## Data Collection Evaluation instruction gold diff generated patch Task 1: Refactor the... Task 2: Prover that... repair & apply commits snapshots <commit-hash> generated files → 1a2b... expand & build → 3d4f... Lean restored content storage \*compile worktrees worktrees 1a2b... → -.lake -.lake\_ 3d4f... → LLM judge-01ab... → -Mathlib -Mathlib One-time Build On-demand Verification

