AI Decision Support for **Nature Credit Project Design**

Nature Tech Collective

CONSERVATION INTERNATIONAL

Earth Genome

With support from the Patrick J. McGovern Foundation

THE PROBLEM

Practitioners developing nature credit projects face a confusing and fast-moving landscape. Early in project design, they need to make decisions that are both technical and consequential, such as: *Which framework or standard should they use? What biodiversity outcomes should they monitor? How do they measure biodiversity in a way that balances scientific rigor with feasibility in the field?*

These questions are difficult to answer because the market is fragmented, siloed, and evolving rapidly. More than 60 frameworks and standards exist for nature credits, each with different requirements, crediting logic, and applicability across ecosystems and geographies. What dimensions of biodiversity matter depend on project goals, local ecology, and the specific credit pathway being pursued. Technologies to support monitoring exist, but are often prohibitively complex to deploy at scale, and practitioners lack clear guidance on how to choose among them or understand the real cost and rigor tradeoffs.

Getting these decisions wrong at the outset is costly — yet no accessible, comprehensive resource exists to guide teams through the process.

OUR APPROACH

We set out to explore whether AI could help practitioners navigate the early design of nature credit projects – as one illustrative example of the broader nature-positive landscape.

The project was intentionally exploratory: the goal was not to build a finished tool, but to test where AI could provide useful support, where it breaks down, and what would be required for more reliable decision-support systems in the future.

To do this, we built a modular, AI-driven prototype structured around major biodiversity-related decisions in early project design. The system uses a supervisor agent to route user queries to a set of specialist tools, each focused on a distinct question:

FRAMEWORK SELECTION

Explains and compares biodiversity and nature credit frameworks, including methodologies, crediting logic, additionality, permanence, leakage, and verification requirements

BIODIVERSITY MONITORING TARGETS

Draws on species distribution data and monitoring protocols to recommend relevant taxa and indicators for different credit pathways

MONITORING DESIGN

Recommends specific monitoring approaches, technologies, deployment strategies, and cost estimates

ACTORS AND ORGANIZATIONS

surfaces relevant platforms, data providers, technology providers, and implementation partners

© Daniel Rosengren

Each tool drew from a curated knowledge library. The prototype was used to test not only whether AI could retrieve relevant information, but whether it could help practitioners reason through decisions in a complex and still-evolving field.

Standard retrieval-augmented generation (RAG) — the most common technique for grounding AI responses in domain documents — proved inadequate here. Biodiversity framework documents are long, technically dense, and highly heterogeneous. During initial testing, we found that the information retrieved was too broad, citations too shallow, and reasoning over unstructured text too prone to error.

In response, the team developed a new approach: **JSON Rule-Augmented Generation (JRAG)**. Rather than retrieving raw text chunks, JRAG converts dense technical documents into structured, human-readable JSON rulesets that encode key requirements, definitions, conditions, and decision logic.

This makes responses more auditable – users and experts can see which rule informed a given answer – and directly editable by domain experts who can review and refine the rulesets over time.

