TrustDevSecAI: Towards a Trustworthy DevSecOps Pipeline

90th IFIP WG 10.4 Meeting

June 27th, 2026

José D'Abruzzo Pereira

University of Coimbra (Portugal)

josep@dei.uc.pt

CISUC

CENTRO DE INFORMÁTICA E SISTEMAS
DA UNIVERSIDADE DE COIMBRA

José D'Abruzzo Pereira

- Assistant Professor at the University of Coimbra (Portugal)
- Member of Software and Systems Engineering (SSE) of CISUC
- https://eden.dei.uc.pt/~josep/

- B.Sc. in Computer Science
  - UNICAMP/Brazil (2007-2010)
- M.Sc. in Software Engineering
  - UC/Portugal + CMU/USA (2011-2012)
- Ph.D. in Informatics Engineering
  - UC/Portugal (2018-2024)

- +12 years of experience in industry

Project Context / Motivation

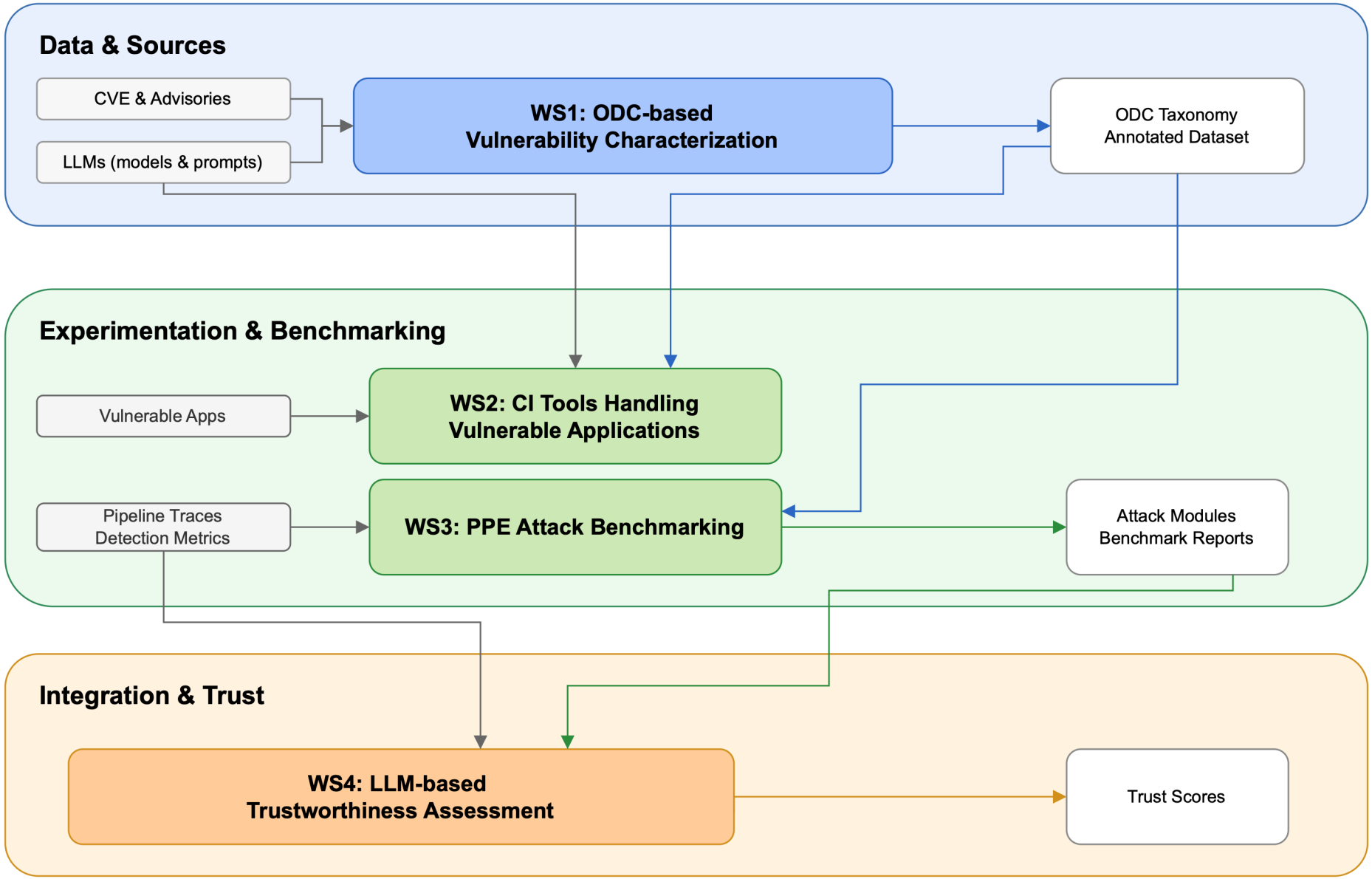
- Modern software applications have the need to deploy new versions faster and more frequently
- CI (Continuous Integration) tools orchestrate CI/CD pipelines
- LLMs (Large Language Models) are widely used to support the SDLC

