

The ALATCHA Project: Improving AI Reliability and Accelerating Mathematical Research at UNR and Beyond

Lead Applicant Information

- Name: Dr. Christopher L. Rogers
- Title: Associate Professor of Mathematics
- Department/Unit: Department of Mathematics and Statistics (DMS)
- Email Address: chrisrogers@unr.edu
- Phone Number: office: 775-682-7537

Additional Project Participants

The following UNR DMS participants will engage in mathematics research as well as testing and developing the proposed ALATCHA proof-discovery environment: Dr. Jonathan Beardsley (Asst. Prof), Dr. Sarafina Ford (Postdoc), Logan Beier (M.S. student), Landon Fox (Ph.D. student). Beier and Fox also have B.S. degrees in CS; they will coordinate closely with Dominik Farr on software development. The external participants are: Dominik Farr (AI software engineer and industry consultant dffarr.com), Dr. Eric Peterson (senior AI scientist and industry consultant chromotopy.org). Farr is responsible for coding the proof-of-concept ALATCHA demo, and will work with students Beier and Fox to get them up to speed. Peterson’s ten years of AI industry experience includes technical and scientific positions at OpenAI, Amazon Web Services, and IBM Research.

Project Description and Objectives

Within the context of AI safety, the goal of AI **reliability** concerns building models that reason carefully and precisely in order to prevent hallucinatory phenomena. Recently, AI labs¹ working at the frontier of such efforts have started training their models to perform rigorous abstract mathematical reasoning. The expectation is that the model will begin to solve problems in a way that imitates how mathematicians logically string together formal definitions, axioms, and postulates in order to construct proofs of new statements: i.e., theorems, corollaries and lemmas. This kind of training is called **formalization**, and it employs a special programming language called **LEAN** which is used to deterministically verify that the model is “writing” its statements and proofs in a mathematically correct way. Ironically, it is here where formalization and research mathematics begin to part ways. Indeed, LEAN requires all statements to be written in a highly constrained formal language. Professional mathematicians, in contrast, think rigorously but communicate using the informal prose of **natural language**. We place a premium value on the underlying intuitive ideas, or “conceptual seeds”, from which a mathematical proof originates, while syntactical fastidiousness, on the other hand, is less appreciated. However, due to unexpected accelerated advancements in LLMs in mid-2025, we have begun to see the advantages of **AI-assisted research** in pure mathematics². Substantial obstacles, nevertheless, remain. For example, most mathematicians are unaware of LEAN, and at the same time, we rightfully do not trust LLM-generated mathematics unless every output is verified.

The PI and his collaborators propose a novel approach towards meeting the dual challenges of formalization both for the sake of AI reliability and AI-assisted research mathematics. **ALATCHA**³ is a human-guided, multi-model, LEAN-verified proof-discovery environment. We are interested in proof discovery specifically in the fields of higher category theory and homotopical algebra. Heuristic evidence suggests that technical lemmas in these two areas are particularly amenable to formal

¹Reuters, “Robinhood CEO’s math-focused AI startup Harmonic valued at \$1.45 billion,” Nov. 25, 2025

²Konstantin Kakaes, “The AI Revolution in Math Has Arrived,” *Quanta Magazine*, Apr. 13, 2026.

³Automated Lemma Assistant Training for Categorical and Homotopical Algebra

proof search, and both areas comprise the research specializations of the PI, Beardsley, and the rest of the Algebra Research Group at UNR. The success of this proposal will allow our group, in collaboration with our AI industry partners, to build and evaluate a prototype system that gives UNR mathematicians the ability to convert real research lemmas into LEAN-verified formal statements, guide proof search using conceptual mathematical insight, and generate structured success/failure data which will be highly valuable for both mathematics and AI reliability.

ALATCHA will provide a front-end for a novel **three-stage workflow** designed to integrate with a research mathematician’s work habits and intuition. It will be distinct from currently available LEAN-based auto-formalization software and **novel** in several key ways.

Stage I will convert a natural-language lemma into a proposed LEAN-formalized target statement through an interface requiring minimal LEAN skill and aided by hidden prompt-engineered **preambles** designed to optimize the fidelity of the formalization process and minimize drift. In particular, we posit that mathematicians can judge the quality of an existing formal translation more easily than writing it from scratch. Stage II involves a novel process called **conceptual seeding**: the user inputs proof-strategy hints based on their intuition for why the statement should be true. Tailored preambles then instruct the agents in Stage III to incorporate this hint into their own proof strategy. Indeed, in Stage III, an agentic proof-search loop calls commercial LLMs (hence, **multi-model**) which interface via MCP with a **new deterministic** proof search tool called **AXLE**, recently developed by the AI lab Axiom Math in March 2026. Regardless of whether a verified proof is found, or if the agents give up, the results are stored in what we will call the **Lemma Ledger**. This is a structured data repository containing both successful and failed attempts. In particular, the failed proof attempts can be easily reattempted when newer models become available. A demo video showing how ALATCHA is used in practice is available at: hoalg.net/chrisrogers/ALATCHA-Demo. Note that while this is an actual functional proof-of-concept, it is not yet a completed prototype.

