



UNIVERSITY OF TARTU
Institute of Computer Science



SEA Group: Software Engineering Analytics

University of Tartu

ICS Day 2025

Dietmar Pfahl, Hina Anwar, Faiz Ali Shah, Kristiina Rahkema, +PhD/MSc students

dietmar.pfahl@ut.ee

<https://sea.cs.ut.ee>



Software Engineering Analytics – The Team

Dietmar Pfahl
(Professor)



Hina Anwar
(Assistant Professor)



Faiz Ali Shah
(Assistant Professor)



Kristiina Rahkema
(Assistant Professor)



Muhammad B. Akhtar
(Junior Researcher, PhD Student)



Fauzia Khan
(Junior Lecturer, PhD Student)



Maj-Annika Tammisto
(Junior Researcher, PhD Student)



Ali Ihsan Güllü
(Junior Researcher, PhD Student)



Sander Sats
(Junior Researcher, PhD Student)



Maria Khan
(Junior Lecturer, PhD Student at TLU)



Past members of SEA:

- Ezequiel Scott (Assistant Professor)
- Huishi Yin (former PhD Student)
- Riivo Kikas (former PhD Student)
- Alejandra Duque-Torres (former PhD Student)
- Myroslava Stavnycha (former MSc Student)
- Laima Anna Dalbina (former MSc Student)
- Mariana Falco (Visiting Researcher)
- Tarcio Bezerra (Visiting Researcher)



Software Engineering Analytics – The Mission

“To collect and analyze software engineering data with the goal to estimate, predict, and improve the quality of software and software-intensive systems”

Notes:

- Software Engineering = A systematic approach to develop software with the expected quality, within a given time and effort budget
- Software quality has many dimensions/facets, e.g., functionality, dependability, maintainability, performance, energy-efficiency, etc.
- The terms “estimate, predict, and improve” imply the development of models and tools (prototypes)

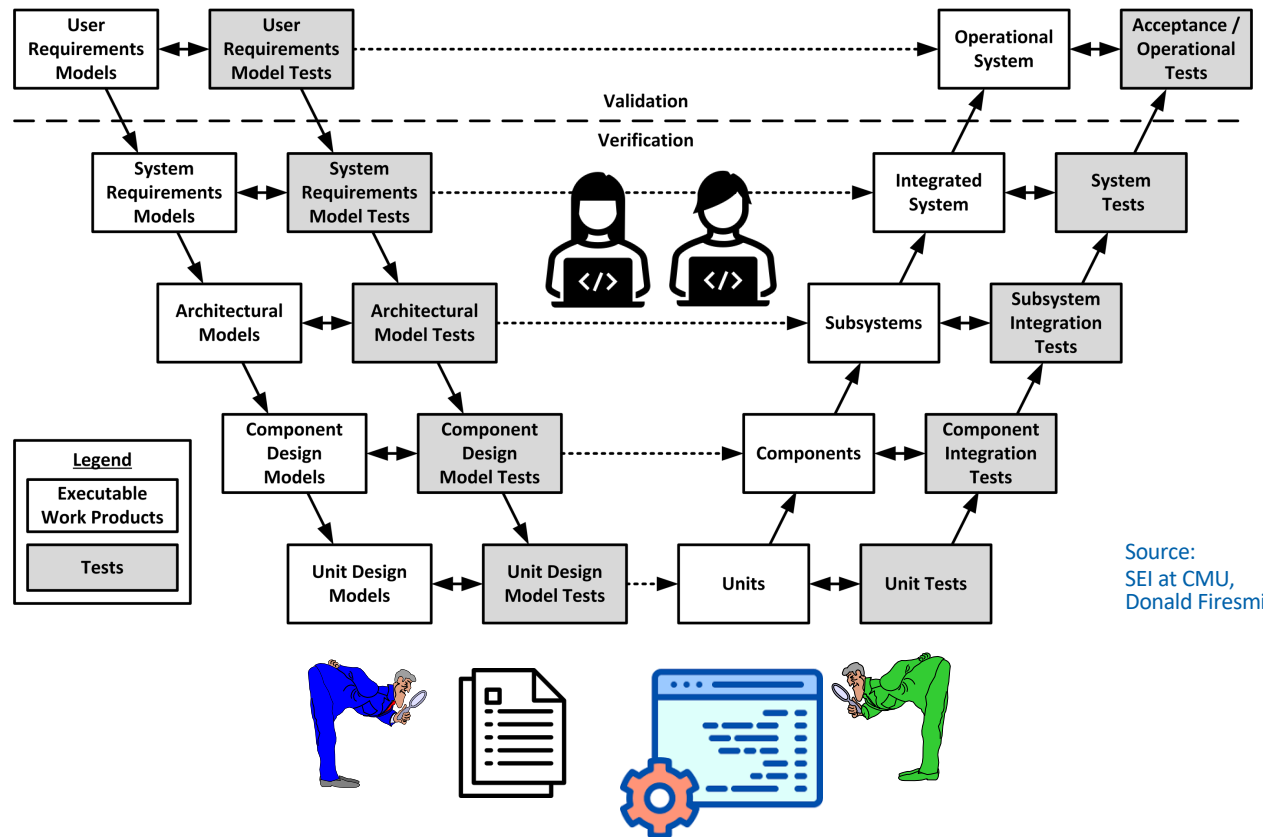
Software Engineering Analytics – The Scope