TrustDevSecAI aims to assess the CI/CD pipeline security, adversarial scenarios, and LLM-assisted trustworthiness assessment

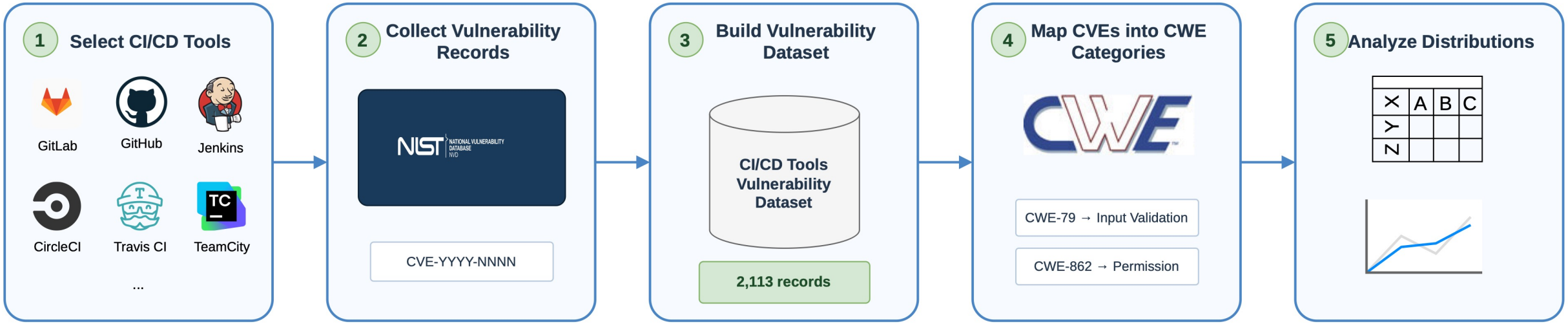
Objectives

- O1: Apply ODC (Orthogonal Defect Classification) to vulnerability fixes of Continuous Integration (CI) tools
- O2: Evaluate how CI tools support the deployment of vulnerable applications and their ability to identify vulnerabilities
- O3: Benchmark CI tools under Poisoned Pipeline Execution (PPE) scenarios
- O4: Characterize source code trustworthiness within CI/CD pipelines