Alignment and Impact

This proposed research strongly aligns with President Sandoval’s characterization of the PACK AI program in August 2025 as a “student-driven artificial intelligence initiative”. The project participants include two UNR faculty, their postdoc and thesis students, thus providing an invaluable career training opportunity. Indeed, in 2025, LinkedIn⁴ reported that, among paid job posts, “mathematics specialist” was the fastest-growing job-demand category. More broadly, the ALATCHA Project aligns with UNR’s Strategic Plan by strengthening graduate training and research infrastructure, and enhancing UNR’s national reputation. As mentioned, this project responds directly to a national shift in mathematical research. Tools for AI-assisted proof verification are moving quickly, and the research discovery rate in pure math will likely soon accelerate significantly. Funding this proposal will place UNR ahead of the curve.

Project Plan and Timeline

July–September 2026: Develop the ALATCHA prototype to handle daily use by approximately 6–10 faculty and student researchers working in the Algebra Group, including deployment of a secure web interface; implementing access control; development of the Lemma Ledger; selection of the first 5–6 lemmas for verification; and obtain any needed IRB or data-governance certification.

October–December 2026: Add OpenAI, Anthropic, and Google API access, develop model-comparison metrics, and cost-ceiling controls. Integrate and optimize the AXLE proof verification and declaration extraction tools. Run controlled trials on the first 8–10 lemmas under two conditions: accepted LEAN target alone versus accepted target plus conceptual seed. PI Rogers and participant

⁴<https://www.linkedin.com/business/talent/blog/talent-strategy/most-in-demand-jobs>

Beardsley will incorporate work on ALATCHA into NSF DMS Topology division grant proposals (mid-November 2026).

January–March 2027: Expand the repository to 15–16 lemmas. Begin experimental development of so-called **compressed context dossiers**, whose goal is to summarize local definitions, notation, examples, and known new theorems, so models can recover project-specific background without exhausting context windows. Begin **mathlib** (LEAN Mathematical Library) submissions for broadly useful definitions and lemmas, which will be crucial for proofs in higher category theory. Present preliminary results in UNR colloquia and symposiums.

April–June 2027: Complete our proposed 20 lemma pilot goal for the Lemma Ledger; analyze proof success rates, costs, human interventions, and conceptual-seeding effects; finalize the Lemma Ledger repository and private cloud archive; prepare curated exports; anticipate participation in workshops and conferences focused on formalization and AI-assisted mathematics research. Submit final report plus external-funding plans.

Use of Award Funds

Award funds will support infrastructure, software access, and cloud services needed to convert the proof-of-concept into a usable research platform. Graduate student summer support: Students will test real lemmas, write and review LEAN code, and contribute to mathlib. Linux laptops will provide standard environments for development and testing. GPT Pro subscriptions and API credits for OpenAI, Anthropic, Google, and AXLE/Axiom will be crucial to ALATCHA. These are needed to support model comparison and research into multi-model proof search. All processed data in the Lemma Ledger will be stored in a NoSQL database hosted on Google Cloud using Customer-Managed Encryption Keys (CMEK).

Expected Outcomes and Sustainability

Deliverables are the working ALATCHA web platform pilot with multi-model proof search and verification via Axiom Math’s AXLE toolkit; approximately 20 real working-mathematician lemmas attempted in higher categorical and homotopical algebra accepted as LEAN statements with a subset accepted as complete proofs; mathlib-submitted definitions and lemmas, undoubtedly a welcomed contribution to the formalization community; a controlled comparison of proof search with and without conceptual seeding; and the ALATCHA Lemma Ledger repository. The Lemma Ledger will be a high-value outcome: curated exports can later provide AI-reasoning benchmarks.

Sustainability will come from external proposals and strategic industry partnerships. Both faculty members will submit NSF three year grant proposals by the November 2026 target date, which will include components supported by ALATCHA. Additional funding target sources include Harmonic’s Aristotle sponsorships (aristotle.harmonic.fun/sponsorships), direct contact with Axiom Math through an existing professional connection to Axiom leadership via the PI’s industry co-participants, and contact with DARPA’s expMath program through the PI’s collaborator on a different project, Prof. Jesse Wolfson at UC Irvine. These pathways will support larger-scale funding for development, dataset curation, and future experiments on fine-tuning mathematical reasoning models specialized for proof search in higher category theory and homotopy theory.

Budget and Justification

Graduate student support: 2 students, two summer semesters at \$2400 each plus fringe 15.8%	\$11,200
Two Framework Ultra X7 358H Linux laptops for grad students; optimal “future-proof” price point for use in other projects	\$4,000
GPT Pro access for 5 UNR participants and D. Farr for one year.	\$14,400
OpenAI GPT 5.5, Anthropic Claude Opus, Google Gemini, and AXLE API credits	\$16,400
Google Cloud storage, NoSQL database, backups, and web hosting.	\$4,000
Total	\$50,000