Related Papers

Existing Bug Repositories

1. Defects4J: A Database of Existing Faults to Enable Controlled Testing Studies for Java Programs
   https://dl.acm.org/doi/10.1145/2610384.2628055

2. Bugs.jar: A Large-scale, Diverse Dataset of Real-world Java Bugs
   https://ieeexplore.ieee.org/document/8595167

3. Defexts: A Curated Dataset of Reproducible Real-World Bugs for Modern JVM Languages
   https://ieeexplore.ieee.org/document/8802677

Training Methods

1. CodeBERT: A Pre-Trained Model for Programming and Natural Languages
   https://aclanthology.org/2020.findings-emnlp.139/

Bug Injection

1. Bug Synthesis: Challenging Bug-Finding Tools with Deep Faults
   https://dl.acm.org/doi/10.1145/3236024.3236084

2. Semantic Bug Seeding: A Learning-Based Approach for Creating Realistic Bugs
   https://dl.acm.org/doi/abs/10.1145/3468264.3468623

3. FixReveRteR: A Realistic Bug Injection Methodology for Benchmarking Fuzz Testing
   https://www.usenix.org/conference/usenixsecurity22/presentation/zhang-zenong