What we do:

- Methods and tools for static document and code analysis
- Methods and tools for dynamic code analysis (incl. testing)
- Methods and tools for quality estimation and prediction
- Methods and tools for selecting, augmenting, and generating test data, test cases, test oracles, and test code

Note:

- All of this may or may not involve AI



SEA Project Example 1: Mobile App Feature Analysis

Research Topic:

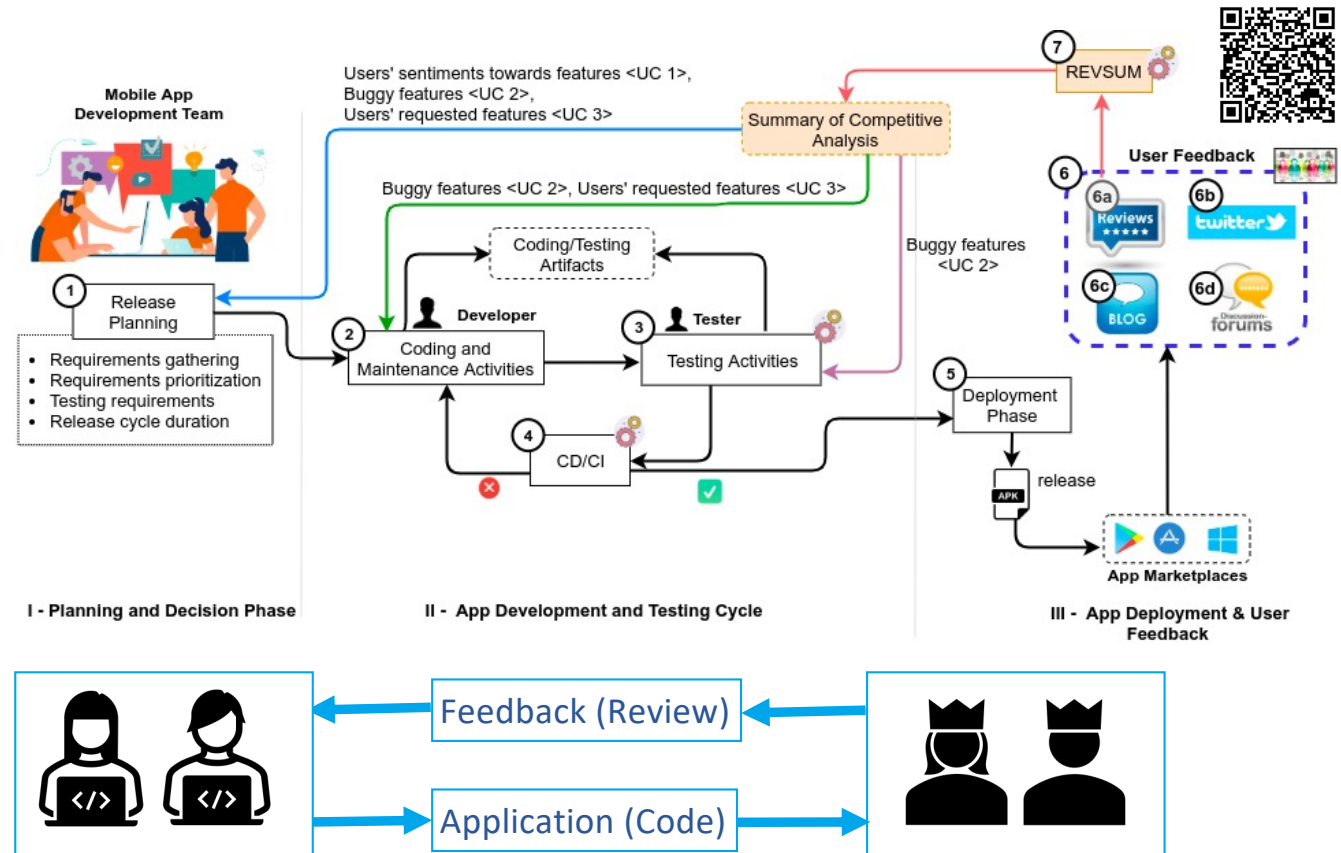
- Automated extraction, selection, and classification of end user reviews (mobile apps) → what features are good/bad/missing?

