Rethinking Issue Resolution for AI/ML Systems

Ahmed Adnan



Mushfiqur Rahman



Antu Saha



Oscar Chaparro



IIT University of Dhaka

William & Mary
CHARTERED 1693

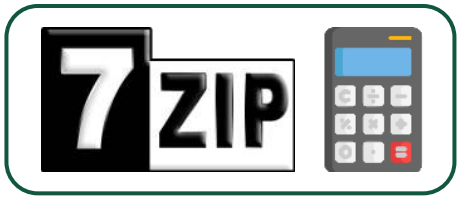
Software Evolution and Analysis Lab

Software Systems

Traditional Software



Follows hard-coded rules to produce exact, predictable results



Software Systems

Traditional Software

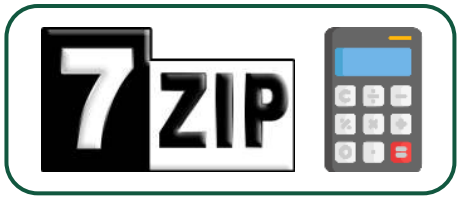


Follows hard-coded rules to produce exact, predictable results

AI/ML-integrated Systems

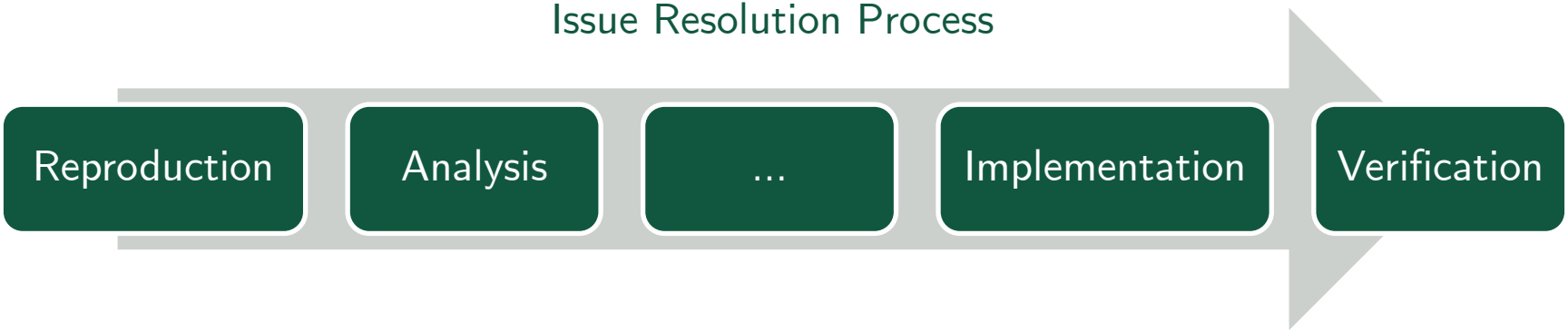


Learn from data to make probabilistic predictions



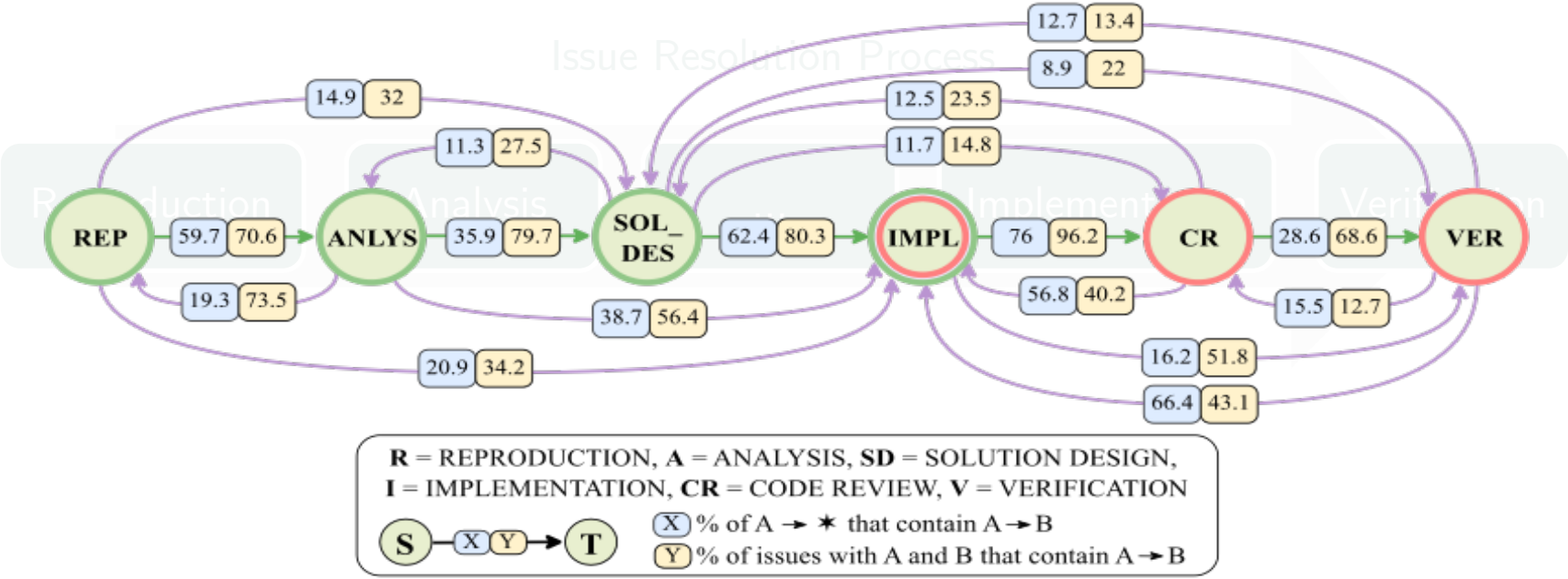


Issue Resolution in Software Systems



Issue Resolution in Software Systems

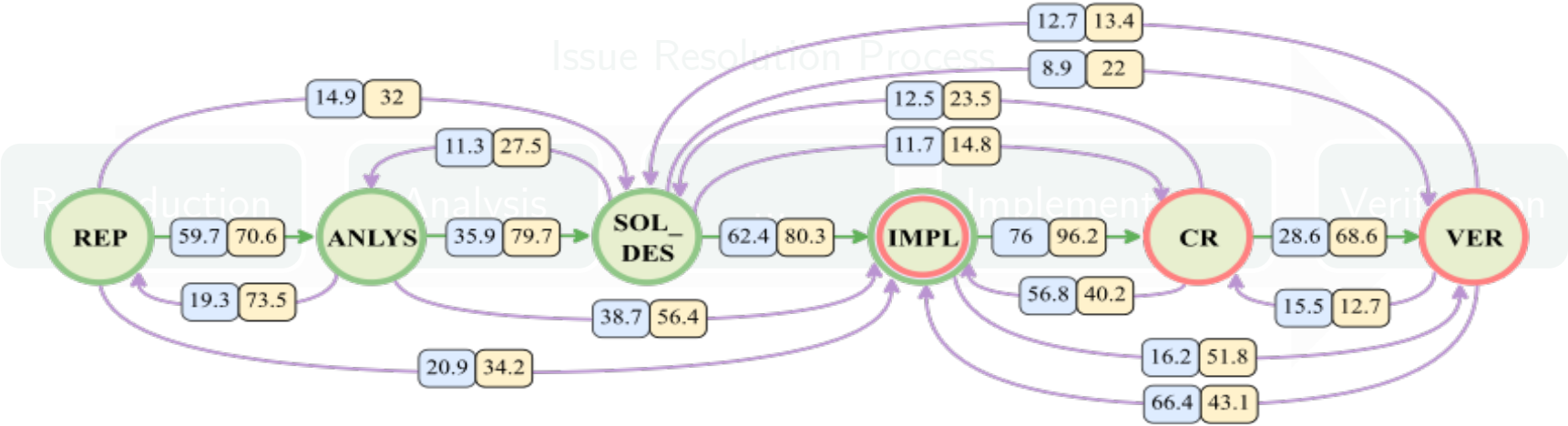
Issue Resolution Process at Mozilla Firefox[1]



1 Saha, A., & Chaparro, O. Decoding the issue resolution process in practice via issue report analysis: A case study of Firefox. ICSE'25.

