2023 DESIGN AND VERIFICATION™ DVCON CONFERENCE AND EXHIBITION UNITED STATES

SAN JOSE, CA, USA
FEBRUARY 27-MARCH 2, 2023

Survey of Machine Learning (ML) Applications in Functional Verification (FV)

Dan Yu, Harry Foster, Tom Fitzpatrick

SIEMENS

accellera SYSTEMS INITIATIVE

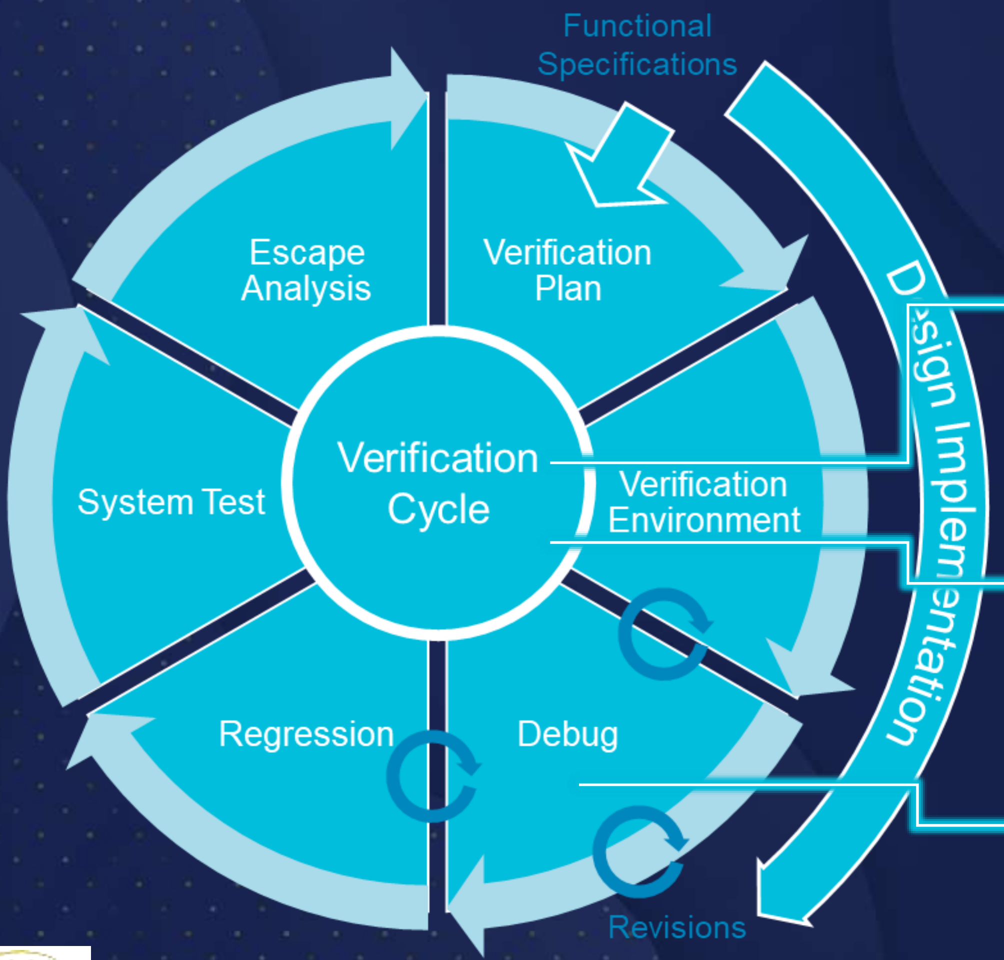
Motivations

Need comprehensive overview on ML for FV

Insufficient attention on data problem

Unaddressed challenges in production environment

Abundance of Data in FV



Example of Data Generated in Verification

Task	Data	Modality
Simulation	Waveforms	Graph
	Timing	Text, Graph
	Test vectors	Text, Table
	Test results	Text, Table
	Logs	Text
	Metrics	Text, Table
	Coverage	Table
Formal	Design specs	Text
	Design	Code
	Proof results	Table
	Traces	Text, Table
	Metrics	Table
	Constraints	Code
	Assertions	Code
	Proof scripts	Code
Debug	Trace logs	Text, Table
	State dumps	Table
	Breakpoint	Table
	Testbench	Code
	Commits	Code, Table

Verification is a Natural Playground of Machine Learning

Problem Formulation

- Well defined inputs & outputs

Baseline Solution

- Rule-based or manual

Data Availability

- from design & verification workflow

How can ML be leveraged to extract value out of data?

