Technische Universität Berlin

@andivogelsang
andreas.vogelsang@tu-berlin.de

Supporting the Development of **Cyber-Physical Systems** with **Natural Language Processing**: A Report

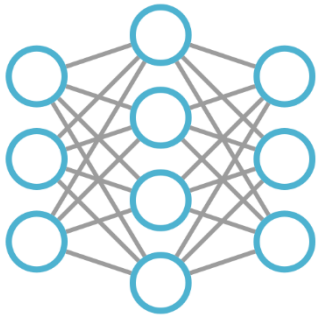
Andreas Vogelsang, Kerstin Hartig, Florian Pudlitz,
Aaron Schlutter, Jonas Winkler

Technical University of Berlin

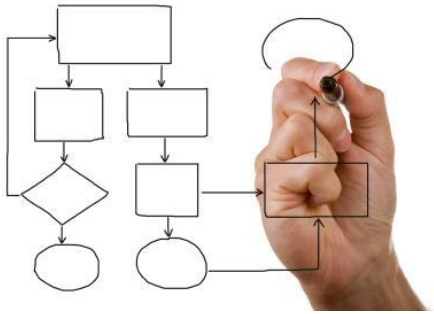
18.03.2019

Automated Systems Engineering Technologies

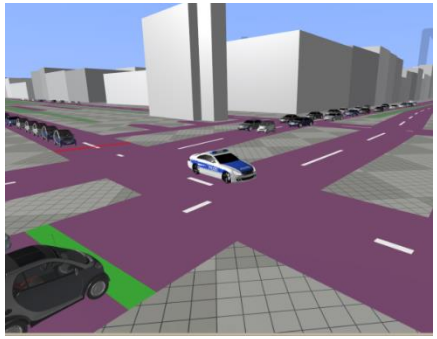
Our research focus: We research and develop technologies to support system engineers and automate time-consuming or error-prone tasks and process steps.



AI for RE



Model-based Systems Engineering



Validation by Simulation

Lead: Andreas Vogelsang

PhD students:

- Tiago Amorim
- Florian Brokhausen
- Patrick Ebel
- Kerstin Hartig
- Florian Pudlitz
- Stefan Rulewitz
- Aaron Schlutter
- Jonas Winkler
- Florian Wiesweg

Website: aset.tu-berlin.de

+ 3 student assistants

Cyber-Physical Systems (CPS)

- Observe environment by sensors, influence by actuactors
- Composed of mechanics, electronics, software
- Software most important and most critical

- Development is interdisciplinary
  - application domains
  - engineering disciplines
- Majority of development information is expressed in **natural language**
- Development is driven by strong safety and security constraints

Acatech: “agendaCPS: Integrierte Forschungsagenda Cyber-Physical Systems”, Springer, 2012

Natural Language in Cyber-Physical Systems

Technische Universität Berlin

Requirements

Test cases

Safety analyses and assessments

Architectural descriptions

...

User Feedback

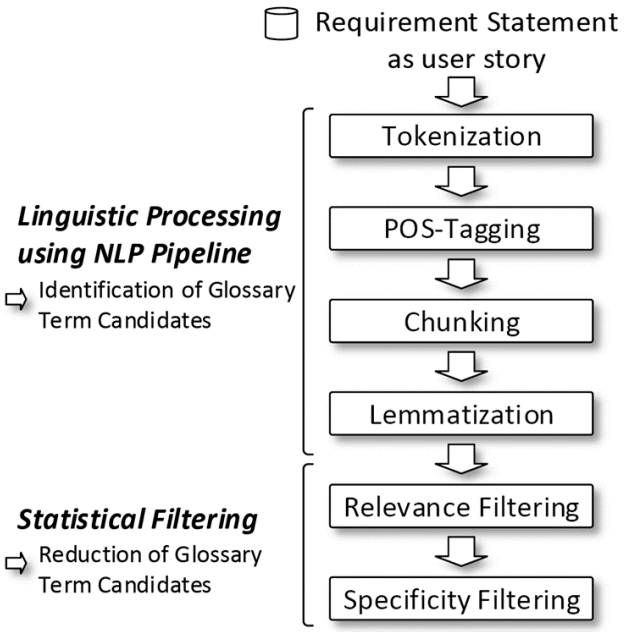
Information spread over hundreds of documents with thousands of entries (e.g., specification repository of a **telematics system** of a modern vehicle: **~30,000 documents**, **~2.5 million textual entries**)