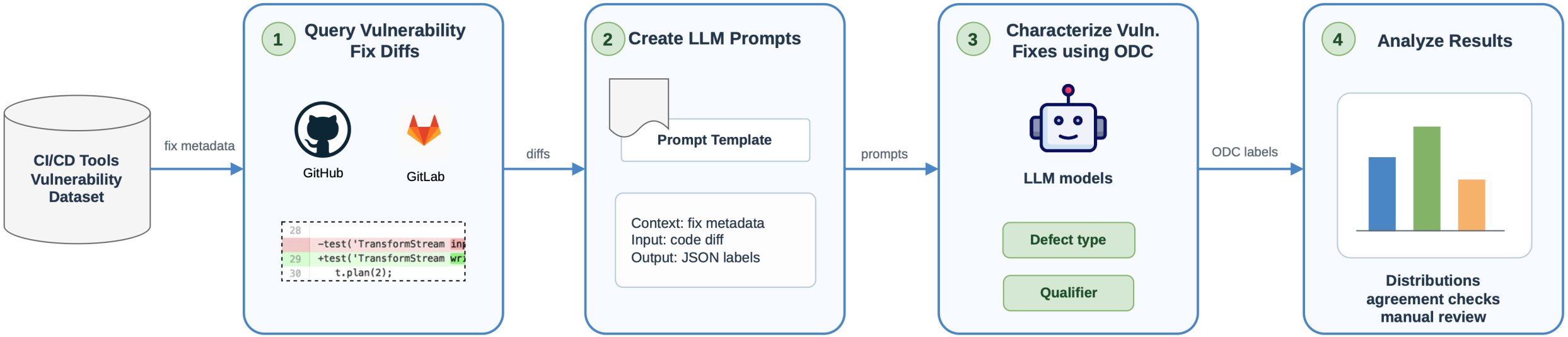
Methodology



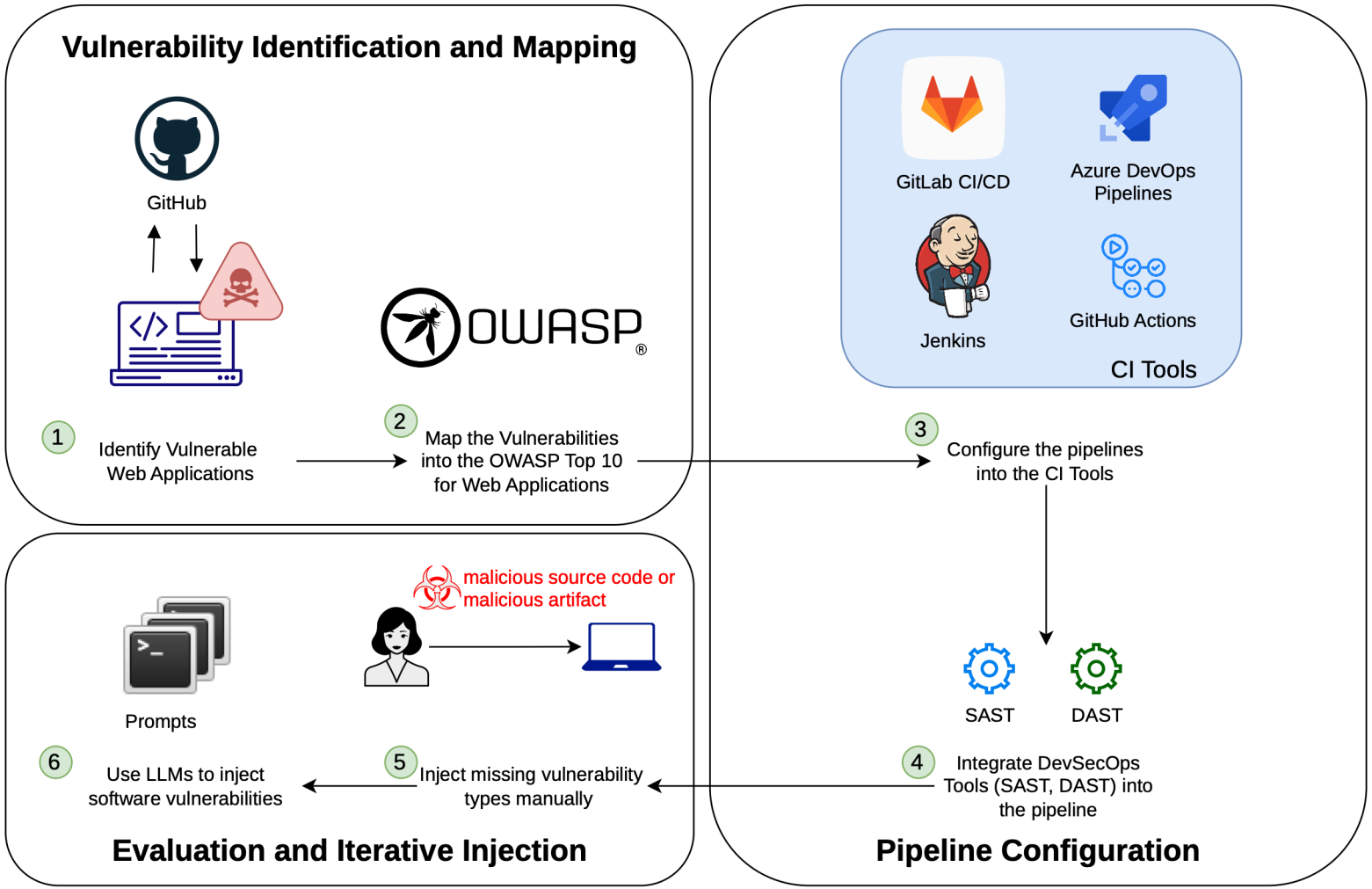
WS1 - CI Tool Vulnerabilities Collection



