CARIAD

Master Thesis: AI Usage in CI/CD/CT Pipelines for Compute Platforms in Automotives

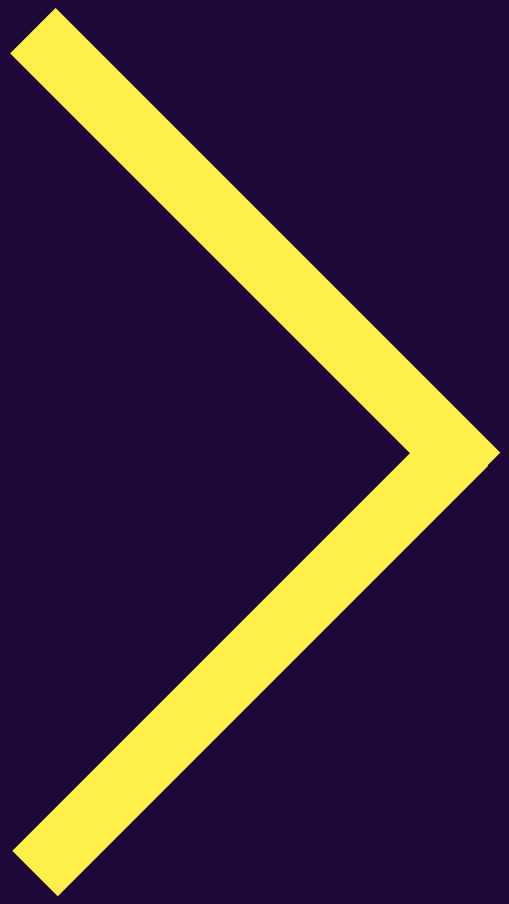
We transform automotive mobility

CARIAD
A VOLKSWAGEN GROUP COMPANY

Agenda

// Status
// Method
// Result
