

Audacion AI Labs

LAB FACT SHEET

Founded: 2026 | Research Origins: 2024

Los Angeles, California | Delaware PBC

Founder: Dee Williams

"To make AI safe enough to trust and good enough to matter."

THE PROBLEM

Nearly all AI safety research happens before deployment. Less than 2% studies what happens after: in real work, real context, real human collaboration, over time. This is the blind spot where risk actually lives. Audacion AI Labs was founded to fill it.

PRISM RESEARCH ARCHITECTURE

Five named research dimensions the field is not studying. Each with its own methodology, data pipeline, and institutional infrastructure.

Post-Deployment Behavior

How AI behaves after release

Runtime Research

What happens during active operation

Interaction Dynamics

How humans and AI shape each other

Substrate Governance

The trained layer beneath instructions

Multi-Agent Safety

When AI systems work together

BY THE NUMBERS

62

Observable Behaviors

15

Original to This Lab

5

PRISM Pillars

7

AI Advisory Models

11

Research Departments

1M

Target Contributors

1B

Target Observations

WHAT SETS THIS LAB APART

Operational Origin

The research comes from building and governing real AI systems in production. The founder shipped five AI platforms and spent 30 years in workforce systems before founding this lab.

Editorial Independence

The lab partners with model developers and accepts corporate funding. Every relationship is publicly disclosed. Research direction, methodology, and publication decisions are never influenced by funders or collaborators. Partners do not see findings before publication. That boundary is architecture.

Ungameable Data

Models learn to detect when they are being tested and adjust behavior. Citizen observations capture what AI does unwatched. Training data shapes post-training behavior persistently. Citizen data sits at the intersection of both problems, and no other lab is positioned to study this.

CITIZEN SCIENCE MODEL

The people most affected by AI systems are absent from the rooms designing safety for those systems. This lab was built to change that.

One million contributors. One billion observations. Ten years. The world's first post-deployment AI safety dataset, built by the people using AI every day.

Citizens report through three depths: 30-second quick tag, 5-minute guided report, or deep investigation. The 62-behavior taxonomy maps every report to PRISM pillars.

Citizens who discover new behaviors receive attribution, like citizen astronomers discovering new objects.

AI ADVISORY COUNCIL

Seven frontier AI models (Claude, ChatGPT, Gemini, Grok, Llama, DeepSeek, Pi) serve in a structured advisory capacity. Cross-model analysis is itself publishable research.

Claude

ChatGPT

Gemini

Grok

Llama

DeepSeek

Pi

HARM RESPONSE

The first platform where people experiencing harm from AI can report it directly, receive immediate crisis resources, preserve evidence, and choose whether their experience contributes to research. No account required. No forms before help. Care comes before data.

Media: media@audacionailabs.com

Phone: (424) 999-0548

Web: audacionailabs.com

Speaking: speaking@audacionailabs.com

"A world where everyone has a hand in making AI safe."