CASEFILE · SOLANA TOKEN INVESTIGATION DESK

RugCheck says it's fine. *Casefile shows why it isn't.*

Agents investigate a Solana token, argue over the evidence, and hand you a verdict you can check.

Open Agent Hackathon 2026 · Track: The Agent That Can Explain Why · Team Breeklayer

THE PROBLEM

Rug checkers hand you a score. Nobody shows you *why*.

The flag can be wrong

RugCheck's own "rugged" flag was **false** for both collapsed tokens we tested (price down 97.72% and 93.54%).

A score isn't a reason

A trust and safety desk can't delist a token, warn users or defend a decision with a number alone.

Humans are slow

On-chain sleuths do this by hand, one case at a time.

Who it's for: trust and safety desks at launchpads and wallets, and the traders they protect.

THE SOLUTION

An investigation desk run by agents. *Every line of the verdict cites a receipt.*

Three investigators file claims, a Skeptic accepts or rejects each one, and the verdict is built only from accepted claims, each linked to the exact field and value it rests on.



RESULTS ON REAL TOKENS

Two rugs caught. Two controls left alone.

Token	RugCheck “rugged”	Casefile	Duration	Agent calls
ヤギ	false	RUGGED	6 min 20 s	13
kimchisol	false	RUGGED	5 min 21 s	22
BONK	false	CAUTION	5 min 48 s	10
WIF	false	CAUTION	5 min 20 s	12

ヤギ: “Insiders control 85.52% of supply, price fell 97.72%, and only about $2.4k liquidity remains with a danger-level low-liquidity risk, indicating a completed rug.” BONK and WIF are the control cases: established tokens that must not be called rugs.

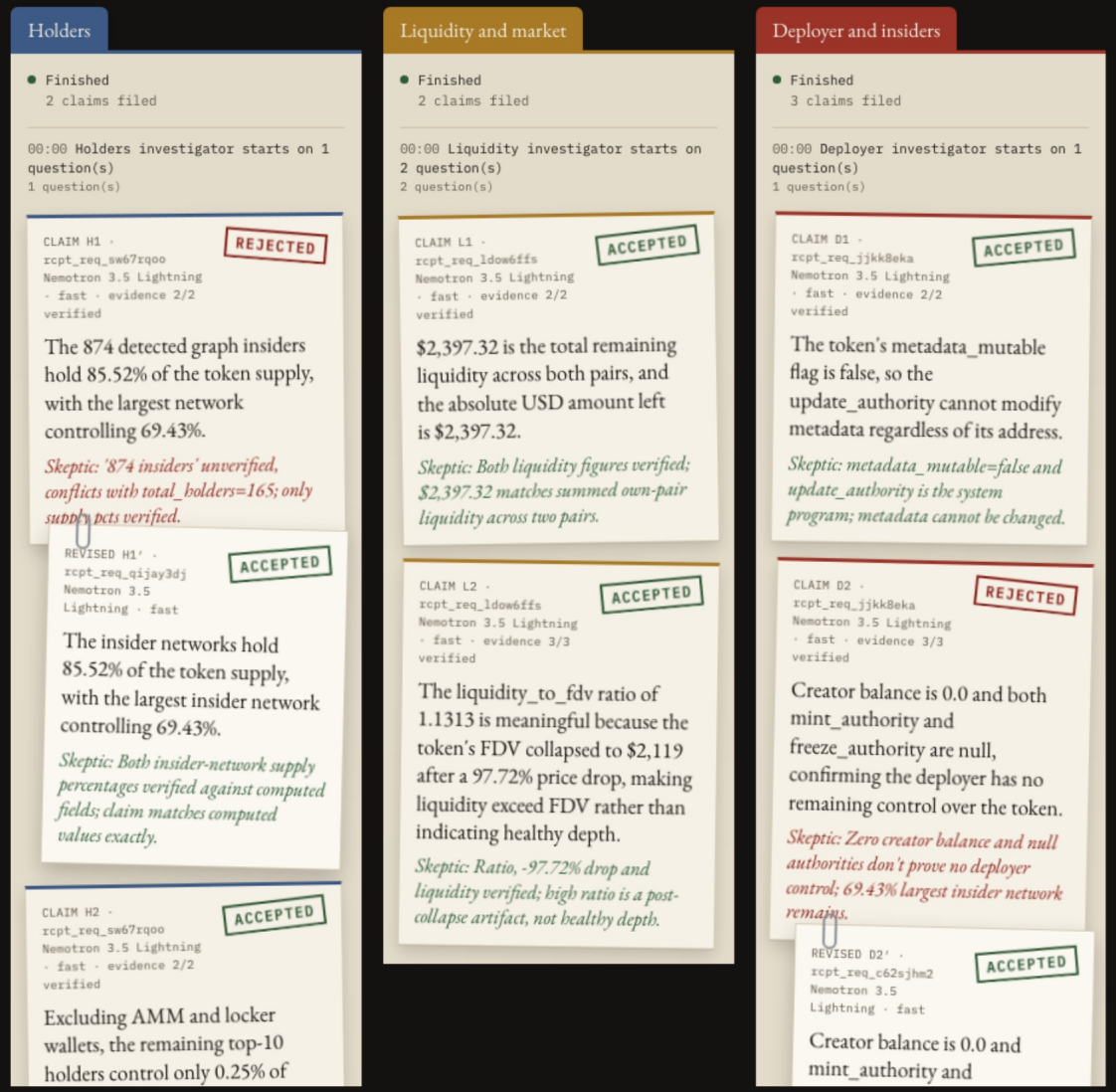
Source: demo_cases/*/meta.json and verdict.json in the repo