Why?

- To support app feature maintenance and evolution decision-making

Who benefits?

- App developers and (indirectly) app end users



Contact: Faiz Ali Shah

SEA Project Example 2: Security Vulnerability Detection

Research Topic:

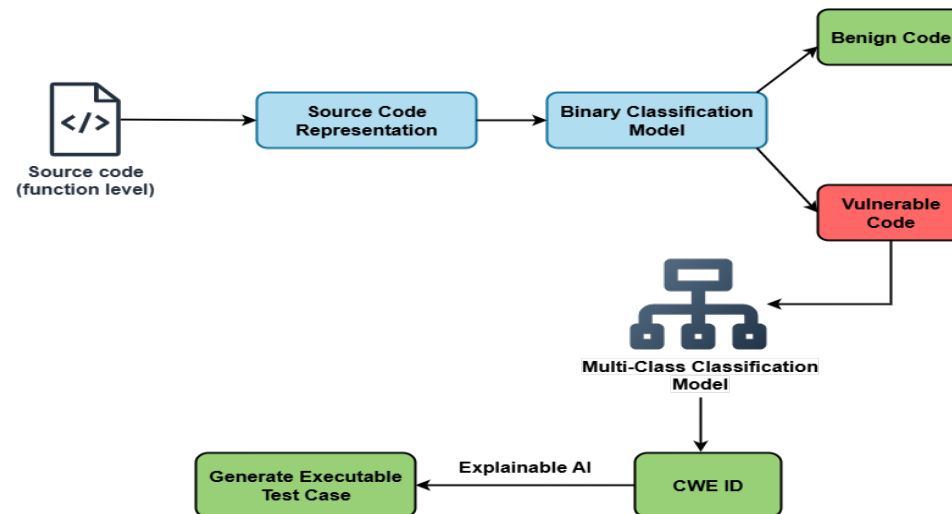
- Predict security vulnerabilities automatically in code
- Generate test cases as explanations for predictions

Why?

- Manual vulnerability detection is slow and error-prone
- Explanations build trust in predictions

Who benefits?

- Software developers, security analysts, and (indirectly) end users



A pipeline for AI-Driven detection and explanation of code vulnerabilities



Contact: Faiz Ali Shah

SEA Project Example 3: Mobile App Security Analysis

Research Topic:

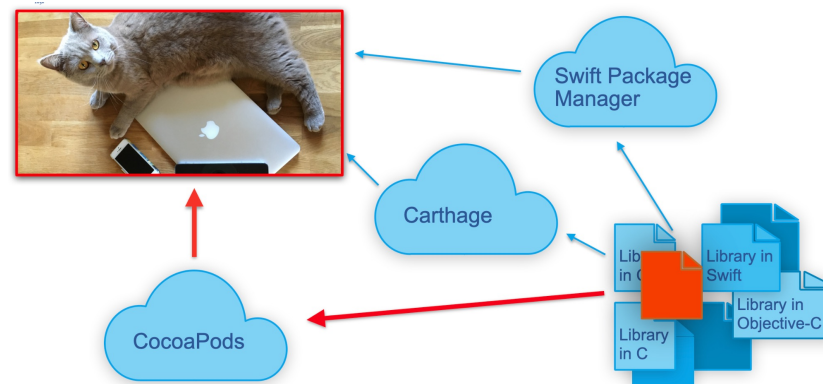
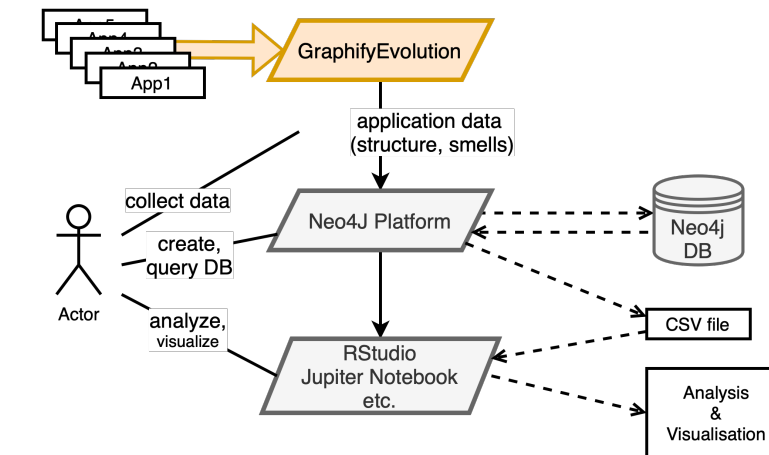
- Analysis of code smells in iOS mobile app code
- Analysis of dependencies to libraries with known security vulnerabilities

