fortiss

# Towards the Estimation of Quality Attributes on System Model Histories

**Konstantin Blaschke and Simon Barner**

fortiss

Research Institute of the Free State of Bavaria
Munich , Germany

**Konstantin Blaschke**

**Simon Barner**

Transformations-Hub Automotive
Software Engineering (TASTE)
grant no. 16THB0009C

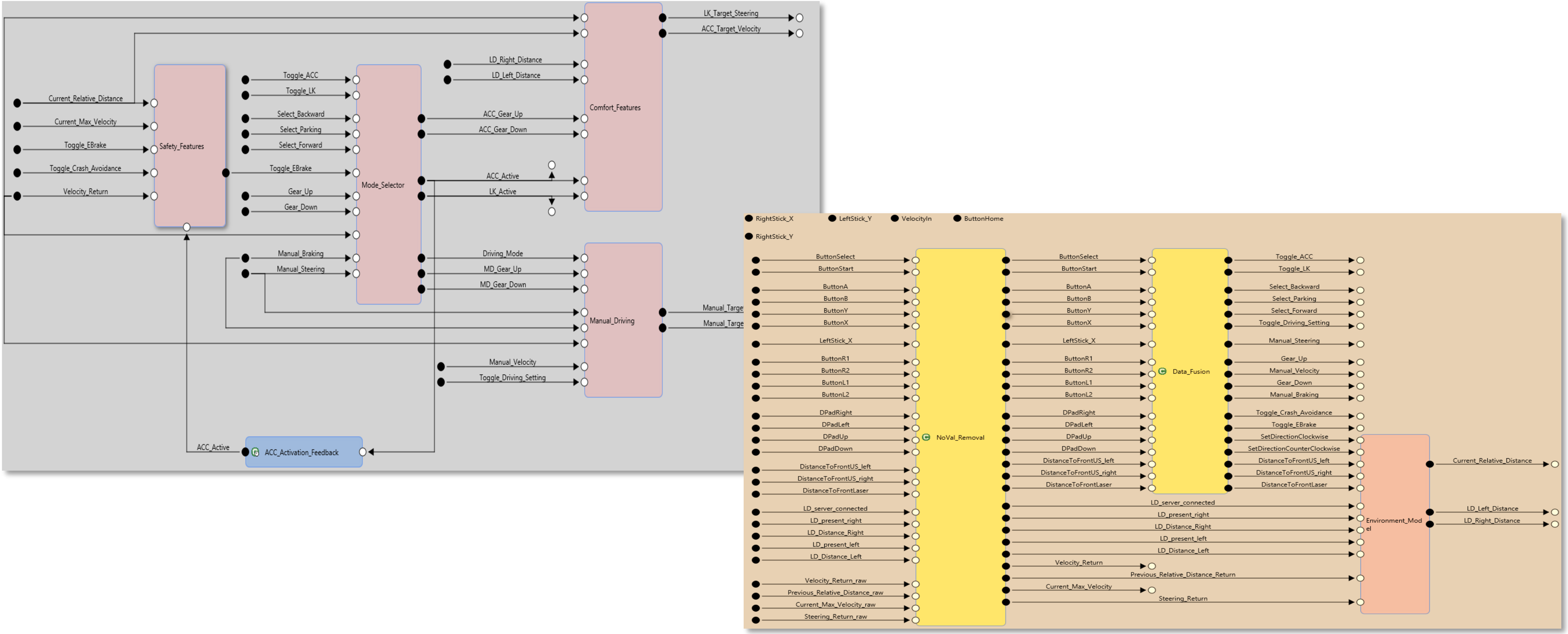
Gefördert durch:
Bundesministerium
für Wirtschaft
und Klimaschutz
aufgrund eines Beschlusses
des Deutschen Bundestages

Plattform für kontinuierliche Analyse
von Modellqualität im Systems
Engineering (KOMET)
grant no. 01IS24012B