// Literature Review
// Next Steps

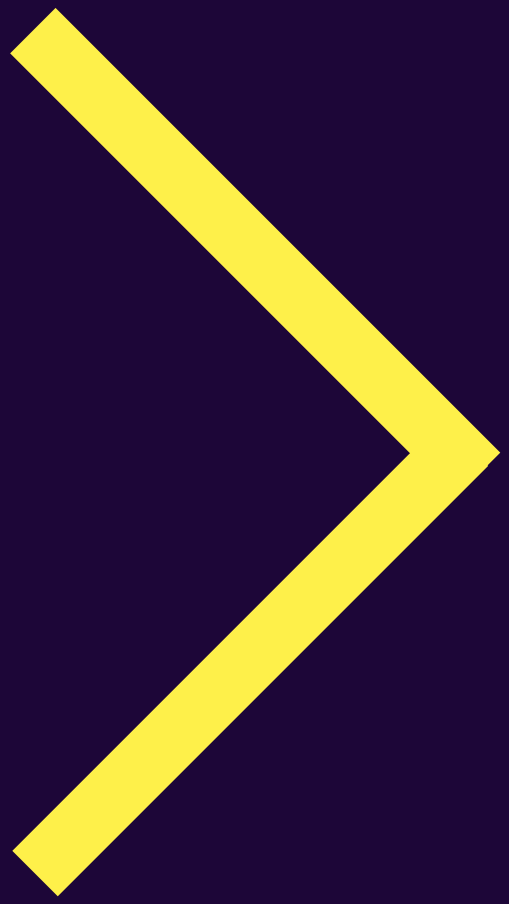
Status



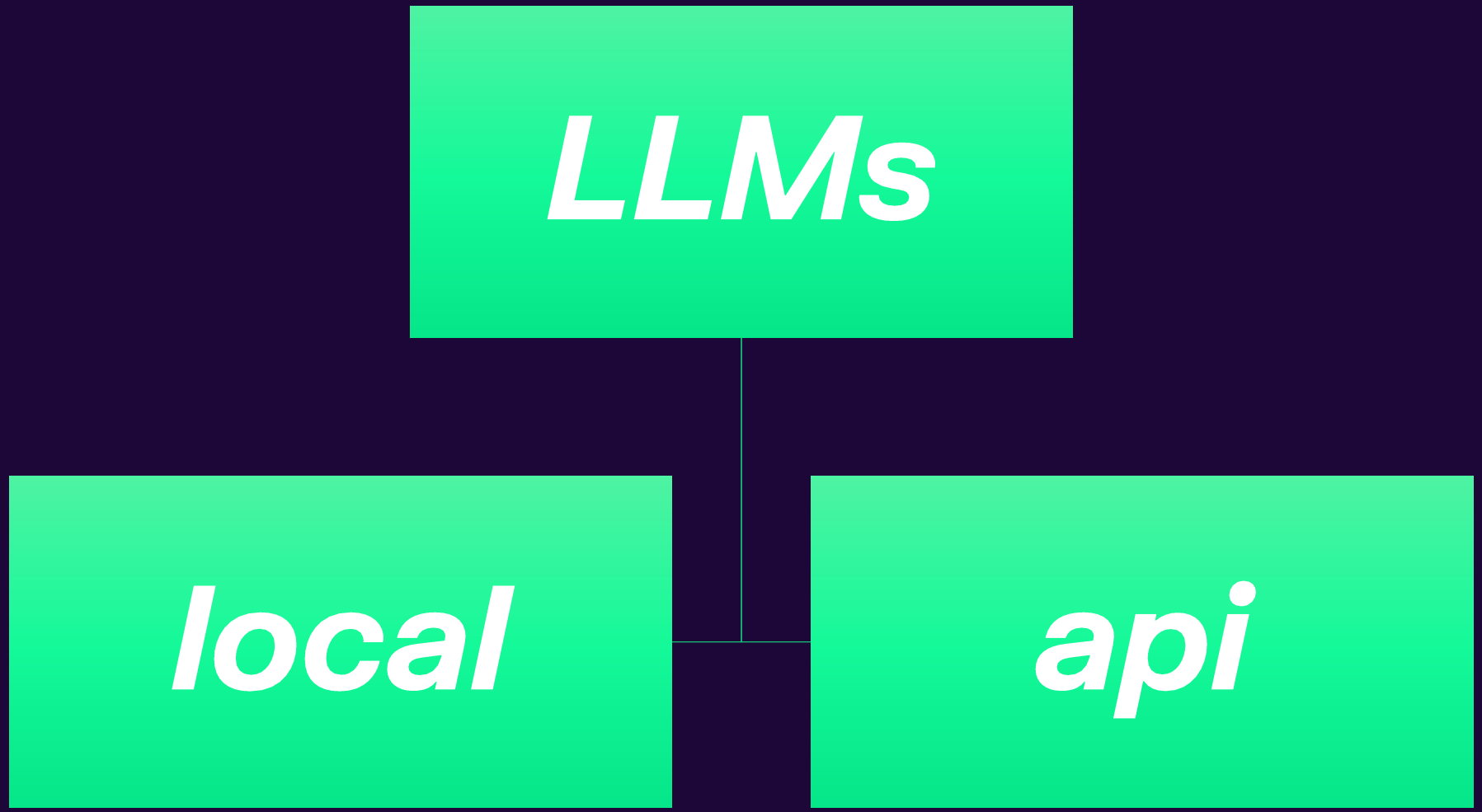
Status

- *All trainings completed*
- *All setups completed*
- *Working on integrating llms to cifuzz*