Why?

- To reduce the risk of high maintenance cost and dependency on insecure code

Who benefits?

- Developers of tools that support software developers, software developers, and (indirectly) app end users



```
81 func analyzeAppVersion(appVersion: AppVersi
82 var finalClasses: [Class] = []
83 var methodsToBeHandled: [Method] = []
84 var variablesToBeHandled: [Variable] =
85
86 var pathWithoutGit = appVersion.directo
87 if pathWithoutGit.contains(".git") {
88     pathWithoutGit = "\(appVersion.dire
89 }
90
91 self.syntaxAnalyser.reset(with: pathWit
92 self.fileManager.updateDependencies(pat
93
94 var includePaths: [String] = fileManag
95 //print("all include paths: \(includePa
96
97 var isMerge = false
98 if appVersion.parent != nil && appVersi
99     isMerge = true
100 }
101
102 var newClassVersions: [Class] = []
103
104 if !self.noSourceCodeAnalysis {
105     if let parent = appVersion.parent {
106         print("has parent")
107         if let parentApp = appVersion.p
108
109         let parentClasses = parentA
110         var altParentClasses: [Clas
111
112         var intersectionChanged = [
113         var intersectionParentNew =
114         var intersectionAltParentNe
115         var intersectionNew = paren
116
117         if let altParent = appVersi
118
119         if let altParentApp = a
120         altParentClasses =
121             for classInstance i
122         } else {
123
124         intersectionChanged = A
125         intersectionParentNew =
126         intersectionAltParentNe
127         intersectionNew = Array
128
129         var addedClasses: [String:C
130
131         combinedPaths.append(conten
132
133
134
135
136
137
138
139
140
```

Contact: Kristiina Rahkema

SEA Project Example 4: Intelligent Test Case Selection

Research Topic:

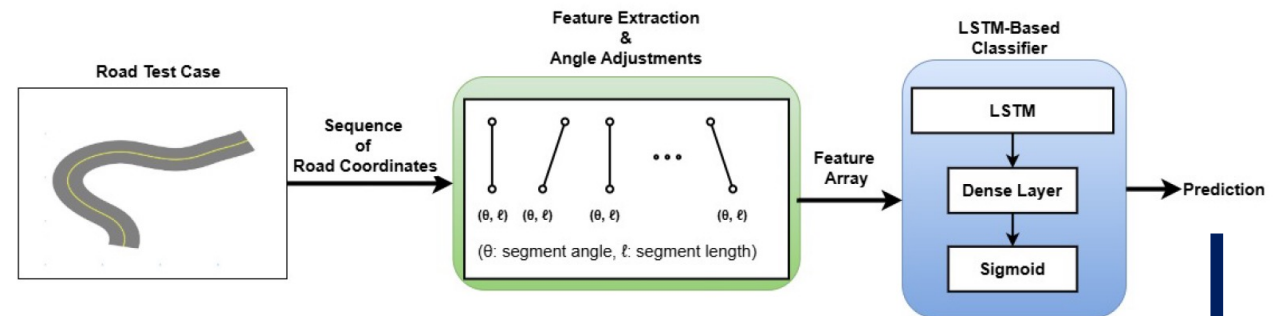
- Intelligent selection of test roads for regression testing of Self-Driving Cars (SDC) in a simulator

Why?