Technische Universität Berlin

Past and Current Research on NLP for CPS Development

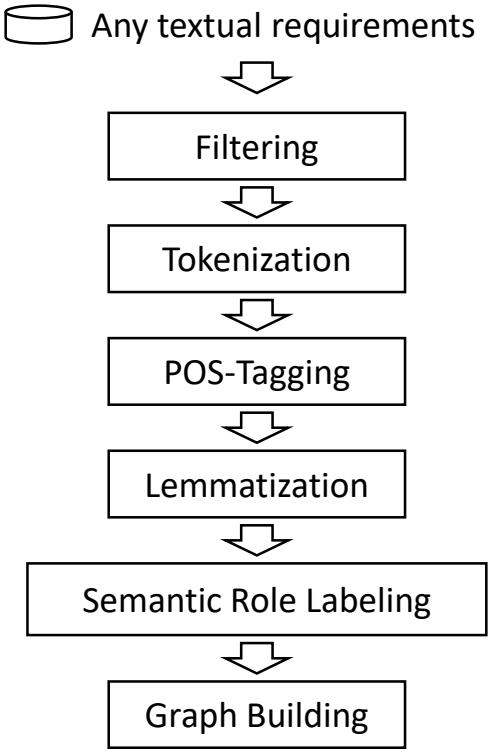
Automatic Knowledge Extraction

Glossary Term Extraction



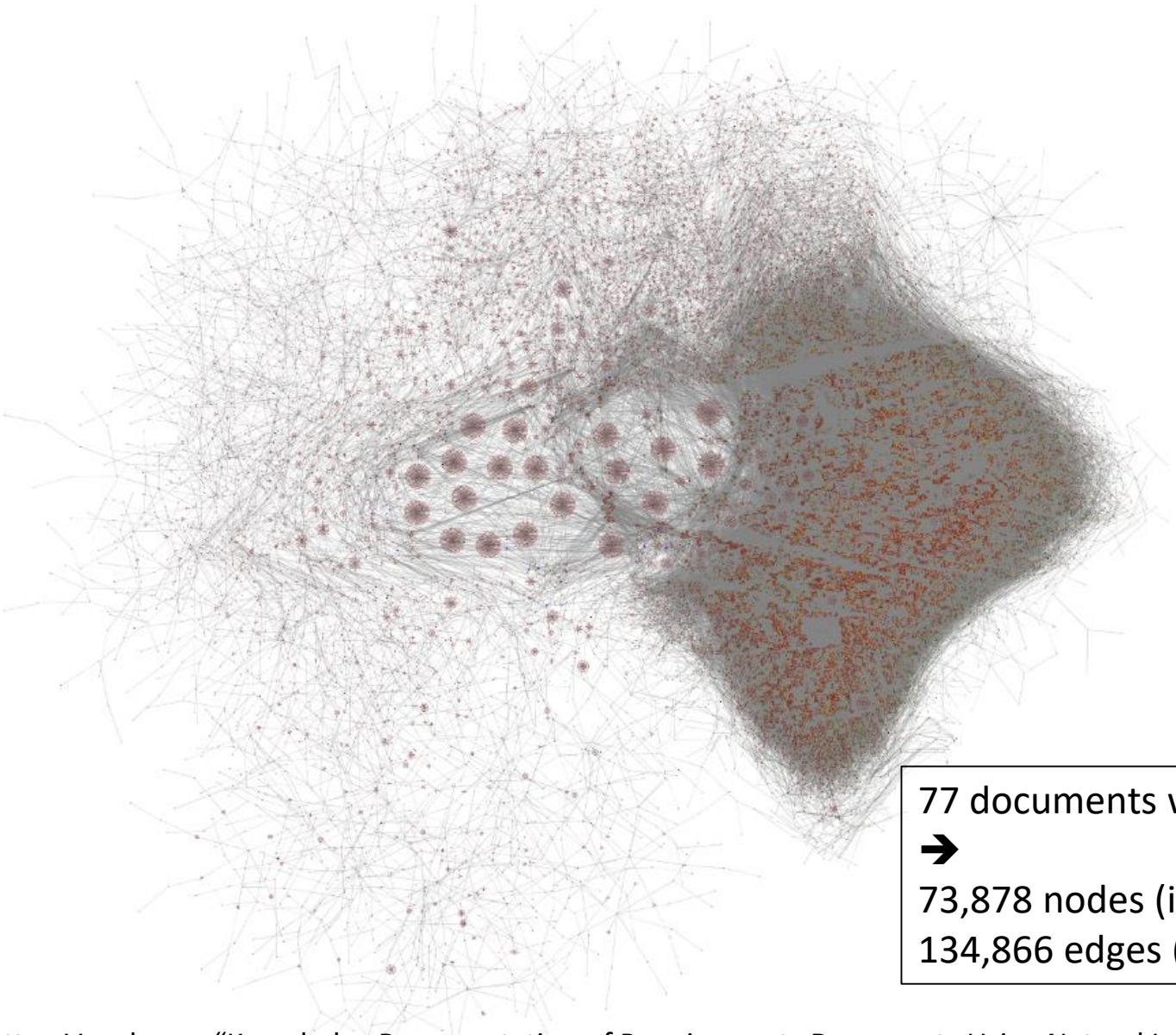
term candidate	no. reqs	term candidate	no. reqs	term candidate	no. reqs
home occupant	1088	water	149	escape	5
home	767	child	146	hair	5
home owner	621	pet owner	138	less energy	5
house	355	person	130	sound system	5
parent	333	temperature	116	drain	5
time	245	food	108	sport fan	5
alert	190	able	107	automatic door	5
door	175	safe	95	fall	5
light	170	phone	93	dirty	5
energy	165	day	92	rural home oocupant	5

