2 Function Scraping Exclude Trivial Functions Repositories, 30,654 Functions

(1) Repository

Selection

(4) Automatic

Annotation

(5) Benchmark

GitHub

68 Repositories, 31,022 Functions

High-quality Repositories

Static Analysis: Signature, Reference

Code, Reference Dependency

Select Functions Based on the

Successful Test Cases
Successful Test Cases

25 Repositories, 301 Functions

- LLMs: Requirement, Domain Label
- 25 Repositories, 301 Functions
- Construction Real-world Program Distribution
 25 Repositories, 275 Functions