Application ML Has Been Studied in Every Corner of Verification

- Static code analysis
- Requirement engineering
- Formal verification
- Simulation acceleration
- Bug detection
- Coverage closure
- …

Requirement Engineering: Spec Translation

Check if ack arrives 3 clocks after a request

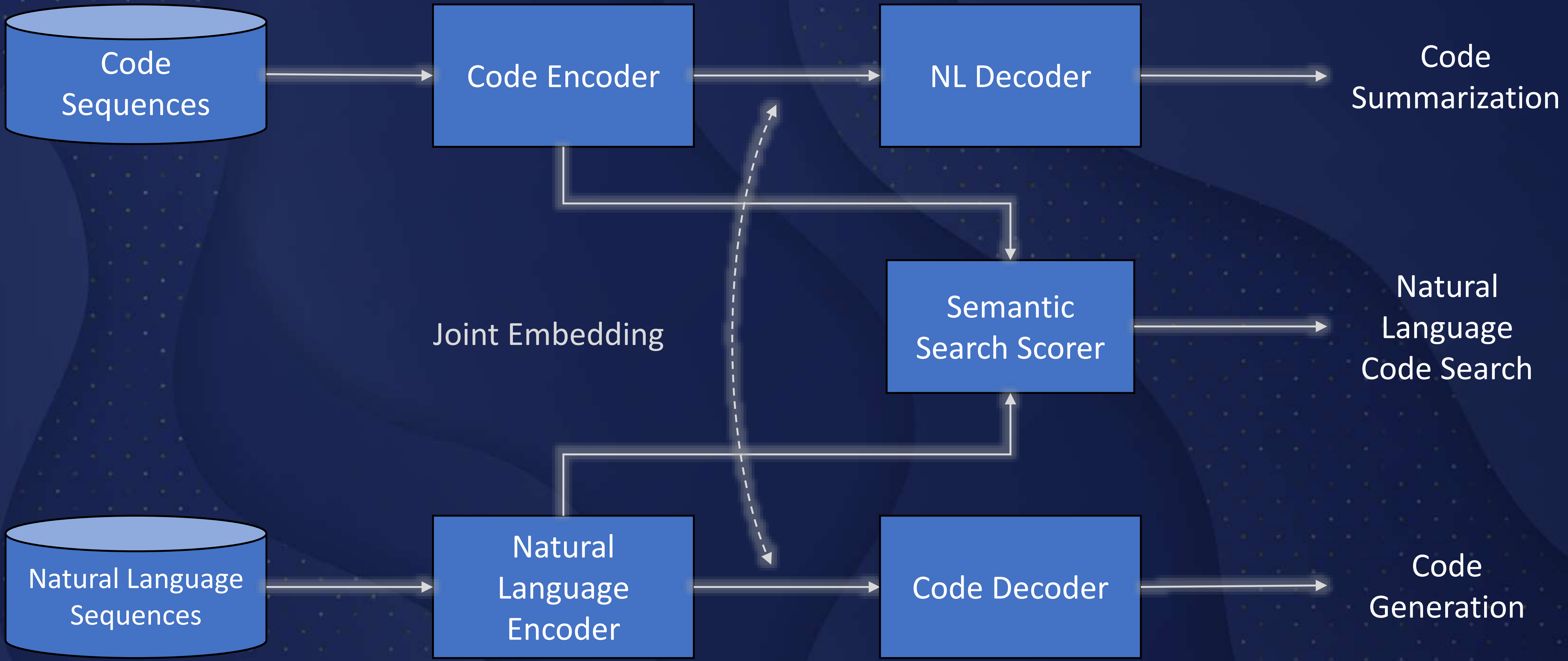
Translation Engine

```
assert property (@(posedge clk) req |-> ##3 ack);
```

Static Analysis: Smell Detection & Quality Assessment

Code smell: malformed code indicating bigger problem, e.g. a big module without submodules