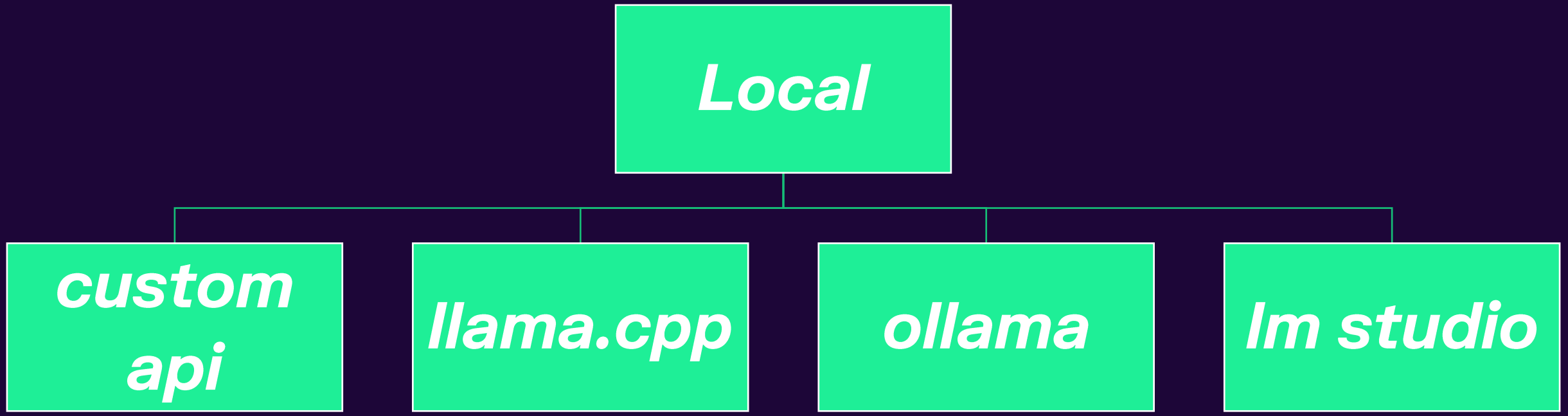
Method

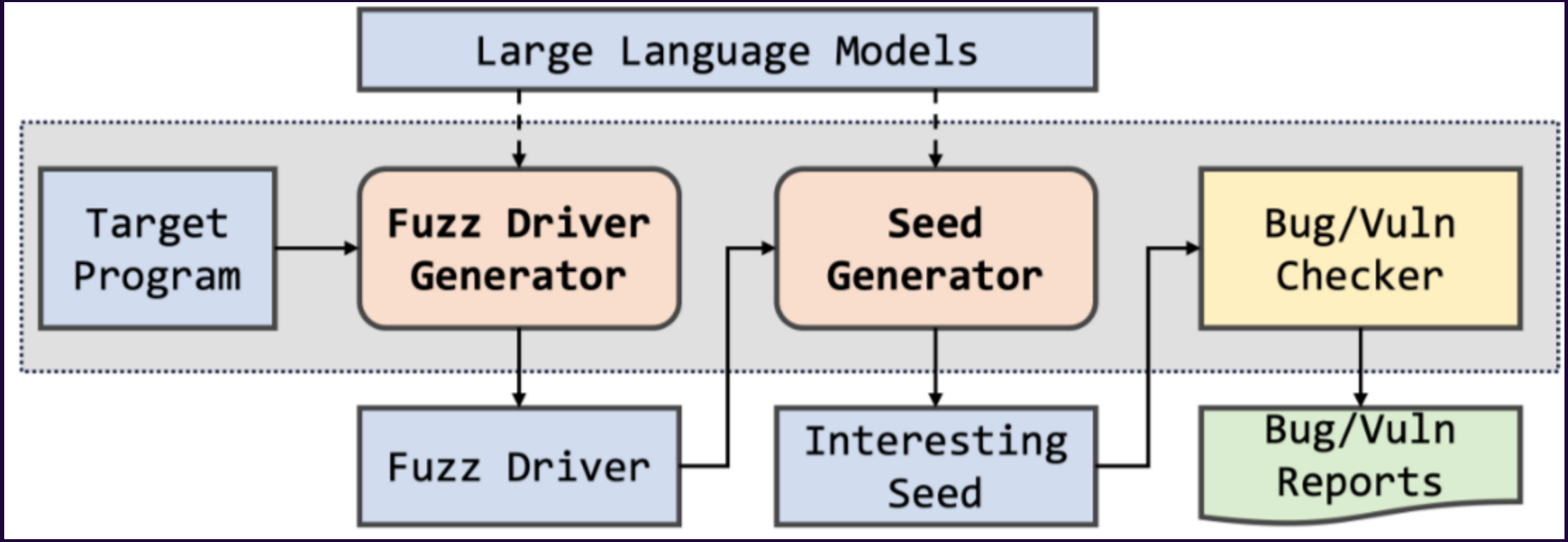


Method

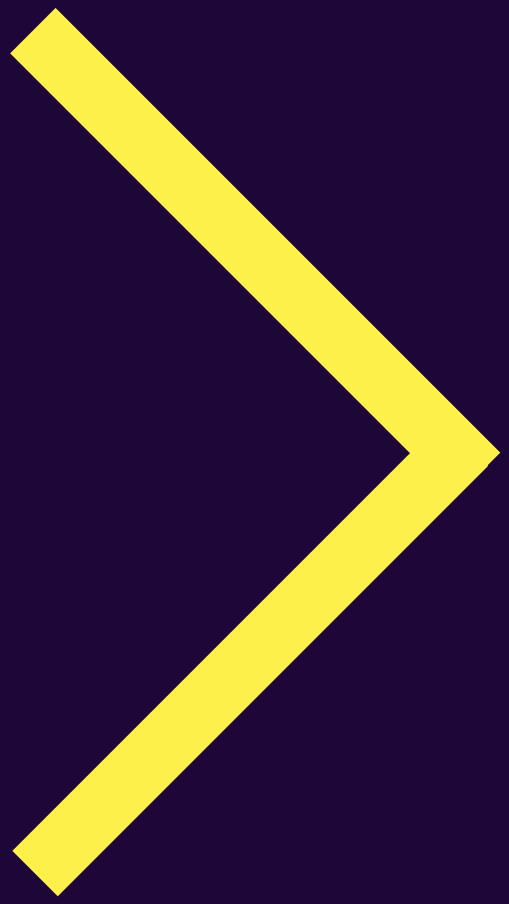


Method