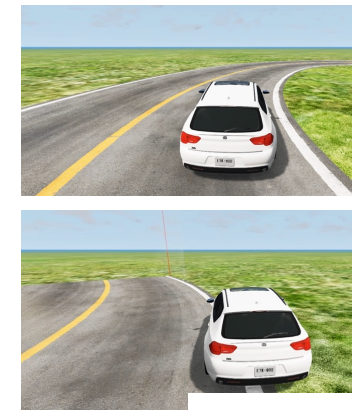
- To make SDC test suites more effective and the test process more efficient

Who benefits?

- Testers and (indirectly) developers of SDC software stacks



LSTM-based test data processing pipeline



will
Pass
or
Fail

Contact: Dietmar Pfahl

SEA Project Example 5: Simulation-Based Safety Test

Research Topic:

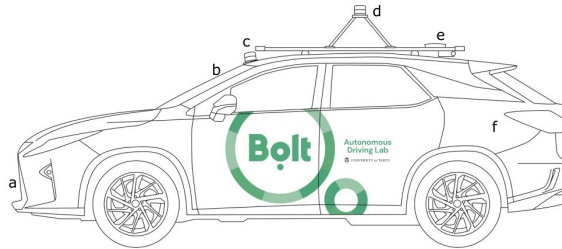
- Methods for simulation-based safety testing of Automated Driving Systems (ADS)

Why?

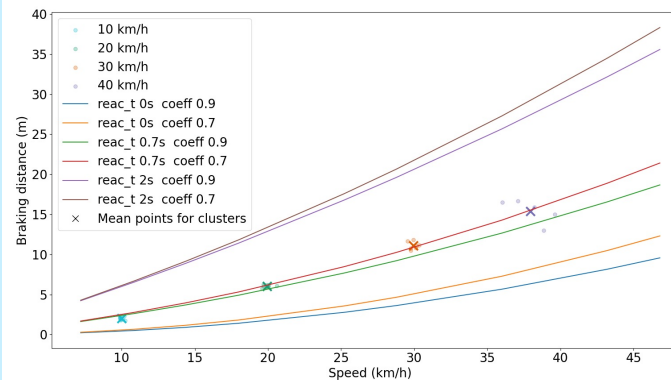
- To help improve the safety of ADS
- To complement on-road testing

Who benefits?

- Testers and developers of ADS and (indirectly) users of ADS and traffic participants where ADS operate



Comparing the emergency braking of simulated ADS versus theoretical reference modes:



SEA Project Example 6: Synthetic Test Data for X-tee

Contacts: Dietmar Pfahl, Faiz Ali Shah

Research Topic:

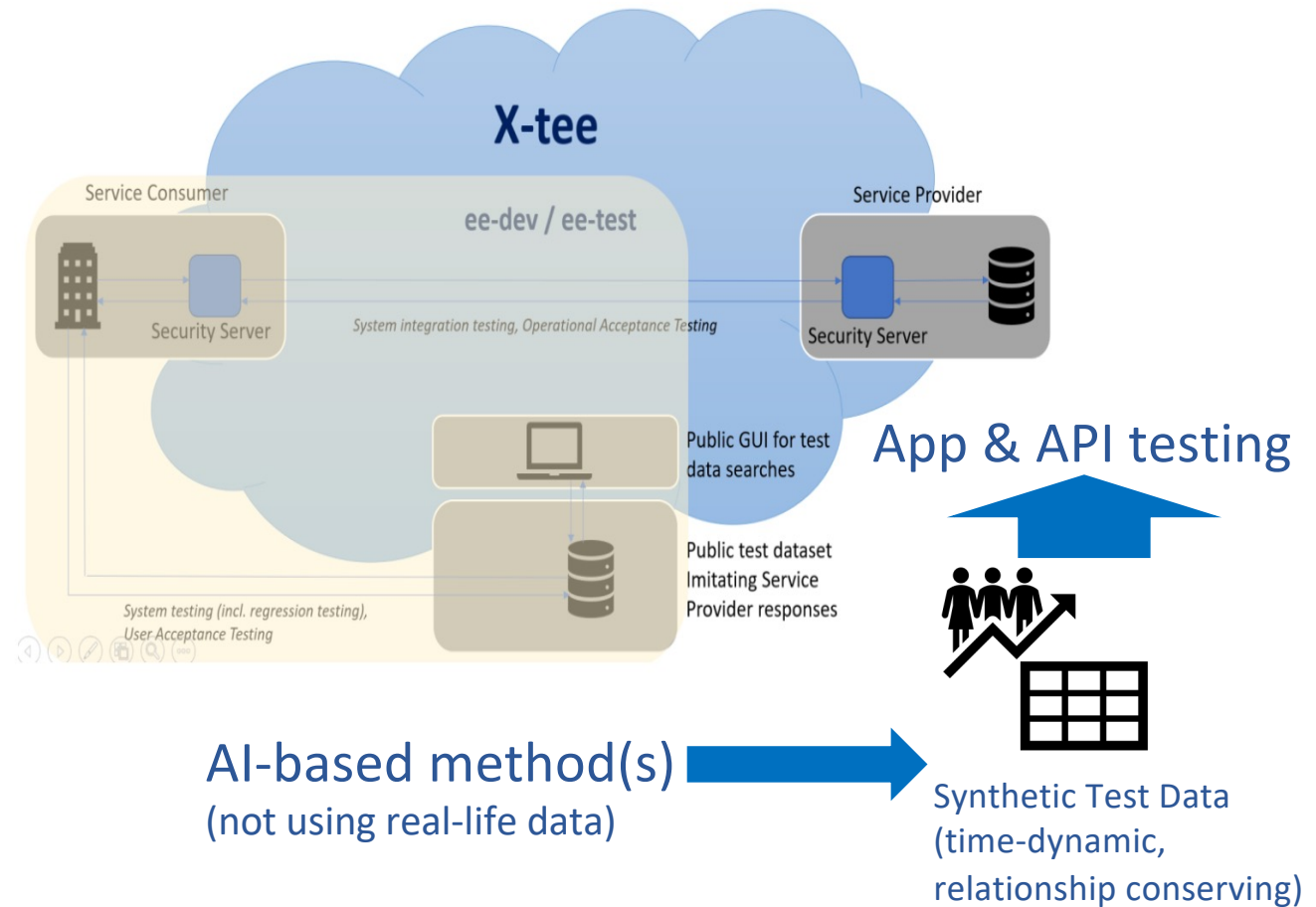
- Methods for automatic generation of relationship-conserving time-dynamic test data sets for the testing of X-tee services

Why?

- To increase test efficiency and effectiveness of X-tee services

Who benefits?

- Testers and developers of X-tee services and (indirectly) end users



SEA Project Example 7: Energy Efficiency in Background Processes

Research Topic:

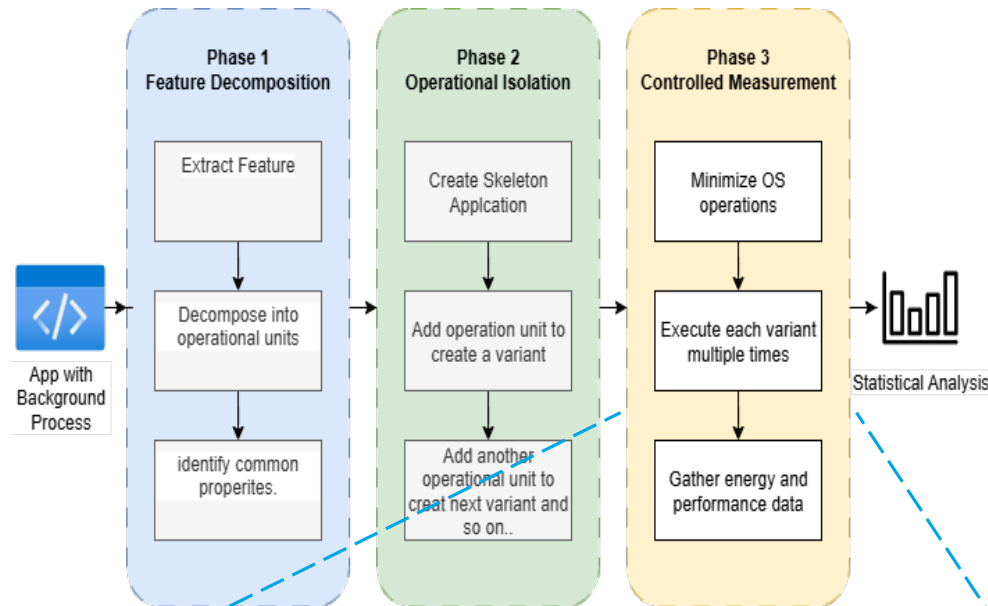
- Evaluate energy use of background features in desktop apps through a reusable, feature-level measurement process.