Static Analysis: Code Summarization, Generation and Semantic Search

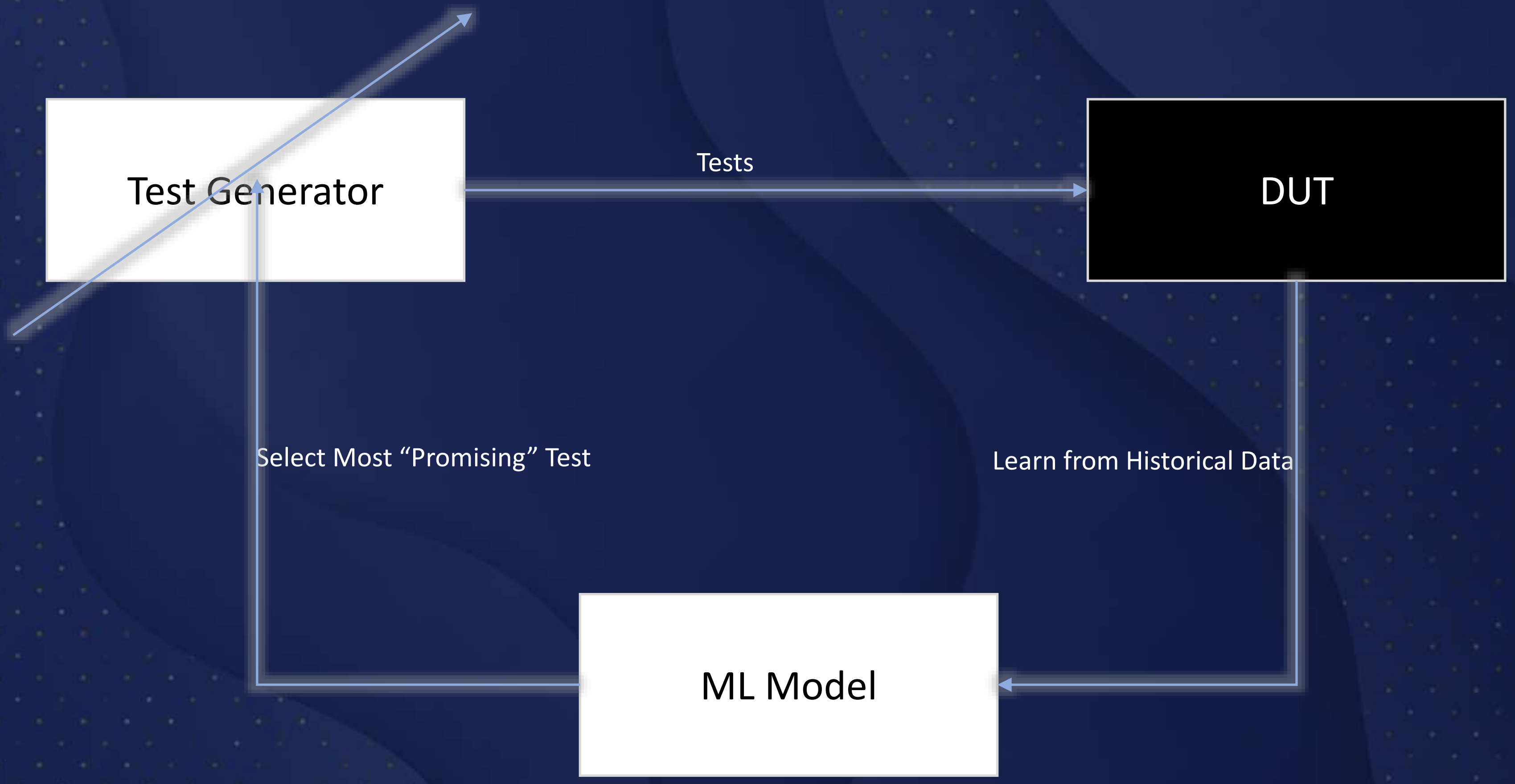


Simulation Acceleration: Approximation with ML models