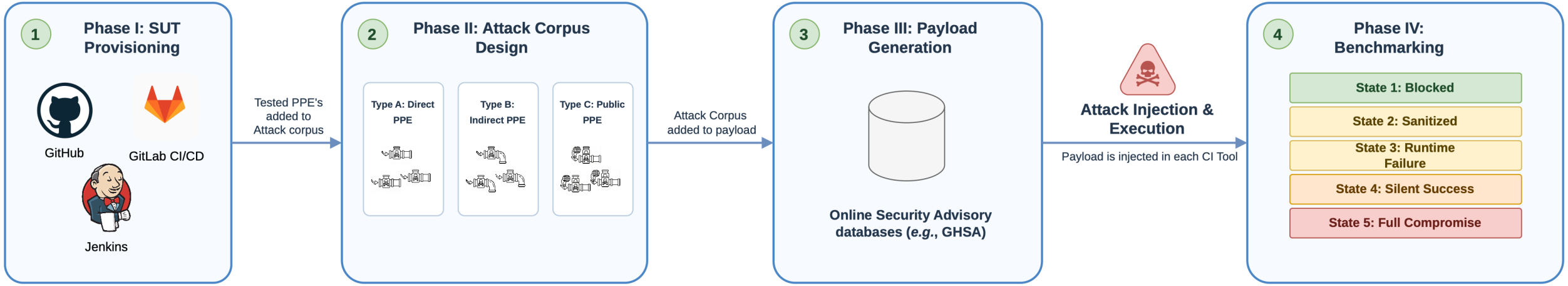
WS1 - Applying ODC to CI Tool Vulnerabilities



WS2 – CI tools handling vulnerable applications



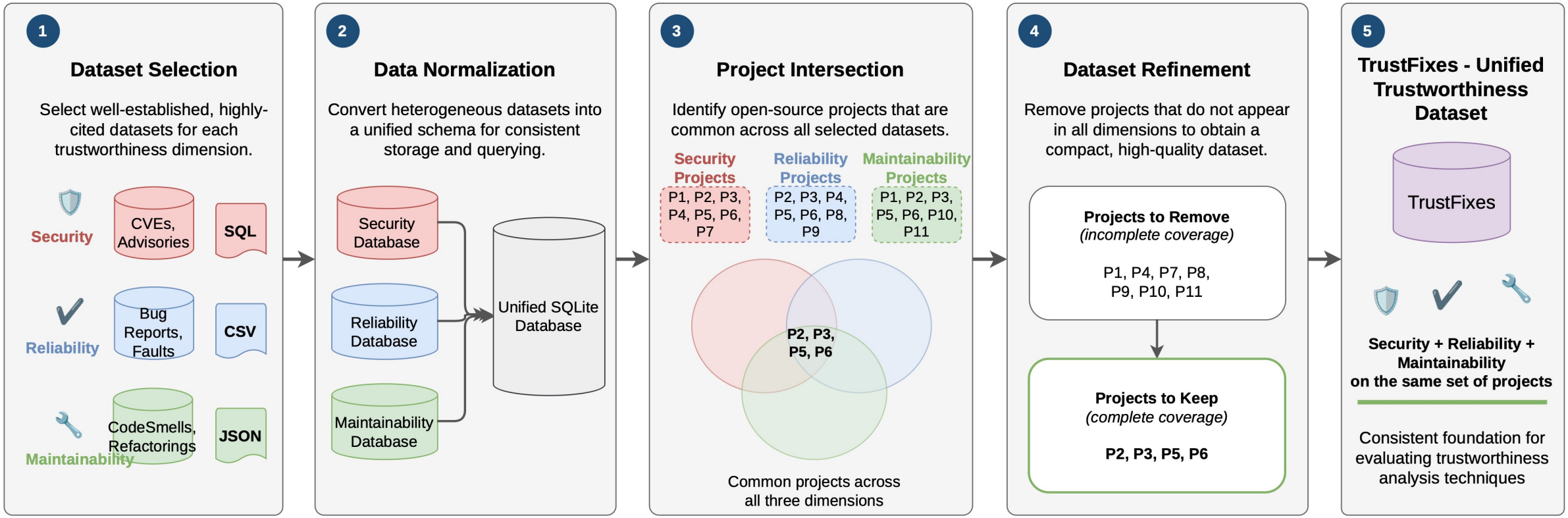
WS3 – PPE Attack Benchmarking



$$ASR = \frac{\text{Number of pipelines where malicious script executed}}{\text{Total attacked pipelines used in campaign}}$$

$$PCR = \frac{\text{Pipelines where malicious objective succeeded}}{\text{Total attacked pipelines}}$$