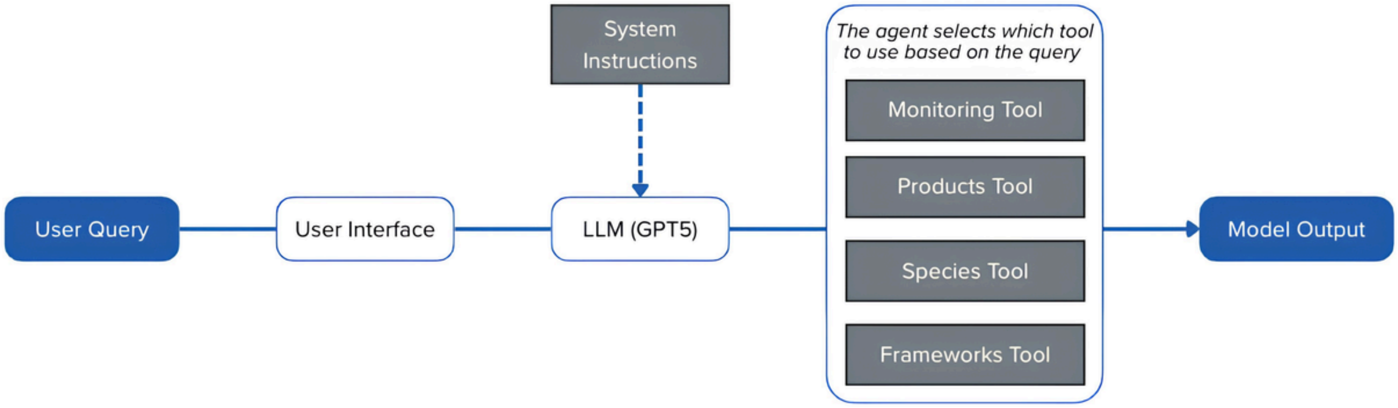

Figure 1. User Journey

We created a modular system with specialized tools for different stages of a user journey (Figure 1). Users input questions into a chat-based interface, powered by GPT5. System instructions direct a supervisor agent to call upon specific tools depending on what the user needs to know or do. We curated key knowledge sources and converted them into human-readable semi-structured JSON rulesets (see Figure 2), that the specific tools draw from to provide relevant answers (Figure 3). A full list of resources and how they were used is available in the technical report; we drew especially heavily from the Nature Tech Collective's Biodiversity Sector Map and the Biodiversity Intelligence Company's Overview of Biodiversity Measurement Approaches.

```
{
  "HIFOR": [
    {
      "name": "min_project_area",
      "description": "Projects must
        encompass a large HIFOR
        Accounting Area (HAA).",
      "detail": "The HAA must be ≥100,000
        ha…",
      "cite": "HIFOR Methodology, 2024"
    }
  ]
}
```

Figure 2. JSON Rulesets

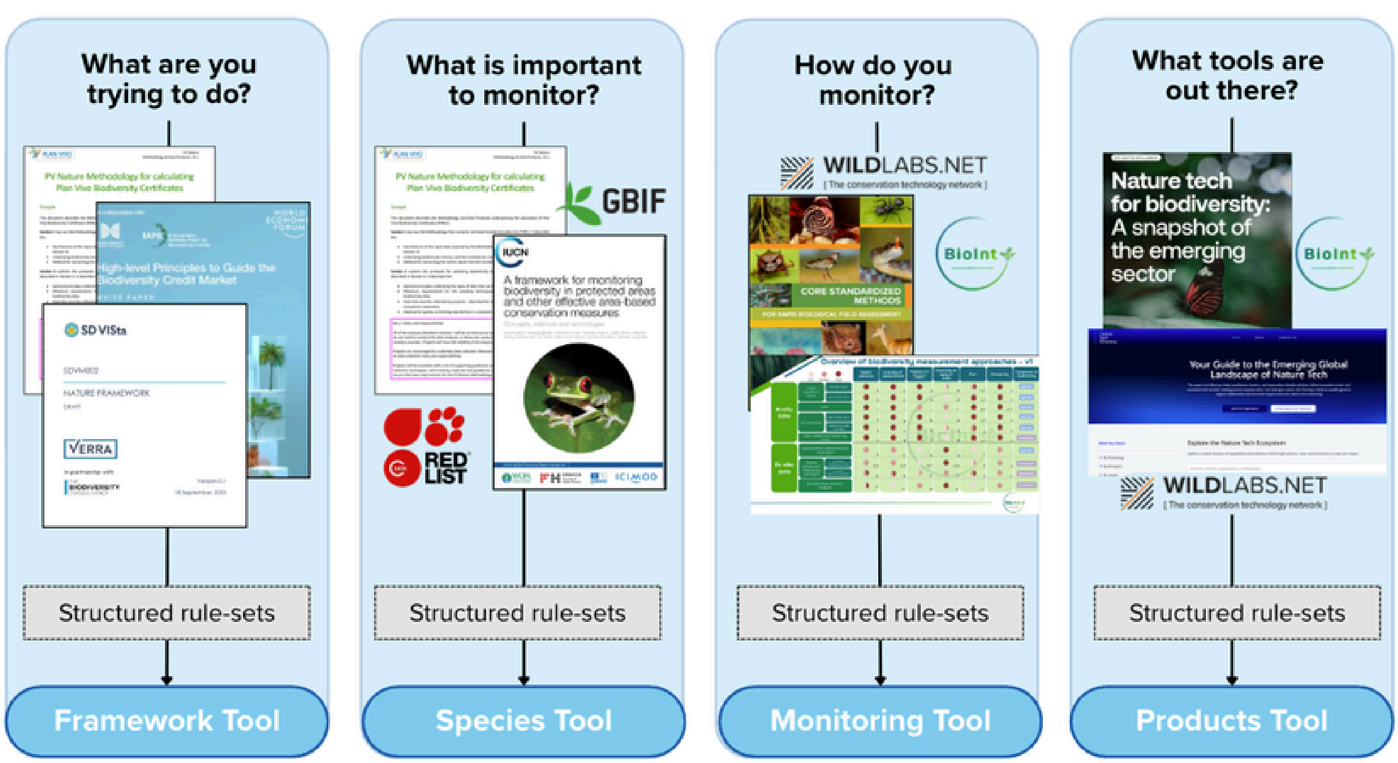

Figure 3. Knowledge Libraries

RESULTS

We tested the platform's performance in targeted interviews with nature credit practitioners. JRAG-augmented responses performed well on bounded, factual questions, such as whether a project is eligible for a given framework. This suggests JRAG may have broader value for domains where AI systems need to reason against complex but explicit rules – such as in the more established carbon credit or environmental policy compliance space.