Issue Resolution in Software Systems

Issue Resolution Process at Mozilla Firefox[1]



How about the issue resolution process in AI/ML-integrated systems?

1 Saha, A., & Chaparro, O. Decoding the issue resolution process in practice via issue report analysis: A case study of Firefox. ICSE'25.

AI/ML Artifacts Sit Inside the Codebase



AI/ML Issues Break Traditional Assumptions

Traditional Software

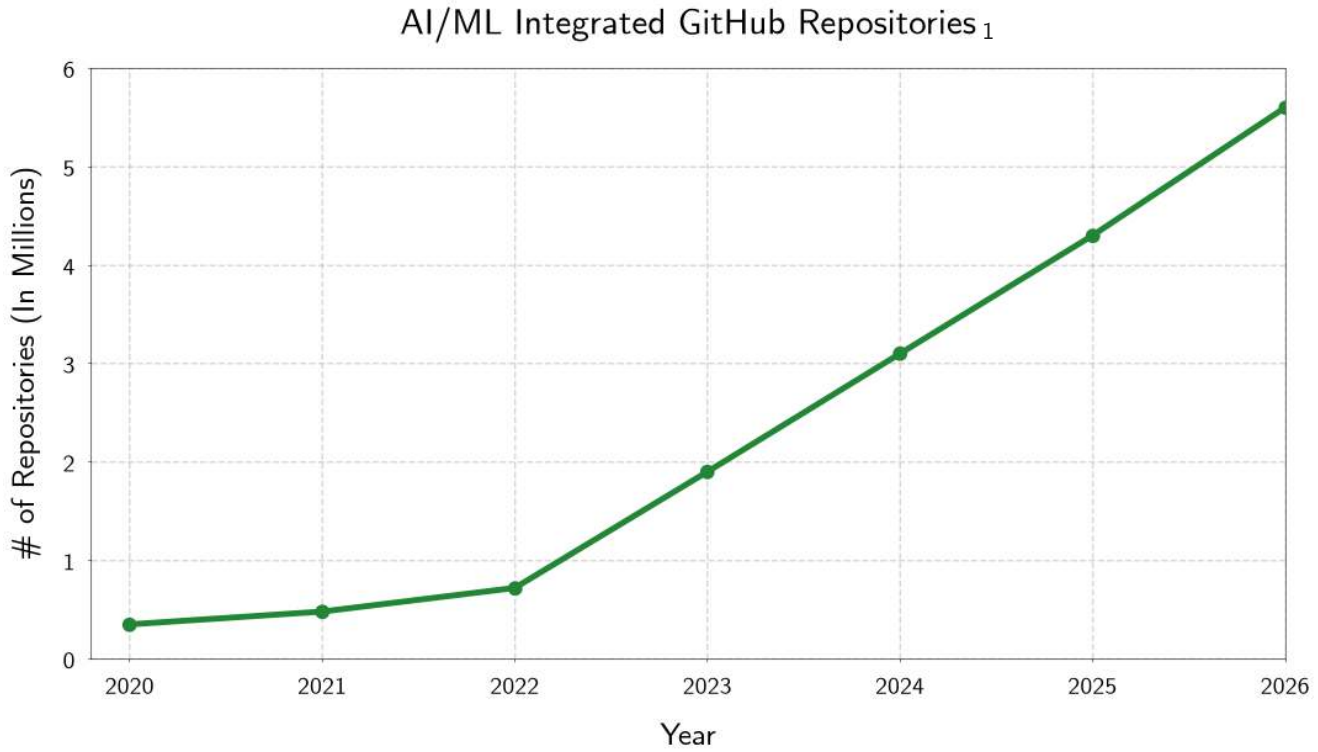
- Deterministic failures
- Code-centric debugging
- Binary pass / fail tests
- Sequential resolution stages