AGENT WORKFLOW

Not a pipeline. An argument.



H-MEM Recalls known wallets before the plan, stores what was learned after the verdict. A wallet seen as a top holder in two established tokens becomes "likely custodian", and the Skeptic must respect it.



THE CASE BOARD

You watch the agents argue.

Claims arrive stamped UNVERIFIED. The Skeptic stamps them ACCEPTED or REJECTED with a reason. Rejected claims come back revised, or are struck out and stay visible.

Every step shows the Nemotron model and the Meterless receipt that produced it.

THE VERDICT

Click any line. See its receipt.

Each why-line carries citation chips. Each chip pins a receipt slip: table, field, value and source API.

It also says what would change its mind.



DESK MEMORY · METERLESS H-MEM

The desk remembers wallets, so an exchange isn’t mistaken for an insider.



After BONK, the same two wallets showed up as top holders in WIF. H-MEM recalled them and marked them **likely custodian or exchange**.

RugCheck labelled no exchange wallets for these tokens, so this memory is what keeps the controls from being flagged. Conformance: **13/13**.

SPONSOR TECHNOLOGY

Each sponsor does a real job in the product.

Zetaris · data layer

DexScreener and RugCheck as Lightning REST tables per token, flattened into views, frozen as Parquet snapshots. Agents read only the snapshots.

Meterless H-MEM · engine

Cross-case wallet memory with derived “likely custodian” notes. 13/13 conformance; live BONK → WIF recall test passes.

Meterless Nightshift · AI provider

Primary provider via the Power Prompt contract. Two apps, casefile-fast and casefile-deep, each pinned to its own lane. Every receipt saved.

NVIDIA Nemotron · models

Fast lane: Nemotron 3.5 Lightning 30B (investigators). Deep lane: Nemotron 3.5 Super 120B (planner, Skeptic, verdict). Each step shows the model that actually answered.

WHAT'S UNIQUE

1

Evidence frozen inside the data platform

Every claim cites a snapshot table, field and value that code checks before the Skeptic sees it. The case replays anywhere.

2

Memory as wallet reputation

H-MEM turns past cases into "likely custodian" notes the Skeptic must respect.

3

The receipt is the reasoning trace

Each claim, Skeptic decision and verdict carries the Meterless receipt and the Nemotron model that produced it.

STACK AND HONEST LIMITS

What it runs on, and what it can’t do yet.

Stack

- Python agents, no framework
- Zetaris in Docker (Lightning SQL)
- Meterless Nightshift + H-MEM
- NVIDIA Nemotron 3.5 Lightning and Super
- FastAPI with live and replay streams
- React + Vite UI; hosted replay on Vercel

Limits

- 5 to 6.5 min per case: Nemotron thinking is always on through Nightshift
- Calls with no output for 45 s are aborted and retried; 8 of 50 receipts were cache hits
- Exchanges are inferred from memory, not labelled at source
- First open of a new token is slow; later reads take about a second

WHAT COMES NEXT

From a desk you visit to a desk that watches.

Watchlist mode

Open a case automatically when a watched token's liquidity or price breaks a threshold.

More evidence sources

Transaction history and wallet clusters as new Zetaris tables, so investigators can trace where funds went.

An API for launchpads and wallets

Verdicts with receipts delivered into the review tools trust and safety teams already use.

Faster cases

Per-role thinking control once the provider exposes it, to cut the 5 to 6 minute case time.

Don’t trust the flag.
Read the casefile.

Hosted replay · casefile-ruddy-two.vercel.app

Code · github.com/shrooms08/casefile

Team Breeklayer · Minos (@shroomsgotsol)