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

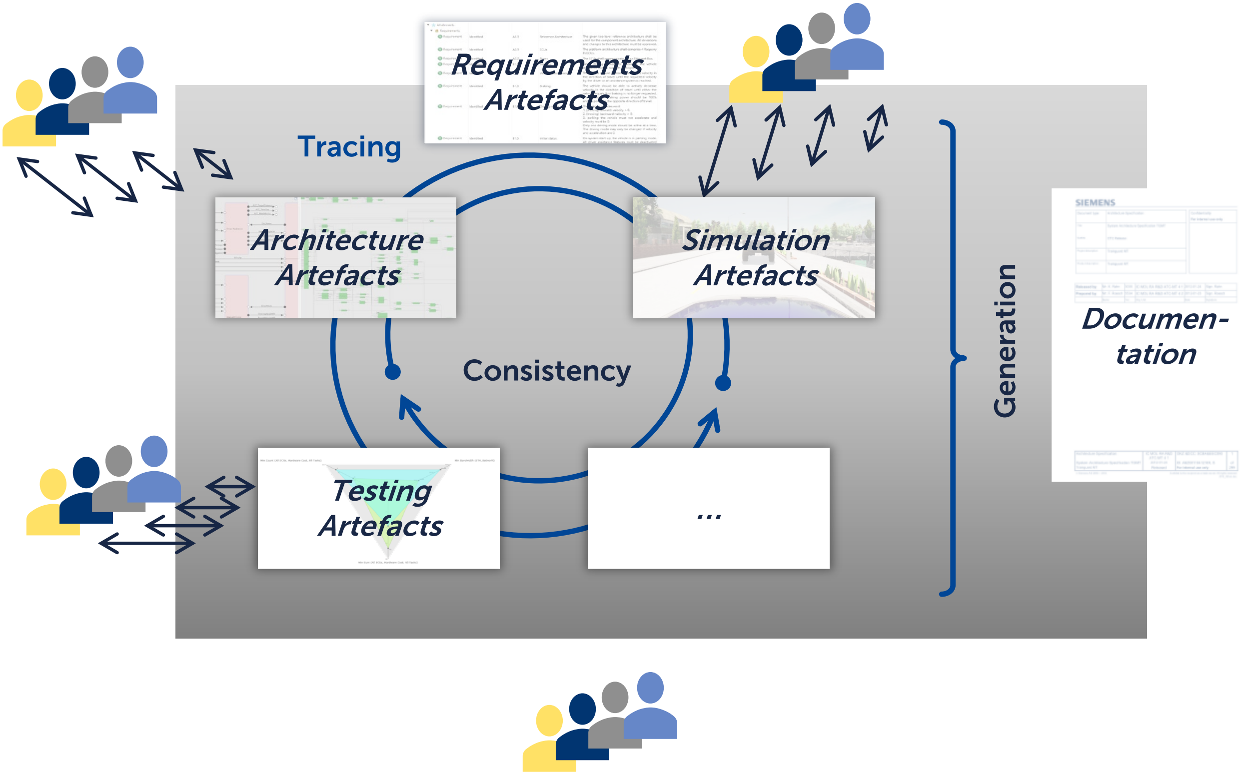
# Quality Assessment in System Model Artifacts

## Assessment of Model Quality in modeling artifacts



# Motivation for Model Quality Assessments

## The rising Complexity in Systems Engineering



*Requirements Artefacts*

Tracing

*Architecture Artefacts*

*Simulation Artefacts*

Consistency

*Testing Artefacts*

…

Generation

*Documen-tation*

SW / HW / Electrical / Management / Requirements

**Challenges:**

- **High collaboration** with diverging **knowledge** of modelers
- Rising **system complexity** and **cross-domain collaboration**
- **Domain** and **project**-**specific** modeling **conventions** and **standards**
- **Model quality** is not **sufficiently** monitored and **communicated** in engineering lifecycle

**Quality degradation over time and comprehensibility issues**

# Related Work & Gap

## State of the Art to Maintain and Improve Model Quality

**Metric-Based**

Extraction of metrics and calculation of quality attributes and complexity values.

[1, 2, 3, 6, 7]

**Rule-Based**

Initiation of rules and syntax checks for quality assurance.

[3, 4, 7, 8]

**Modeling Guidelines**

Textual guidelines document modeling conventions and best practices.