V
S

AI/ML-integrated Systems

- Stochastic behavior
- Heterogeneous artifacts (data, models, prompts)
- Statistical & qualitative evaluation
- Iterative, cross-stage experimentation

Increasing Adoption of AI/ML Components



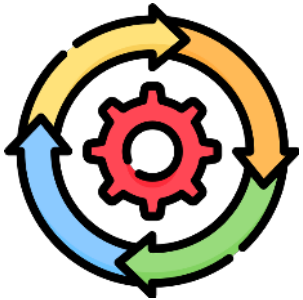
1 Data source: GitHub Innovation Graph. (https://innovationgraph.github.com/)

Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems

Our Vision

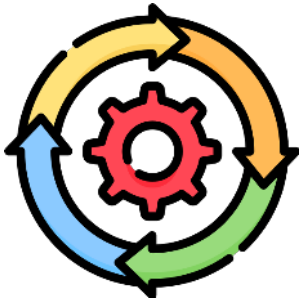
Adapted issue resolution frameworks for AI/ML-integrated systems



AI/ML specific Activities

Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems



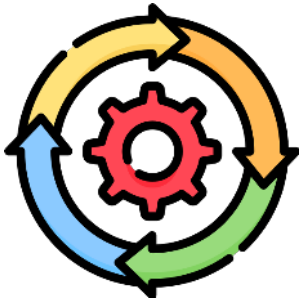
AI/ML specific Activities



Heterogenous Artifacts

Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems



AI/ML specific Activities



Heterogenous Artifacts



Non-deterministic Reproduction & Verification

Research Methodology

4 AI/ML Projects



100 Issue Reports



Research Methodology

4 AI/ML Projects



100 Issue Reports



Four Researchers



Activities & Challenges



Qualitative Analysis



Recurring AI/ML-Specific Activities

Traditional Issue Resolution Stages



Reproduction



Analysis



Solution Design



Implementation



Code Review



Verification

Recurring AI/ML-Specific Activities

Traditional Issue Resolution Stages


Reproduction


Analysis


Solution Design


Implementation


Code Review


Verification

AI/ML Specific Issue Resolution Activities

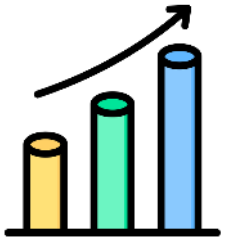

Model Performance Monitoring
22 Issues (27%)


Parameter Tuning & Training
9 Issues (11%)


Data Modification
6 Issues (7%)


Model Functionality Analysis
4 Issues (5%)

Issue Resolution Challenges in AI/ML Systems



Scaling fixes to larger datasets or distributed training

22 Issues (27%)



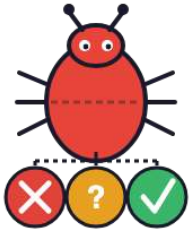
Understanding model architectures

9 Issues (11%)



Nondeterministic solution verification

6 Issues (7%)



Nondeterministic reproduction

4 Issues (5%)

AI/ML Activities Span Multiple Stages

Challenges

Activities

Stages

AI/ML-Tailored Issue Resolution Frameworks

01

Iterative & Cross-Stage Workflows

02

Reproducibility & Performance-Aware Verification

03

Data-, Model- & Artifact-Aware Resolution

04

Coordinating Heterogeneous Artifacts

05

Human-AI Collaboration

AI/ML-Tailored Issue Resolution Frameworks

01



Iterative & Cross-Stage Workflows

02



Reproducibility & Performance-Aware Verification

03



Data-, Model- & Artifact-Aware Resolution

04



Coordinating Heterogeneous Artifacts

05



Human-AI Collaboration

AI/ML-Tailored Issue Resolution Frameworks

01



Iterative & Cross-Stage Workflows

02

Reproducibility & Performance-Aware Verification

03

Data-, Model- & Artifact-Aware Resolution

04

Coordinating Heterogeneous Artifacts

05

Human-AI Collaboration

AI/ML-Tailored Issue Resolution Frameworks

01

Iterative & Cross-Stage Workflows

02

Reproducibility & Performance-Aware Verification

03

Data-, Model- & Artifact-Aware Resolution

04

Coordinating Heterogeneous Artifacts

05

Human-AI Collaboration

Coordinating Heterogeneous Artifacts

04



Coordinating Heterogeneous Artifacts

- Traceability and provenance across artifacts
- Trace regressions back to evolving datasets, prompts, or runtime environments
- Understand how fixes propagate downstream across interconnected artifacts

Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems



Recurring AI/ML Activities



Heterogenous Artifacts



Non-deterministic Reproduction & Validation

Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems



Recurring AI/ML Activities

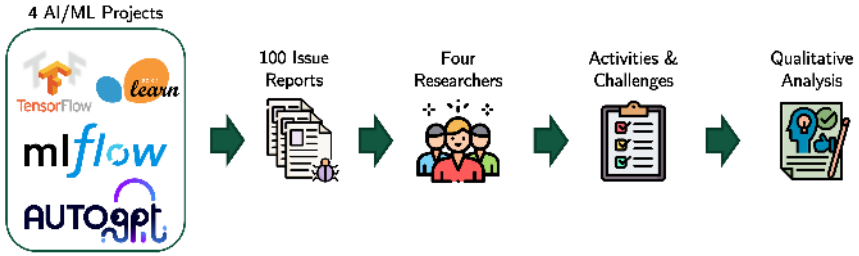


Heterogenous Artifacts



Non-deterministic Reproduction & Validation

Research Methodology



Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems



Recurring AI/ML Activities

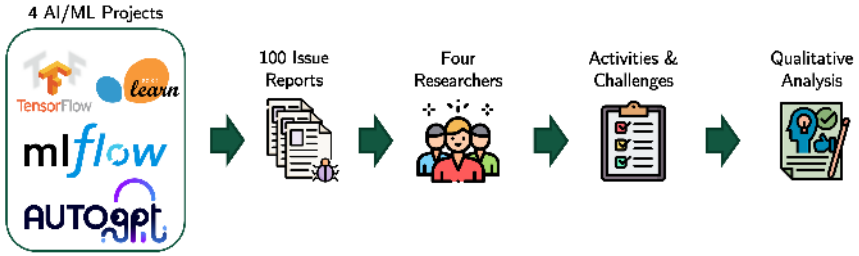


Heterogenous Artifacts



Non-deterministic Reproduction & Validation

Research Methodology



AI/ML-Tailored Issue Resolution Frameworks

01 Iterative & Cross-Stage Workflows

02 Reproducibility & Performance-Aware Verification

03 Data-, Model- & Artifact-Aware Resolution

04 Coordinating Heterogeneous Artifacts

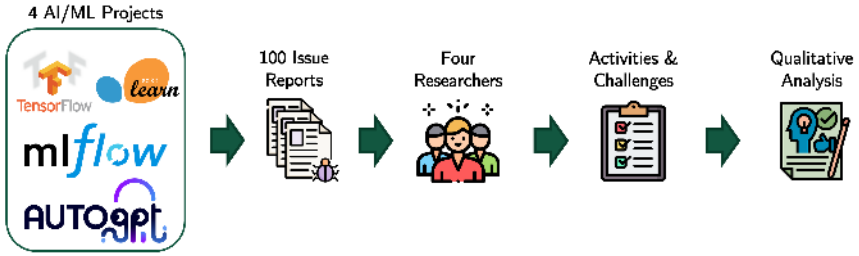
05 Human-AI Collaboration

Our Vision

Adapted issue resolution frameworks for AI/ML-integrated systems

- Recurring AI/ML Activities
- Heterogenous Artifacts
- Non-deterministic Reproduction & Validation

Research Methodology



AI/ML-Tailored Issue Resolution Frameworks

01 Iterative & Cross-Stage Workflows

02 Reproducibility & Performance-Aware Verification

03 Data-, Model- & Artifact-Aware Resolution

04 Coordinating Heterogeneous Artifacts

05 Human-AI Collaboration



SCAN ME

Let's get connected!

Antu Saha
asaha02@wm.edu