

Artisan

Researcher



Output

- Repro. Script

LLM Agent

Bash cmd & out

Bash Tool

Input

- Exec. Log
- Obfus. Table

Output

- Fmt. Table

Input

- Repro. Script

Output

- Repro. Result

Format Tool

Judge Tool

Exec-Based
Output Judging

LLM-Based
Output Judging