[INCOSE, OMG, 4, 5, 7]

**Gap: Missing cost-efficient model quality assessment approach that considers project-/domain- specific knowledge.**

---

1. Francesco Basciani, Juri Di Rocco, Davide Di Ruscio, Ludovico Iovino, and Alfonso Pierantonio. 2019. A tool-supported approach for assessing the quality of modeling artifacts. Journal of Computer Languages 51 (2019), 173–192. https://doi.org/10.1016/j.cola.2019.02.003
2. Marcela Genero, Mario G. Piattini, and Coral Calero. 2005. A Survey of Metrics for UML Class Diagrams. J. Object Technol. 4 (2005), 59–92
3. Fáber D. Giraldo, Sergio España, and Oscar Pastor. 2014. Analysing the concept of quality in model-driven engineering literature: A systematic review. In 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS). IEEE, Marrakech, Morocco. https://doi.org/10.1109/RCIS.2014.6861030
4. Myron Hecht and Jaron Chen. 2021. Verification and Validation of SysML Models. INCOSE International Symposium 31, 1 (2021), 599–613. https://doi.org/10.1002/j. 2334-5837.2021.00857.x
5. John Krogstie. 2012. Quality of Models. Springer London, London, 205–247. https://doi.org/10.1007/978-1-4471-2936-3_4
6. Christian F. J. Lange. 2007. Model Size Matters. In Models in Software Engineering, Thomas Kühne (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 211–216.
7. Parastoo Mohagheghi, Vegard Dehlen, and Tor Neple. 2009. Definitions and approaches to model quality in model-based software development – A review of literature. Information and Software Technology 51, 12 (2009), 1646–1669. https://doi.org/10.1016/j.infsof.2009.04.004 Quality of UML Models.
8. https://website.incquery.io/validator-for-enterprise-architect

# Problem Statement

## Challenges of assessing model quality

**Manual** model **reviews** are **time** and **cost-intensive.**

**Current automated approaches** are **rule-based** and cannot capture **higher-level properties** such as comprehensibility.

High **expertise** needed for **project-specific** model **quality** evaluations.

# Contribution

## Work in Progress on a new Approach for Quality Estimation

- **Method for creating automated data-driven model quality estimators based on expert ratings.**
- **Machine-learning pipeline for automated exploration of most accurate quality estimators.**

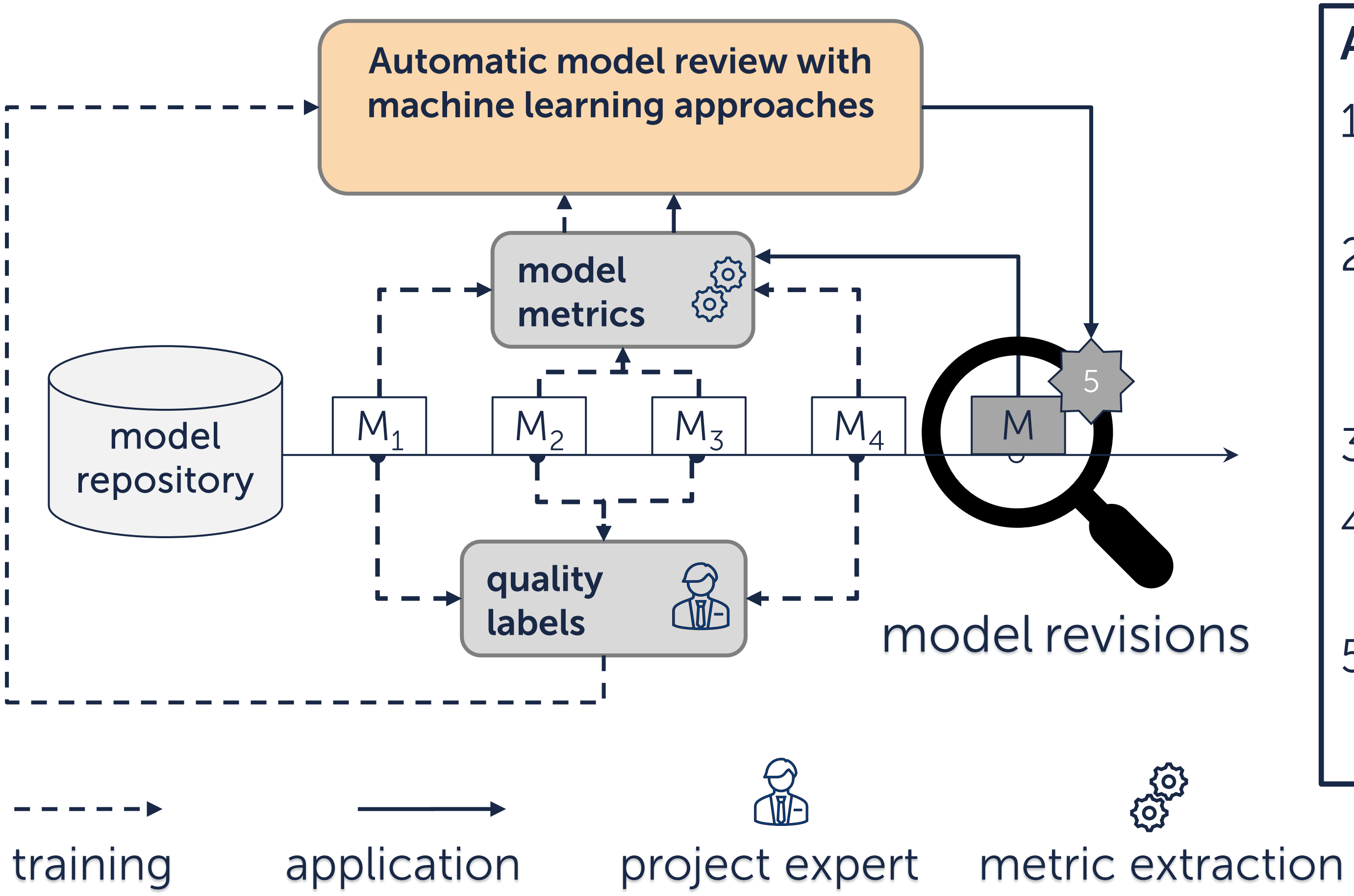
Reduction of **model review effort** by means of **automation**.