Why?

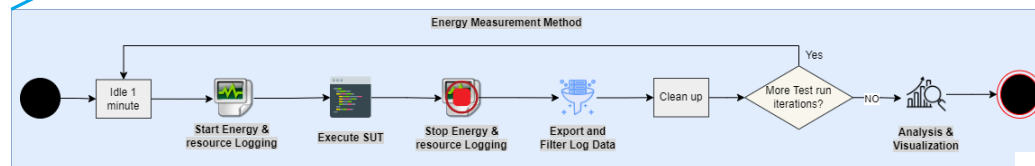
- To reduce overlooked energy waste in automated software features (such as AutoSave) and guide greener design decisions.

Who benefits?

- Software developers and researchers optimizing desktop application efficiency.



Energy measurement process for Background Processes



Measured Metrics	Unit
Duration	s
Mean power draw	W
Energy	Wh or J
CPU usage	%
RAM usage	MB or %

Contact: Hina Anwar

SEA Project Example 8: Energy Efficient AI Solutions

Research Topic:

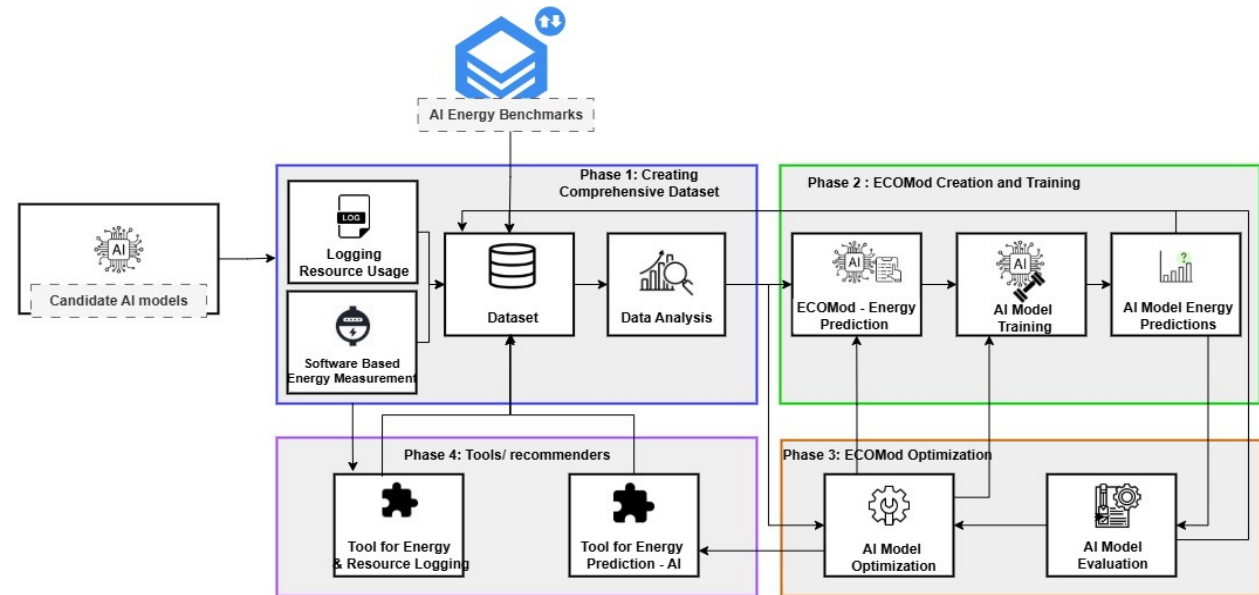
- Develop ECOMod: A model to predict and optimize energy consumption of AI models during early development.

Why?

- To enable responsible, energy-efficient AI solutions by giving developers early-stage tools to predict and optimize energy use.

Who benefits?

- AI developers, (Indirectly) policymakers focused on sustainable, responsible AI.