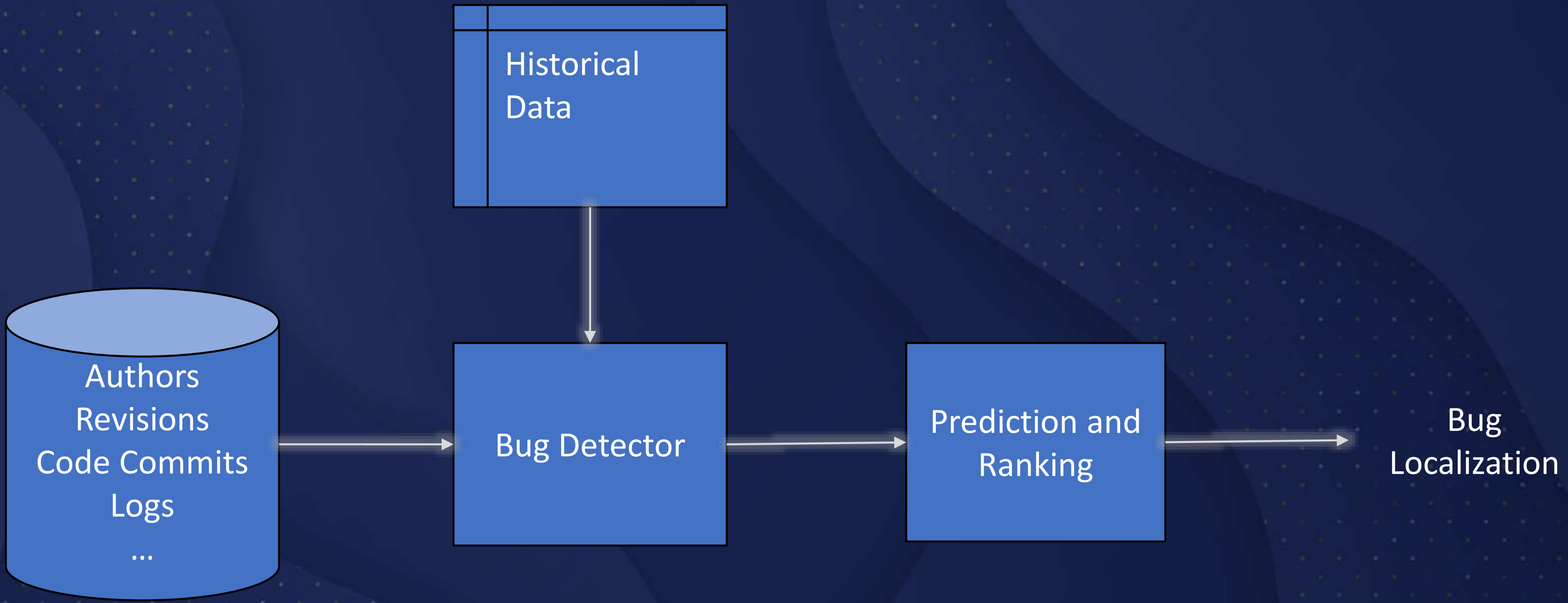
universal approximation theorems

IC Design Simulation ≈ ML Model (Large-scale Highly-parallel Model) + AI Accelerator (Highly-parallel computation)