Going beyond rule-based approaches by including **project-/domain knowledge.**

**Reducing expert knowledge** required to **assess model quality.**

# Solution Approach

## Quality Estimator Concept

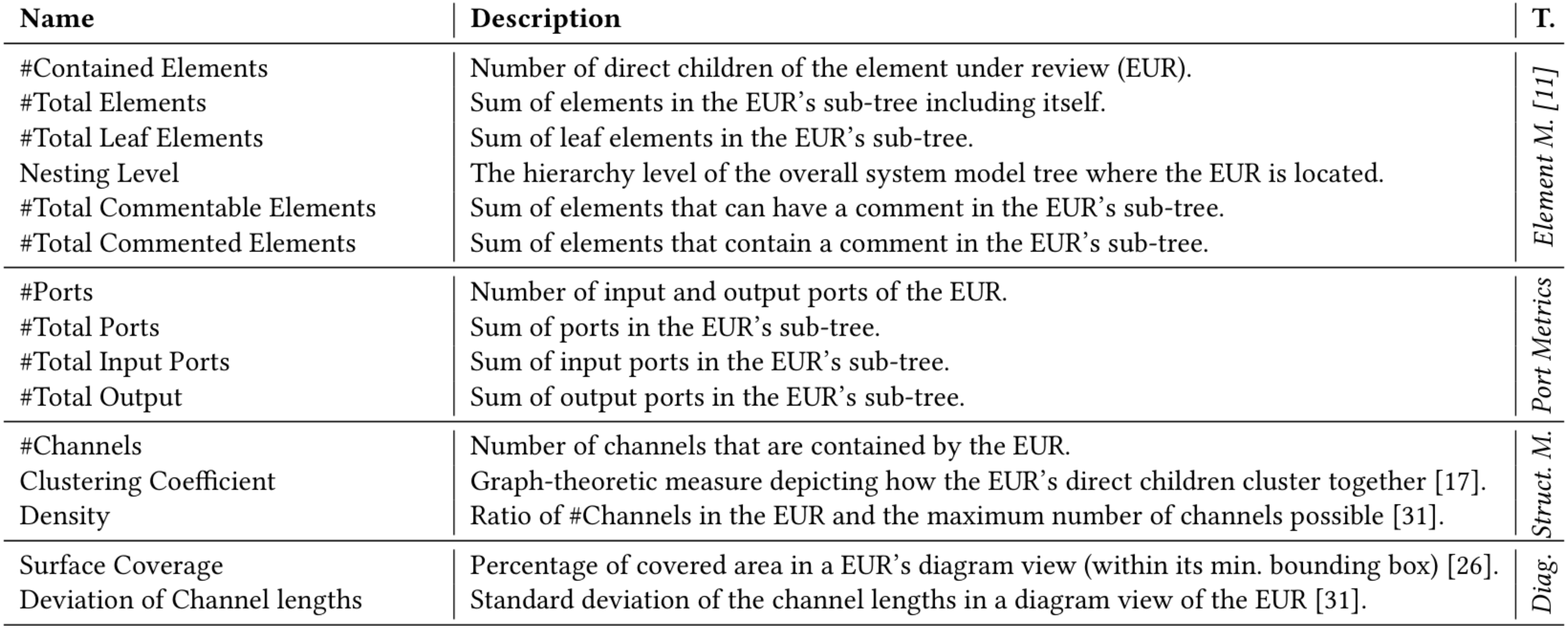


**Approach Steps:**

1. Setup **and extract model metrics** from model version history.
2. Let **experts** decide on a **company/project** specific **quality catalogue**.
3. Rate **model elements** manually.
4. Store **metric** and **quality attribute** history.
5. Train and **explore estimators semi-automated** and **apply** them.