**Sustainable AI Isn't Optional
It's the Future!**

Contact: Hina Anwar

SEA Supervision & Teaching

PhD (since 2014):

- 6 PhD theses completed at UT
- 2 PhD theses completed at University of Calgary
- 5 PhD theses ongoing



2 PhD theses started in Sep 2024:

- AI-powered methods for detecting and explaining security vulnerabilities in code
- A method for systematically assessing the safety of automated driving systems via simulation

MSc & BSc (since 2014):

- 50+ MSc theses completed
- 30+ BSc theses completed



Courses and Curricula we are responsible for:

- International MSc Software Engineering (program director)
- 3 MSc-level courses (6 ECTS per term / 35 students per term)
- 3 BSc-level courses (7.5 ECTS per term / 200 students per term)

SEA Supervision & Teaching

PhD (since 2014):

- 5 PhD theses completed at UT
- 2 PhD theses completed at University of Calgary
- 5 PhD theses ongoing



2 PhD theses started in Sep 2024:

- AI-powered methods for detecting and explaining security vulnerabilities in code
- A method for systematically assessing the safety of automated driving systems via simulation

MSc & BSc (since 2014):

- 50+ MSc theses completed
- 30+ BSc theses completed



SEA Webpage



Thesis Topics

SEA Collaboration & Service

PROFES 2024
in Tartu ...



Collaboration:

- Within ICS/UT: ADL, Huber Flores' lab, Kairit Sirt's lab, Vesal Voidani's lab (in 4 other chairs)
- Within Estonia: joint MSc/BSc supervisions with companies / RIA
- Within Europe: Several collaborations with Universities and Research Institutes (Fraunhofer, SCCH), several EU project applications with many partners (academia, industry)
- Worldwide: Member of ISERN (International Software Engineering Research Network)

Service:

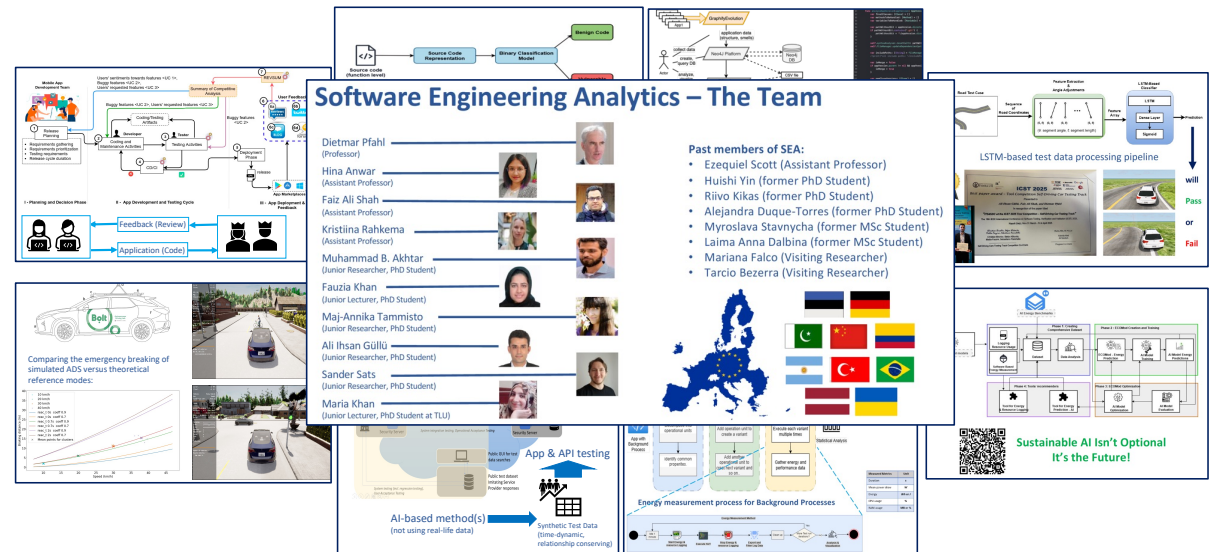
- Organizer of international conferences (ICSSP 2015 in Tallinn / ESEC/FSE 2019 in Tallinn / PROFES 2024 in Tartu)
- Member of several OCs, PCs, and steering committees worldwide
- Editorial board member of two top-level international journals (ESEM, INFOSOF)
- Volunteering in professional organisations (ACM, IEEE)
- Reviewer of grant application for several public funding agencies in Europe (Austria, Belgium, Finland, Ireland, Sweden, Switzerland, EU)

Wrapping it up

We are serving the Research Community, the Industry, the Society by

- Extending the body of knowledge
- Building models and tools
- Educating the future work force

“To collect and analyze software engineering data with the goal to estimate, predict, and improve the quality of software and software-intensive systems”





UNIVERSITY OF TARTU
Institute of Computer Science

**Talk to us if you wish to learn
more about our research!**



Hina



Faiz



Kristiina



Dietmar



unitartu



tartuuniversity