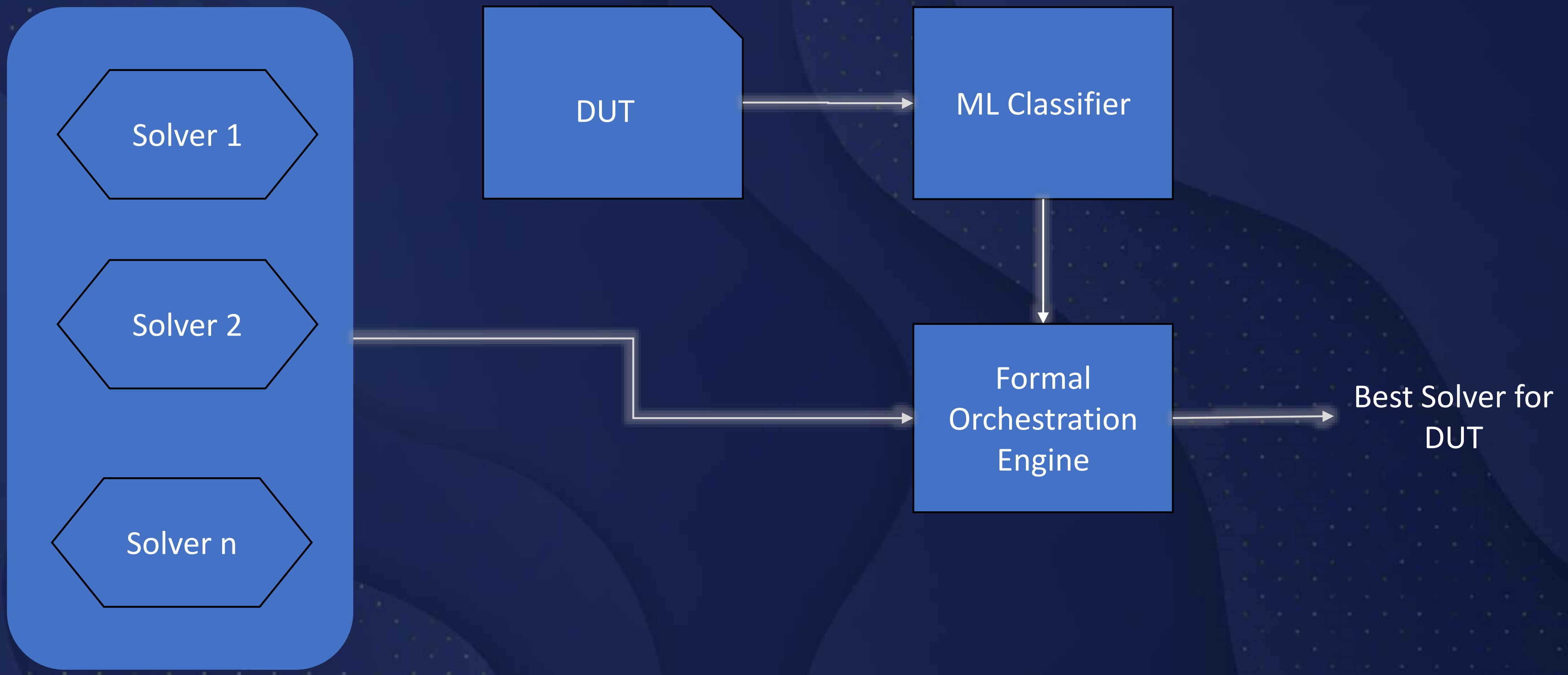
Coverage Closure: Test Generation



Bug Detection and Localization



Formal Orchestration



Most Research Restricted by Data

1

Data Quality: More efforts spent on preparing data for research work

- Most open-source data are not ready to experiment
- Lack of complete lifecycle verification data

2

Repeatability: Results are not easily verifiable or directly comparable without the same data