# Approach: Metrics for Component-based Models

## Step 1: Extract **metrics** for each revision in model **version history**



| **Name** | **Description** | **T.** |
|---|---|---|
| #Contained Elements | Number of direct children of the element under review (EUR). | *Element M. [11]* |
| #Total Elements | Sum of elements in the EUR's sub-tree including itself. | |
| #Total Leaf Elements | Sum of leaf elements in the EUR's sub-tree. | |
| Nesting Level | The hierarchy level of the overall system model tree where the EUR is located. | |
| #Total Commentable Elements | Sum of elements that can have a comment in the EUR's sub-tree. | |
| #Total Commented Elements | Sum of elements that contain a comment in the EUR's sub-tree. | |
| #Ports | Number of input and output ports of the EUR. | *Port Metrics* |
| #Total Ports | Sum of ports in the EUR's sub-tree. | |
| #Total Input Ports | Sum of input ports in the EUR's sub-tree. | |
| #Total Output | Sum of output ports in the EUR's sub-tree. | |
| #Channels | Number of channels that are contained by the EUR. | *Struct. M.* |
| Clustering Coefficient | Graph-theoretic measure depicting how the EUR's direct children cluster together [17]. | |
| Density | Ratio of #Channels in the EUR and the maximum number of channels possible [31]. | |
| Surface Coverage | Percentage of covered area in a EUR's diagram view (within its min. bounding box) [26]. | *Diag.* |
| Deviation of Channel lengths | Standard deviation of the channel lengths in a diagram view of the EUR [31]. | |

# Approach: Quality Attributes for Component-based Models

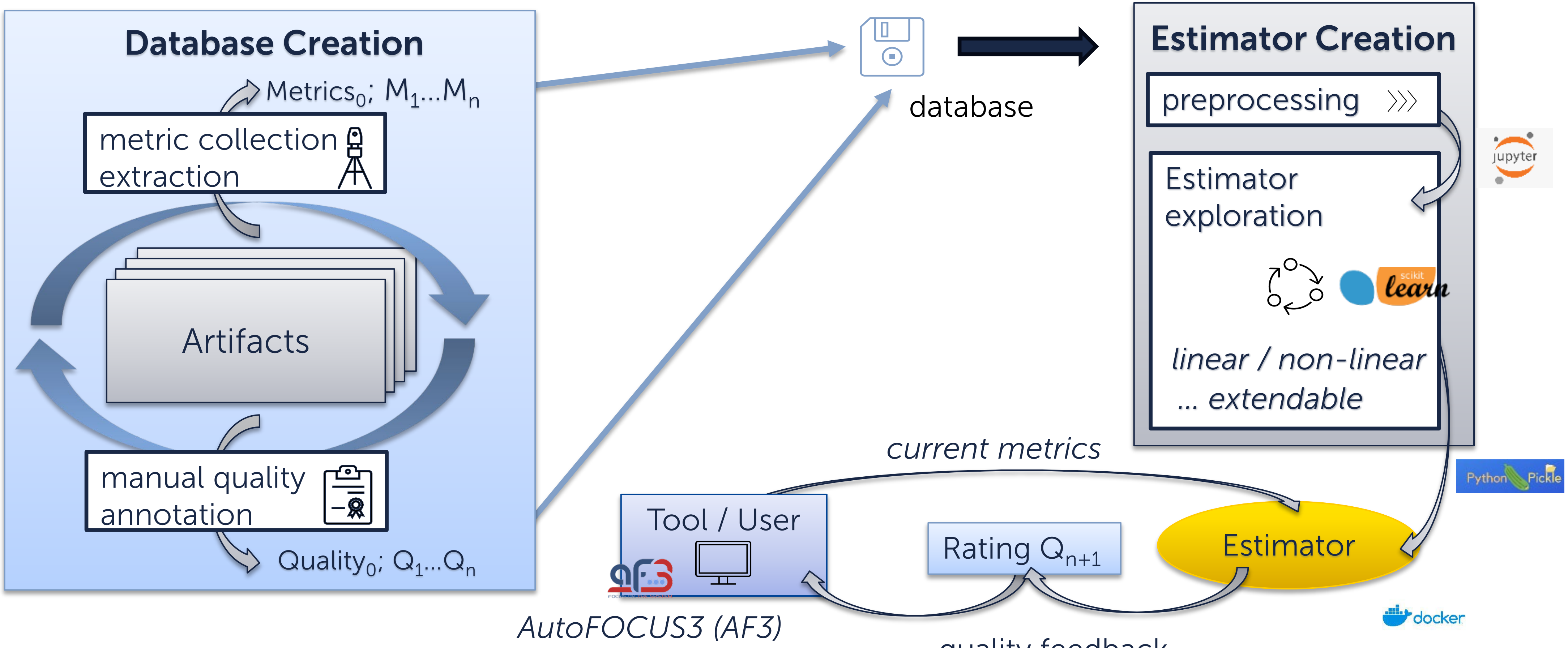
## Step 2: Define **attributes** and (manually) **rate** model **elements**