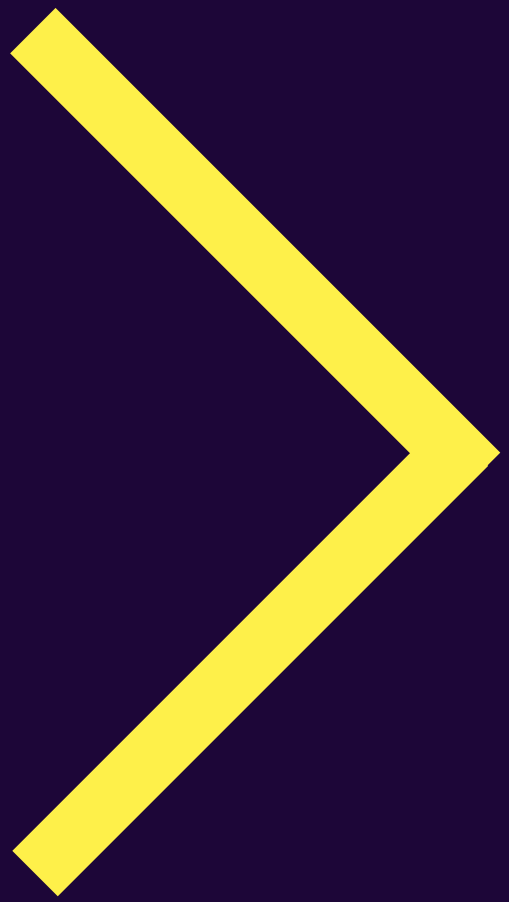

Large Language Models
Target Program
Fuzz Driver Generator
Seed Generator
Bug/Vuln Checker
Fuzz Driver
Interesting Seed
Bug/Vuln Reports