Knowledge Graph Extraction



Gemkow, Conzelmann, Hartig, Vogelsang: "Automatic glossary term extraction from large-scale requirements specifications", *RE'18*

Automatic Knowledge Extraction



77 documents with 45,092 objects
➔
73,878 nodes (i.e., concepts used)
134,866 edges (i.e., concept relations)

Schlutter, Vogelsang: “Knowledge Representation of Requirements Documents Using Natural Language Processing”, *NLP4RE’17*

Expert Systems for Hazard and Risk Analysis

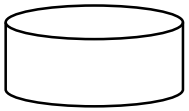
Technische Universität Berlin

Safety Expert

Structured Query

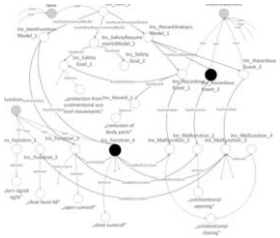
Give me all elements of type <X> related to element <Y>

HaRa repository

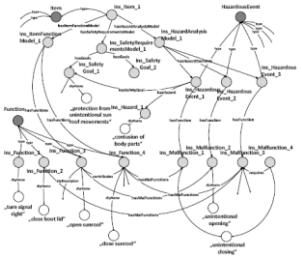


Representation

Graph-based query

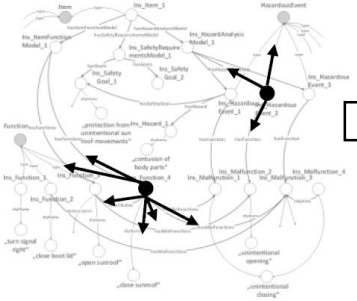


Semantic knowledge graph



Retrieval

Spreading Activation



Result

Filter elements of type <X>

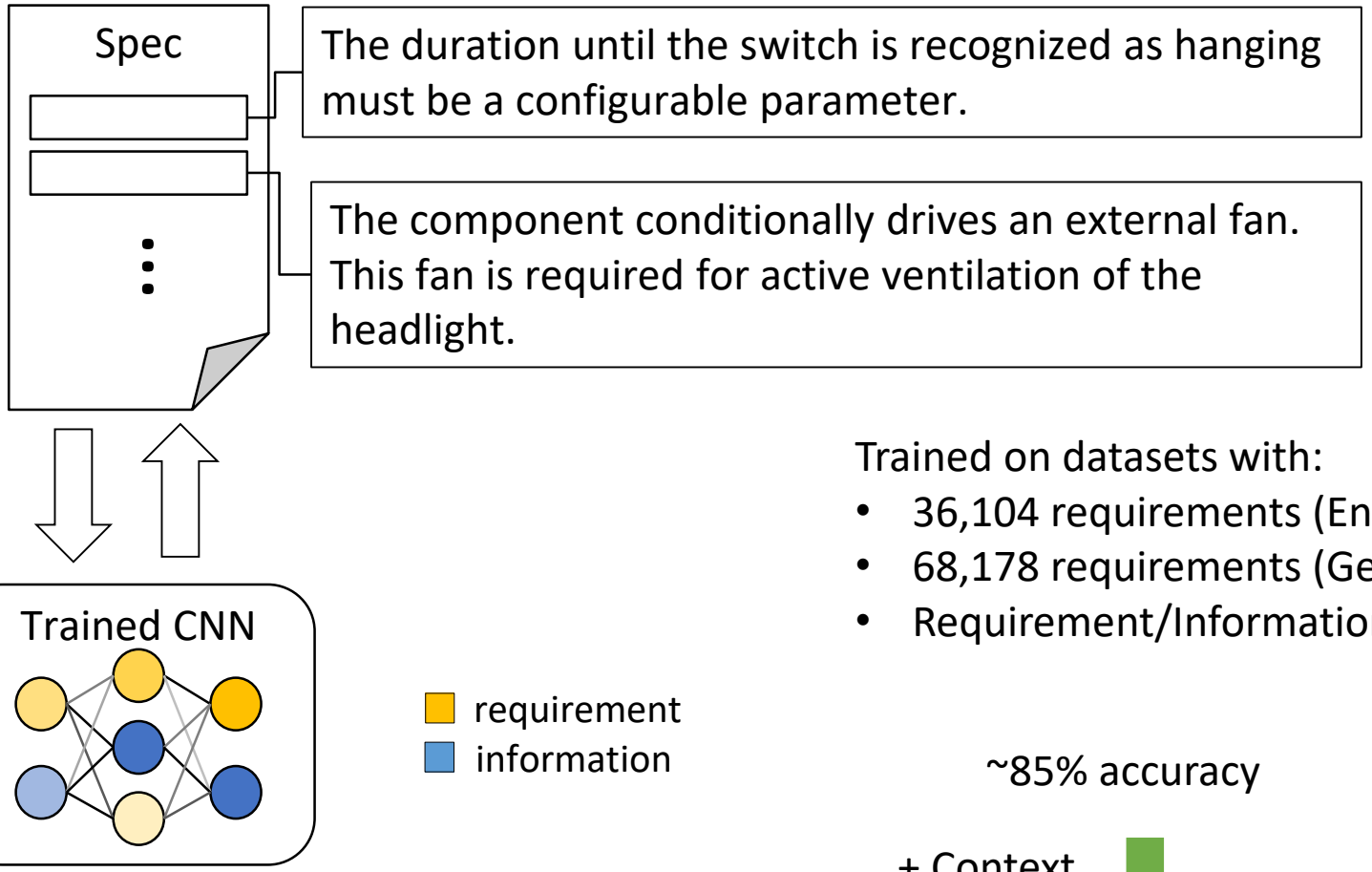
Recommendations

Elements	Relevance
...	
...	
...	
...	

155 HaRa documents:
600 functions; 1,700 malfunctions; 4,200 hazards; 540 safety goals

Expected Precision (EP): 0.66

Automatic Requirements Classification



Trained on datasets with:

- 36,104 requirements (English)
- 68,178 requirements (German)
- Requirement/Information: 60:40

~85% accuracy

+ Context

~90% accuracy

Winkler, Vogelsang: "Automatic Classification of Requirements Based on Convolutional Neural Networks", AIRE'16

Technische Universität Berlin

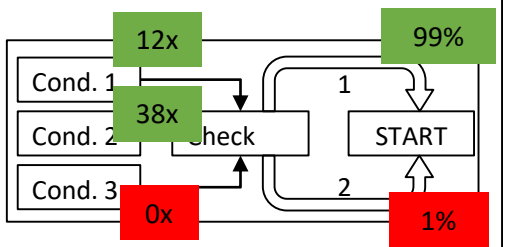
Future Research on NLP for CPS Development

Connecting NL Requirements and Simulation

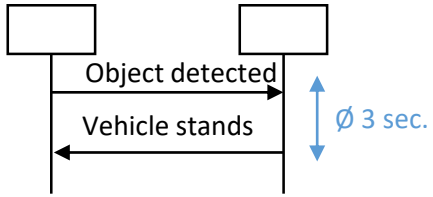
Technische Universität Berlin

Specification

1. Automatic windscreen wiping is activated in case of detected rainfall
2. The motor stops in the following cases
   - Crash 1%
   - Start/Stop 70%
   - Manual operation 29%
   - Key out of vehicle 0%
3. Procedure of function



4. Scenarios



Lightweight mapping between requirements and simulation events

Detected rainfall	⟷	Rain
Windscreen wiping	⟷	WS_Whip
Cond. 1	⟷	Ign_on
Cond. 2	⟷	v > 10 m/s
Cond. 3	⟷	Crash det