Example: Quality Attributes from Case Study (focus: comprehensibility)

| Name | Description | Aspects |
|---|---|---|
| Graphical Element Representation | Does the visualization of the element under review represents its impact and role for the system under review? | size, presentation, position, icon, … |
| Graphical Data Flow | Does the graphical representation convey an intuitive understanding of the data flow? | which elements consume / produce / transform data? |
| Element Naming | Do the element names fit in the context of the element under review?<br>Do they convey an intuitive understanding of the element's functionality? | name of element under review in relation to name of its constituent elements |
| Interface Representation | Are the port and channel names as well as the data types comprehensible in the context of the system under review? | port naming, datatype names and definitions |
| Abstraction Level | Is the decomposition of the element under review comprehensible?<br>Are its constituent elements at the expected granularity level? | Aspects: assess if the children of the element under review are on a similar abstraction level |

# Approach: Quality Estimator Architecture

## Step 4 & 5: Prototype Implementation



**Database Creation**

Metrics$_0$; M$_1$…M$_n$

metric collection extraction

Artifacts

manual quality annotation

Quality$_0$; Q$_1$…Q$_n$

database

**Estimator Creation**

preprocessing

Estimator exploration

*linear / non-linear*

*… extendable*

*current metrics*

Tool / User

*AutoFOCUS3 (AF3)*

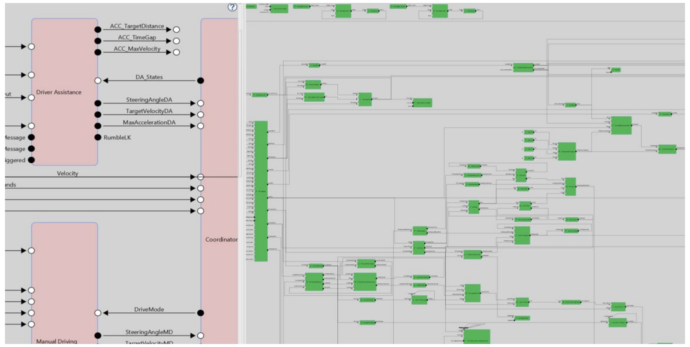
Rating Q$_{n+1}$

Estimator

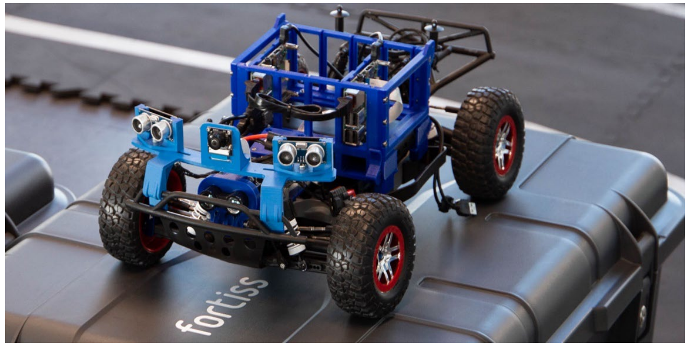
quality feedback

# Experiment: Platform and Data

## AutoFOCUS3 system models describing assisted driving functions



https://af3.fortiss.org/



https://git.fortiss.org/ff1/

**Experiment Data and Context**

- Software architecture models describing decomposition, data-flow and system behavior.
- 103 model versions from 14-week practical course at TU Munich
- 1837 elements rated for experiment.
- Quality ratings from 1 (bad) to 3 (good).
- Training data 80%; Test data 20%

# Experiment: Data Processing and Training

## Pipeline for Automatic Exploration of different Estimation Models

### Linear Regression

Correlation Analysis with Pearson, Spearman & Kendall

None of our extracted metrics highly correlates with one of our annotated comprehensibility ratings.