- Many proprietary / undisclosed design data used to test performance

3

Data Volume: Scale of datasets too small to train advanced ML models

- Individual efforts limits the useable dataset size for the experiments

Scarcity of Quality Data

Open-source Data has Great Potential, Proceed with Caution

GitHub

Verilog	19,861
VHDL	16,352
SystemVerilog	3,212

RISC-V

Verilog	59
VHDL	11
SystemVerilog	25
Other	16

OpenCores
www.opencores.org

Verilog	470
VHDL	467
Other	30+

OSS Projects Hosted 10+ mln LoC

Data Preparation

Gather / Assess / Cleans / Labeling / Transform / Enrich / Validate

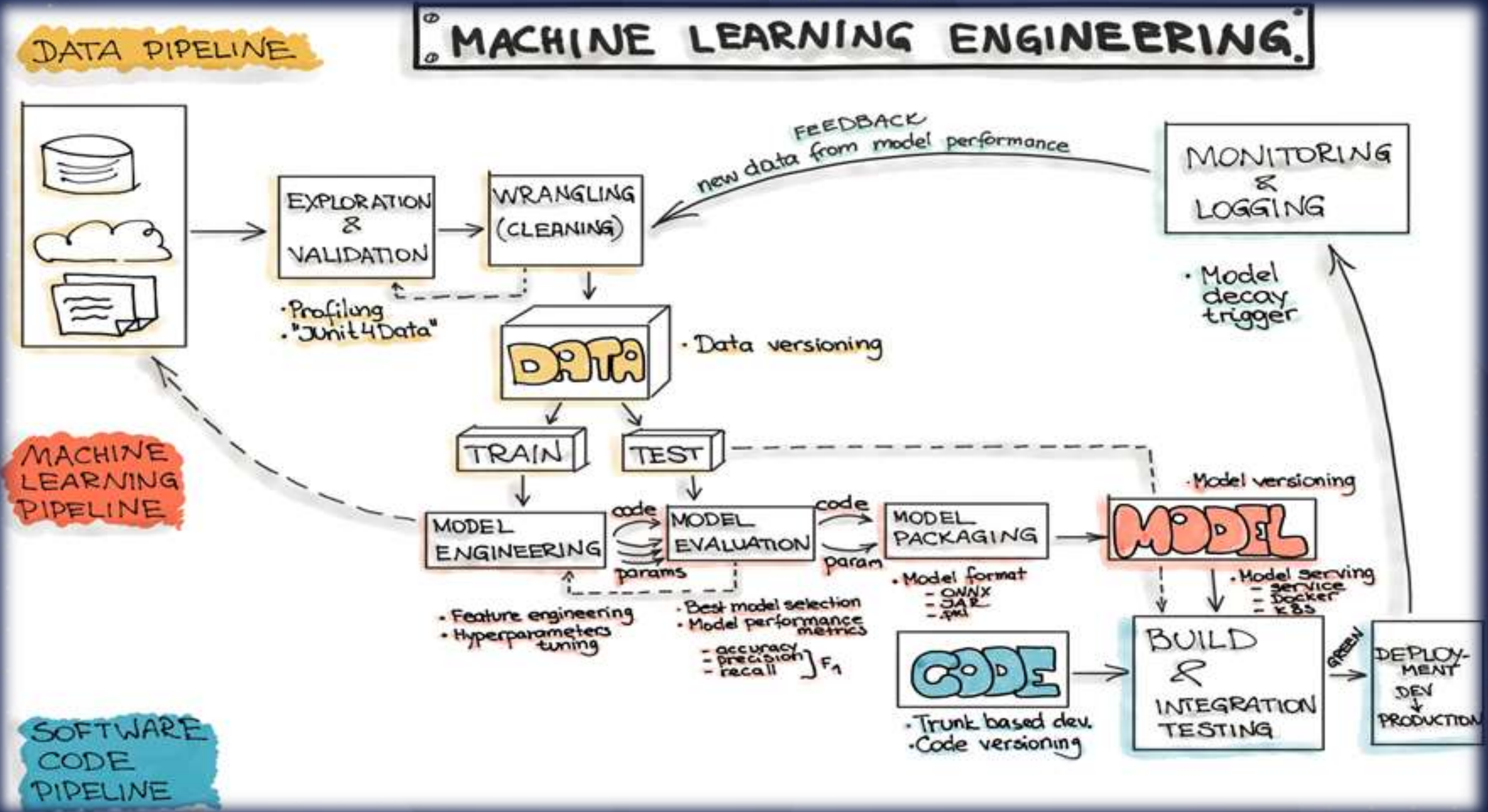
ML Useable Verification Data

ML Model’s Generalizability

How does a model perform on new data never seen before?

	Design 1	Design 2	Design 3	New Design 4	
Model A	95%	97%	95%	29%	“Rote Learning”
Model B	95%	93%	97%	94%	Generalizable Model

Scalability: Data Centric MLOps

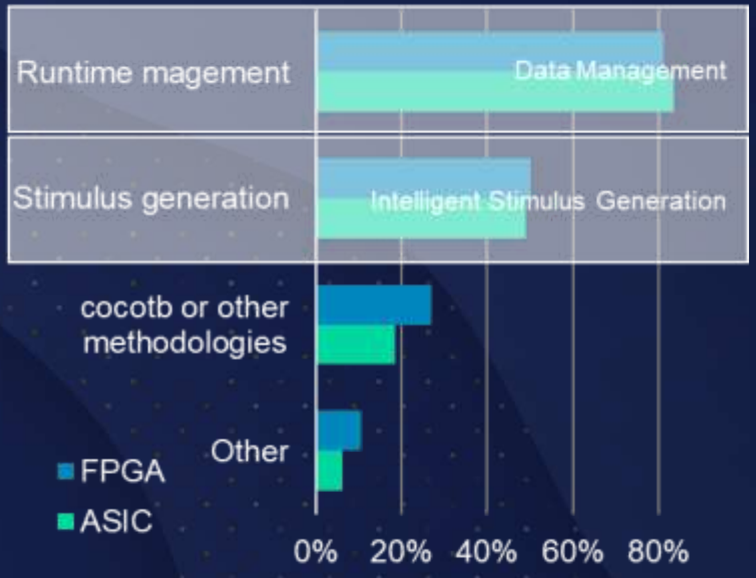


Source: End-to-end Machine Learning Workflow (ml-ops.org)

ML for Verification: What's the Next?

Big Data &
Large (Language) Model

Complex Relationships & Graph
Learning



Python &
Data Analytics w/ ML

Contact

Dan Yu, Harry Foster, Tom Fitzpatrick

Digital Verification Technology
Siemens EDA
{dan.yu, harry.foster, tom.fitzpatrick}@siemens.com