Result



Result

Literature Review



Literature Review

1. When Fuzzing Meets LLMs Challenges and Opportunities
2. An Empirical Study of OSS-Fuzz Bugs
3. Towards LLMs Guided Kernel Direct Fuzzing
4. LLMs for fuzzing parsers

Literature Review

When Fuzzing Meets LLMs Challenges and Opportunities

Results for database fuzzing:

- Driver correctness improved by up to 94% (Wingfuzz vs. direct LLM)
- Branch coverage increased by up to 56%
- Semantic correctness of SQL inputs up by 159%
- Bug detection: false positives reduced from 99% to 36%, with real bug detection rate rising from 0.5% to 64%

Yu Jiang et al., FSE 2024

An Empirical Study of OSS-Fuzz Bugs

- *Analyzed 23,907 bugs across 316 open-source projects over 4 years.*
- *Majority (52%) of bugs impact availability (e.g., crashes, timeouts).*
- *13% of all bugs are flaky; 86% of flaky bugs remain unfixed.*
- *22% of all bugs go unfixed; especially timeouts, OOM, and assertion violations.*

Zhen Yu Ding & Claire Le Goues, 2021

Towards LLMs Guided Kernel Direct Fuzzing

Proposes SyzAgent: Integrates LLMs with Syzkaller for real-time kernel fuzzing guidance.

Results:

- ***On 27 Linux kernel functions, SyzAgent outperformed Syzkaller in 67% of cases (≥10% higher hit rate).***
- ***Broke Syzkaller's coverage plateau in several deep code paths.***

Xie Li et al., 2025

LLMs for fuzzing parsers

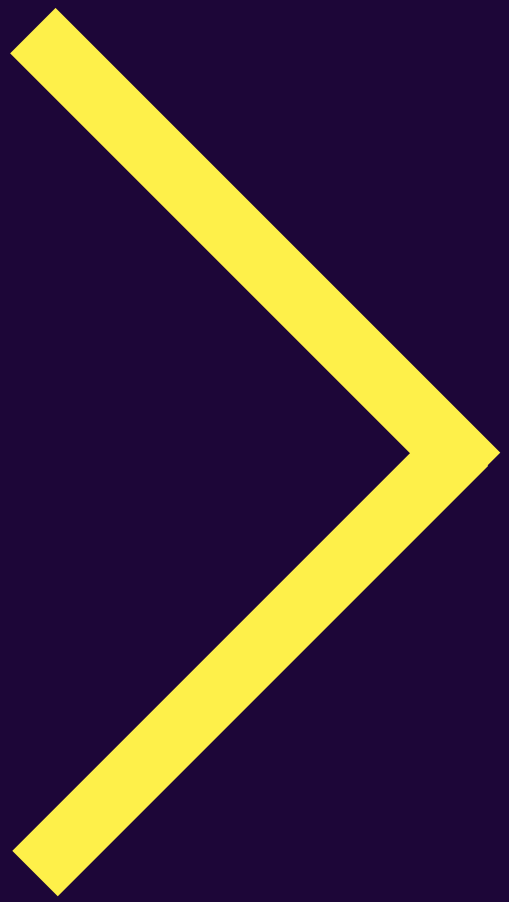
Proposes using LLMs to generate fuzzing seeds by interpreting natural language format specifications.