### Support Vector Machines

Grid Search with varying C and Gamma Values and Radial Basis Function kernel.

Best accuracy at reaching 0.79 on test data for the graphical element representation rating.

### Random Forests

Exploration of parameter constellations resulting in high accuracy on the test data.

Best results on the test data with average accuracy of 0.94 on the test data for all attributes.

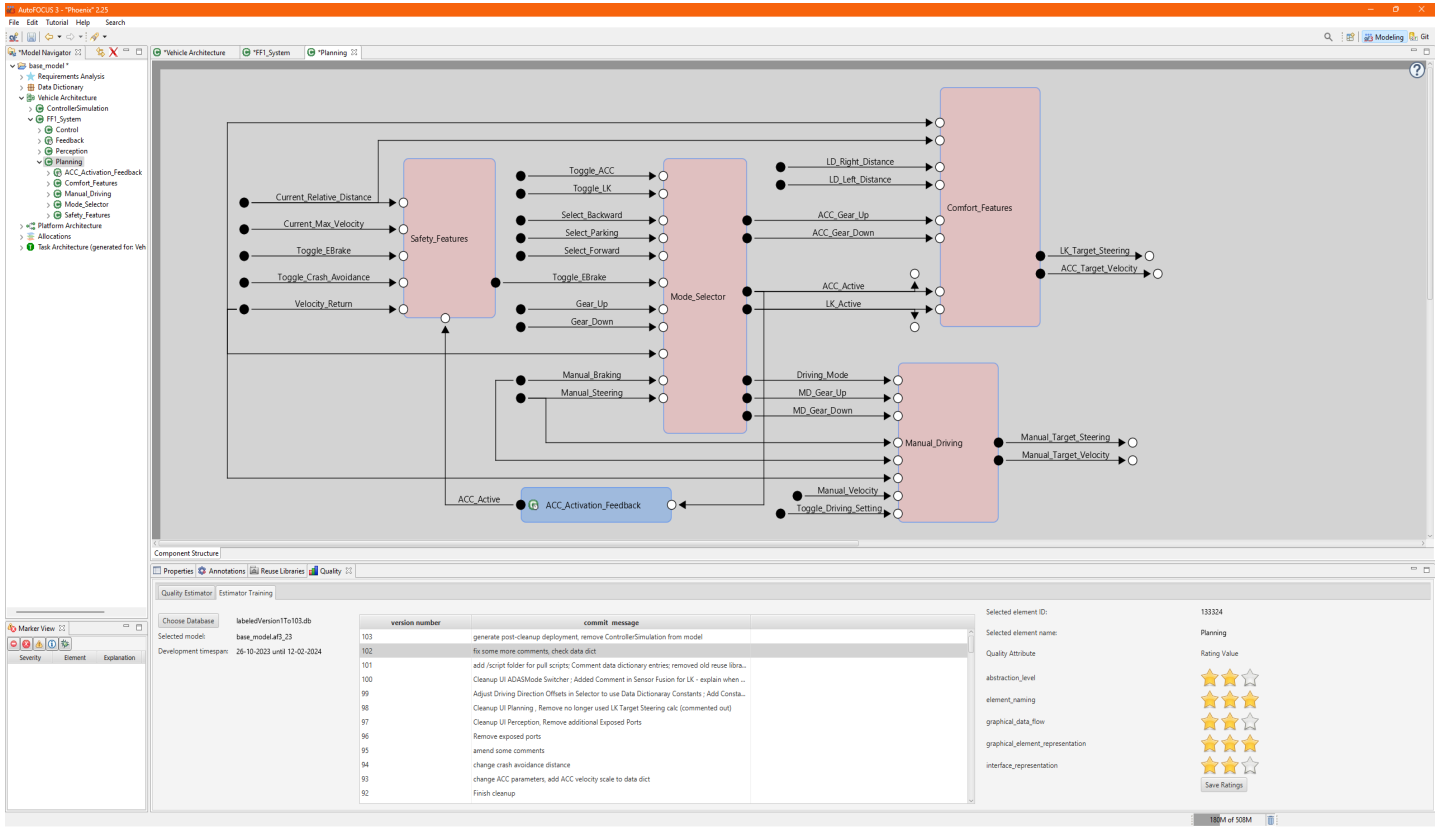
# Experiment: Random Forest Results

## Exploration of 3472 different Predictor Sets per Quality Attribute

| Predictors | Ratings | Accuracy |
|---|---|---|
| Nesting Level, #Channels, Surface Coverage, Deviation of Channel Lengths | Abstraction Level | 0.94 |
| Nesting Level, #Ports, #Total Ports, Density, Surface Coverage | Element Naming | 0.92 |
| Nesting Level, #Contained Elements, #Total Elements, Deviation of Channel Lengths | Graphical Data Flow | 0.94 |
| Clustering Coefficient, #Channels, #Total Elements, Deviation of Channel Lengths | Graphical Element Representation | 0.95 |
| Nesting Level, #Total Ports, #Total Leaf Elements, Surface Coverage | Interface Representation | 0.94 |

# Prototype: Quality Estimator Tooling

## AutoFOCUS 3 Quality Plugin



labeledVersion1To103.db
Selected model: base_model.af3_23
Development timespan: 26-10-2023 until 12-02-2024

| version number | commit message |
|---|---|
| 103 | generate post-cleanup deployment, remove ControllerSimulation from model |
| 102 | fix some more comments, check data dict |
| 101 | add /script folder for pull scripts; Comment data dictionary entries; removed old reuse libra... |
| 100 | Cleanup UI ADASMode Switcher ; Added Comment in Sensor Fusion for LK - explain when ... |
| 99 | Adjust Driving Direction Offsets in Selector to use Data Dictionaray Constants ; Add Consta... |
| 98 | Cleanup UI Planning , Remove no longer used LK Target Steering calc (commented out) |
| 97 | Cleanup UI Perception, Remove additional Exposed Ports |
| 96 | Remove exposed ports |
| 95 | amend some comments |
| 94 | change crash avoidance distance |
| 93 | change ACC parameters, add ACC velocity scale to data dict |
| 92 | Finish cleanup |

Selected element ID: 133324
