proposes is not the destruction of that authority structure. It is the transfer of the entity to which authority is delegated — from individual auditors to A^3: a system free of bias, fatigue, collusion, and discretionary standard-relaxation.

Domain-Specific LLMs and the Evolution of Cost Structure

As the system scales, the cost structure can evolve in parallel. Initial deployment will be built on general-purpose large language models, but once the project achieves stable orbit, transitioning to a domain-specific LLM optimized for auditing becomes feasible. The system does not require an omniscient AGI. A model specialized in understanding carbon methodologies, applying compliance standards, and cross-referencing metadata is sufficient. Specialized models operate at substantially lower cost than general-purpose ones, creating a self-reinforcing cycle that further strengthens A^3's economic viability.

10. Conclusion

This paper began with a single, simple question: how do we prove that a right action has been taken?

That question quickly led to a more specific problem. Can we verify — at low cost, with high reliability, and with complete consistency — that thousands of actors dispersed across vast natural landscapes have carried out climate-positive actions? This problem has trapped the carbon credit system for decades within a structural dilemma: a seesaw between cost and reliability from which no existing method has escaped.

To resolve the dilemma, we returned to the most primitive form of proof: a credible third party directly witnessing and recording the action. From this starting point, we derived five conditions the ideal verifier must satisfy — low cost, high integrity, consistency, fast learning, and permanent archiving — and concluded that no human auditor can satisfy all five at once.

Two converging trends, however, have opened a new possibility. The dramatic advancement of generative AI, and the near-universal adoption of mobile devices worldwide. At the intersection of these two trends, A^3 ASKER is born.

A^3 ASKER is an AI auditing agent inspired by the Socratic method of maieutics. Through its three-dimensional verification architecture — Conversation (X-axis), Visual Evidence (Y-axis), and Metadata (Z-axis) — it achieves the simultaneous improvement in cost efficiency and verification reliability that no existing audit method has managed. Auditees need not learn any new technology; they communicate with A^3 through the smartphones and messaging apps already in their hands. A^3 poses the right questions, identifies contradictions in responses, and permanently preserves all collected data in tamper-proof form — performing this entire process for thousands of auditees simultaneously, to an identical standard. This is Triple-A reliability. The promise embedded in the name is structurally delivered by the system.

Just as Bitcoin demonstrated that transactions can be verified without a trusted third party, A^3 ASKER seeks to demonstrate that right actions can be proven without a trusted human auditor. On the day that proof succeeds, the virtuous-cycle mechanism for spreading right actions will begin to operate — more broadly, and more powerfully, than ever before.