Results:

- ***Outperforms basic mutation and random fuzzers in code and branch coverage.***
- ***LLM-generated seeds work on both real and novel, handwritten formats (accuracy: 61–100%).***

Ackerman & Cybenko, 2023

Next Steps



Next steps

Evaluation

Comparison of different local llms

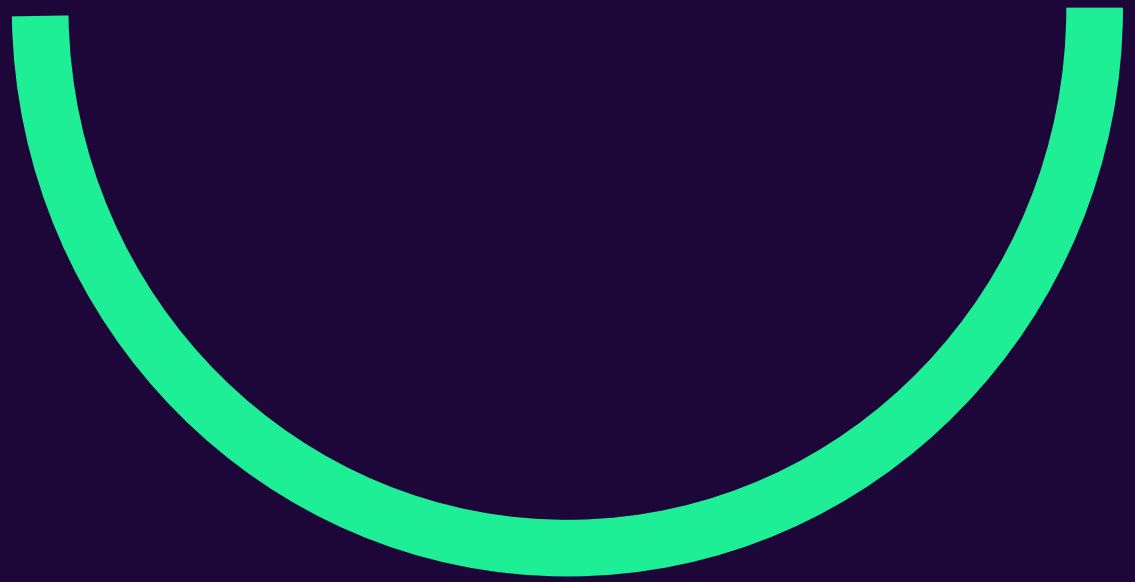
Comparison between local and cloud llms

Evaluation Framework

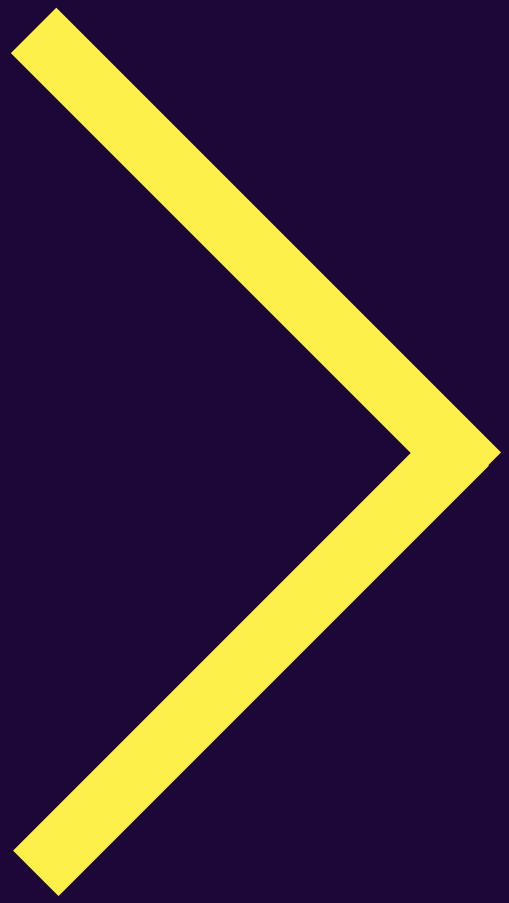
Metrics:

- Number of bugs found
- Code coverage percentage
- Time to generate fuzz tests
- Quality of generated code (measured via static analysis or manual review)
- Number of tokens

Any Questions



Thank you!



We transform automotive mobility

CARIAD
A VOLKSWAGEN GROUP COMPANY