Selected element name: Planning

| Quality Attribute | Rating Value |
|---|---|
| abstraction_level | ★★☆ |
| element_naming | ★★★ |
| graphical_data_flow | ★★☆ |
| graphical_element_representation | ★★★ |
| interface_representation | ★★☆ |

### Prototype Functionality

- **Automated Extraction** of **model metrics** from model version history into SQLite databases.
- **Navigation module** to checkout commits and **rate** selected model elements manually.
- **Load** and **store estimators**
- **Docker service** running with API waiting for model metrics to estimate **quality attributes.**

# Future Work on Model Quality Estimation

## Approaches to reduce Limitations and Threats to Validity

| Threats to Validity and Limitations | Future Work and Extensions |
|---|---|
| **Data Quality** | **Data Quality** |
| • Bias in labeling due to one person labeling the data along with low granularity in the assessment | • Study with 4x the data labeled by multiple people with more sophisticated quality ratings with higher granularity |
| • Initial selection of metrics originating from AF3 metamodel | • Metric selection also developed for industry case study |
| • Distribution of ratings on existing model data. | • Researching ways to improve training data quality label distribution. |
| **Approach** | **Approach** |
| • Industrial applicability | • Case study with industrial partner |

# Thank you for your Attention!

## Key take aways, benefits, limits and challenges

⊕

- **Domain-/Project-** specific **quality assessment**
- **Live** quality feedback
- **Saving costs** by detecting quality **degradation** early
- **Analysis** of quality development in model histories

- **Scarcity, quality** and **constraints** of ground truth data
- **Causality** between **metrics** and **quality attributes**
- **Adaption** of tooling to **industry standards**
- Approach **Industry Applicability**

## Contact

Konstantin Blaschke

**+49 160 93525405**

**blaschke@fortiss.org**

**Feel free to reach out to me for further discussions.**

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Gefördert durch:
Bundesministerium
für Wirtschaft
und Klimaschutz
aufgrund eines Beschlusses
des Deutschen Bundestages