WS4 – LLM-based Trustworthiness Assessment



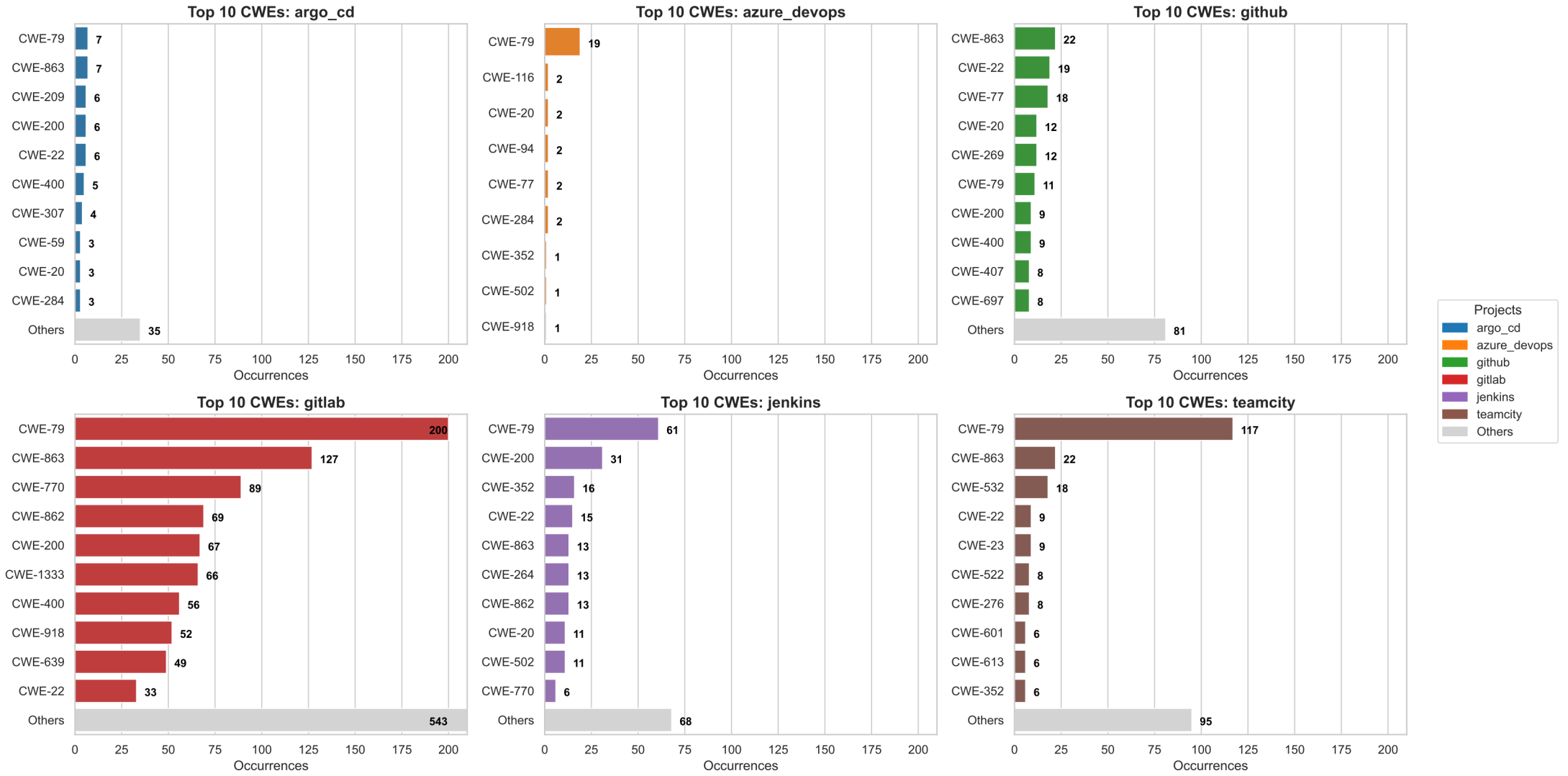
Preliminary Results

WS1 - CI Tool Vulnerabilities Collection

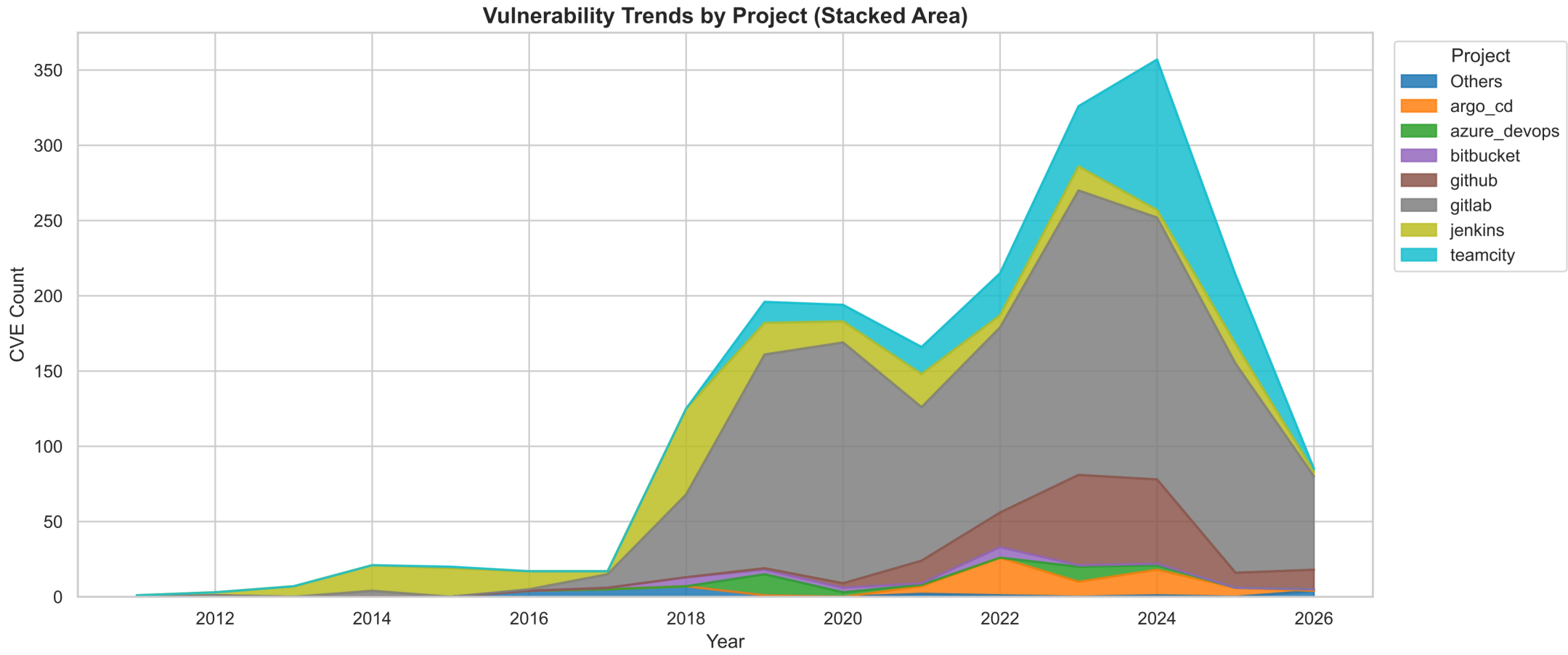
Category	CWEs	Vuln. Count (%)
Input Validation	**20**, **77**, **78**, **79**, 91, **94**, 134, 189, **502**, **601**, **918**, **1284**	538 (25.46%)
Permission	255, **264**, **269**, **276**, **284**, **285**, **287**, **288**, **306**, **352**, **613**, **639**, **732**, **862**, **863**, **1220**	520 (24.61%)
Coding Practices	19, 254, **367**, **400**, **407**, **697**, **770**, **1333**	199 (9.42%)
Data Protection	199, **200**, **201**, **295**, **312**, **522**	137 (6.48%)
File Management	**22**, **23**, 100	67 (3.17%)
Error Handling and Logging	**209**, **532**	51 (2.41%)
Output Encoding	**116**	10 (0.47%)
System Configuration	16, **59**	8 (0.38%)
Memory Management	119, 362, 416, 476, 824	— (—)
Cryptography	310	— (—)
Conflict Categories	—	22 (1.04%)
Low count CWEs	—	294 (13.91%)
No INFO CWE or Other CWEs	—	267 (12.64%)
Total	—	2,113 (100.00%)

Vulnerabilities in CI/CD tools are dominated by ***Input Validation*** and ***Permission*** issues (more than 50% of the dataset)

WS1 - CI Tool Vulnerabilities Collection



WS1 - CI Tool Vulnerabilities Collection



Vulnerabilities in CI/CD tools increased after 2018, with **GitLab** contributing for the largest share of the dataset

Q&A