# 1 Task Generation **Policies** API graph Examples Domain Prompt data Template & metadata Persona Task = {query,actions,outputs}

# 2 Task Validation

# 1. Action Validation

- **Format Check Execution Check** Policy Check
- @@ -60631,11 +60631,11 @@ "name": "Smart Watch" "product\_id": "6945232052" "item\_id": "1706622510", "price": 328.67, "item\_id": "2860956907", "price": 315.61,

Regenerate

on validation

failure

#### Aggregate Diff Patch

### 2. Alignment LM Committee {Task data, Query, Diff patch} Reflect & Score V<sub>X</sub> avg\_score > threshold Validated 3. Final Review Task Reflections & Summarize

Realign task

using feedback

## 3 Simulated Agent-Human Interplay Prompt to reveal query details in turns Test Agent Human **↓** Env LM ✓ State based reward Output based reward / Successful Trajectory