The harder challenge was with less structured questions with which practitioners needed the most help: how to weigh frameworks against each other, how to navigate tradeoffs between monitoring rigor and cost, and how to make defensible decisions when there is no clear right answer. On these questions, the agent tended to be vague or would try to fit all possible frameworks to the use case rather than help users reason through the decision. This eroded user trust, making it difficult for practitioners to rely on the agent as a meaningful decision-support tool.

Several factors likely contributed to this. This prototyping project operated under tight budget and timeline constraints and was intentionally exploratory in scope. Evaluation was also genuinely difficult, as complex biodiversity project design decisions rarely have a clear, black-and-white answer or a single "correct" response. That makes it hard to measure AI performance, identify failure modes, or attribute them to data gaps, reasoning limitations, or insufficient decision structure in the prompts.

THE CRITICAL GAP: FOUNDATIONAL KNOWLEDGE AND DESIGN LOGIC

This project revealed that the main barrier to using AI for effective nature credit project design is knowledge infrastructure rather than just the technology itself.

Unlike more mature sectors, biodiversity credit markets have few published, standardized project examples for AI models to learn from. Instead, methodological frameworks are proliferating faster than practical application can consolidate, and processes vary widely by geography, ecosystem, credit type, and stakeholder preferences.

© Diego Pérez - SPDA / Conservación Internacional

Contrast this with fields such as software development. AI coding assistants perform well not just because the underlying models are capable, but because code exists within rich, structured infrastructure: explicit syntax rules, vast repositories of documented examples, and clear feedback mechanisms like tests and compilers. That structure gives AI something reliable to reason against — and practitioners something reliable to verify against. Biodiversity project design lacks an equivalent foundation.

Our implementation of JRAG successfully improved information retrieval by imposing structure on existing data, but it could not fix a fundamentally inconsistent knowledge base. Consequently, the agent excelled at bounded, specific tasks, but lacked the comprehensive grounding required to navigate the end-to-end complexities of early-stage project design.

WHERE WE GO NEXT

This project produced a working prototype, a novel RAG technique, and a significantly clearer view of where the real barriers lie. Three priorities emerged for the next phase of work:

1. CURATE CRITICAL KNOWLEDGE

The most important near-term investment is not in AI systems themselves, but in assembling and structuring the underlying knowledge base — specifically the tradeoff landscape between monitoring rigor and cost across different project types, technologies, ecosystems, and credit frameworks. This is the material AI would need to draw on to give practitioners genuinely useful guidance.

2. BUILD EVALUATION BENCHMARKS

Improving AI performance requires being able to measure it. Complex biodiversity design questions often have no consensus right answer, which makes it difficult to build the benchmarks needed to assess whether a system is actually helping. The field needs agreed-upon evaluation standards to make meaningful progress.

3. DEVELOP SHARED INFRASTRUCTURE ACROSS THE SECTOR

None of this can be done well in isolation. Progress requires common evaluation benchmarks, open documentation of agent architectures, collaboratively maintained knowledge bases, and curated examples of applied projects. The knowledge assembly work alone is substantial and demands a collective effort.

AI has genuine potential to accelerate biodiversity conservation outcomes, but realizing that potential depends on the domain knowledge and shared infrastructure that make reliable AI assistance possible in the first place.

FOR MORE INFORMATION

Read more about the full project:
https://earth-genome.github.io/ci-agent-app/the-knowledge-problem.html

Interested in exploring this space together?

Ali Swanson
Director, Nature Technology & Innovation
AI Hub
aswanson@conservation.org

CONSERVATION INTERNATIONAL

ACKNOWLEDGEMENTS

This work was made possible by the generous support of the Patrick J. McGovern Foundation. It was led by Conservation International (Ali Swanson, Carly Batist, Emily Nguyen, and Jack Seevers), in partnership with Earth Genome (Mikel Maron and Dan McCarey), who built the prototype and developed the JRAG methodology, and the Nature Tech Collective (Gilad Goren, Owen Dehmler-Buckley), who helped define the user needs and whose prior collaboration with CI inspired the project. This work builds on the Nature Tech Collective Biodiversity Sector Map and looks ahead to their broader vision for a comprehensive Nature Tech Directory. We are grateful for their enduring collaboration in the effort to fully harness tech for nature positive impact.

We are especially grateful to Joshua Berger of the Biodiversity Intelligence Company, whose partnership on the uses, costs, and complexities of different technologies scaffolded the prototype's technology-selection component. We also thank Simas Gradeckas of Bloom Labs, the WILDLABS team, the many conservation and nature finance practitioners whose candid feedback and thought partnership directly shaped this work.

© Daniel Rosengren