Data Analytics

Simulation

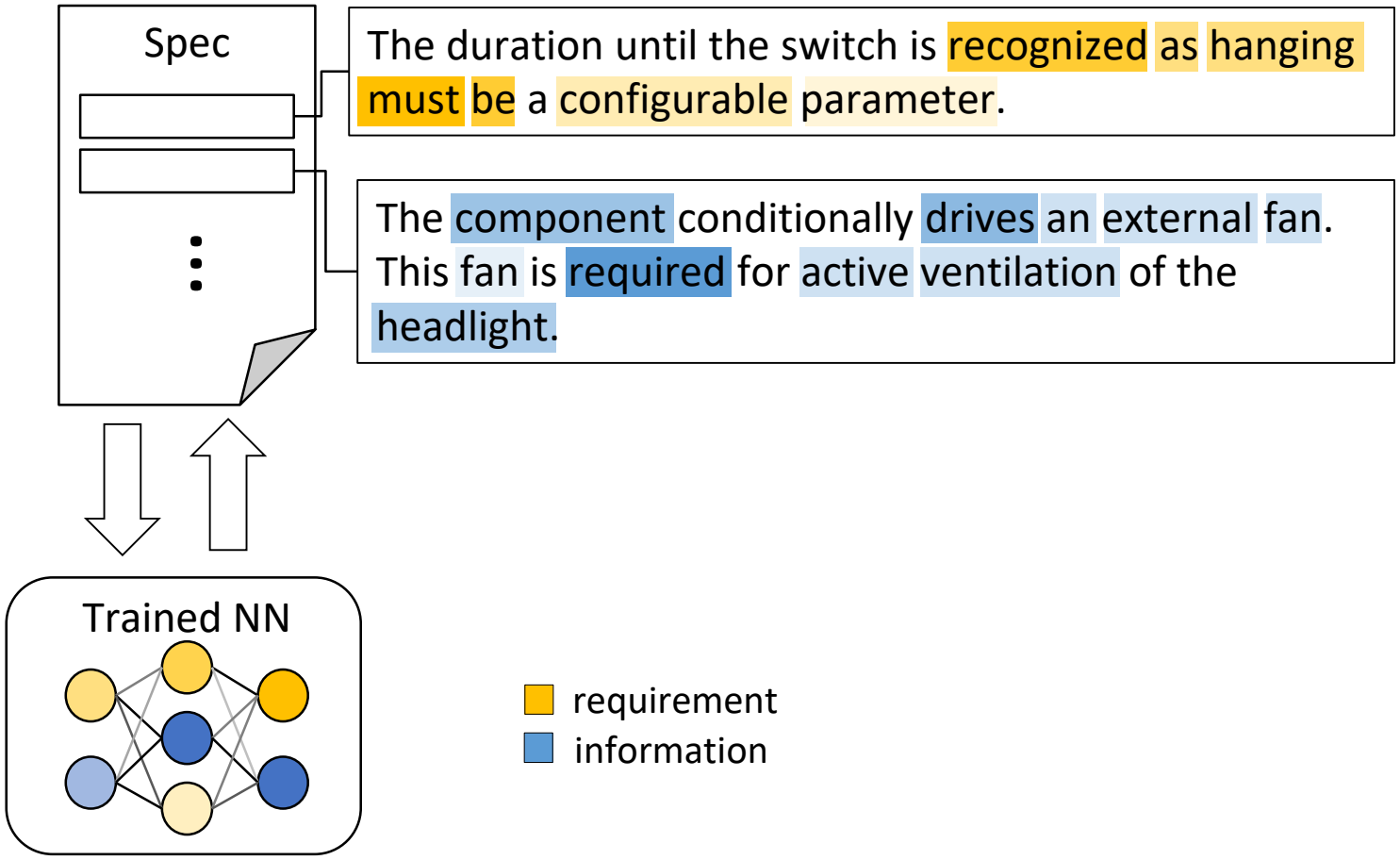
Log file

1. Ign_on
2. v = 50 km/h
3. Oncoming traffic
4. Rain starts
5. WS_Whip sent
6. Signal X
7. Break – Obstacle detected

…

Pudlitz, Vogelsang, Brokhausen: “A Lightweight Multilevel Markup Language for Connecting Software Requirements and Simulations”, *REFSQ’19*

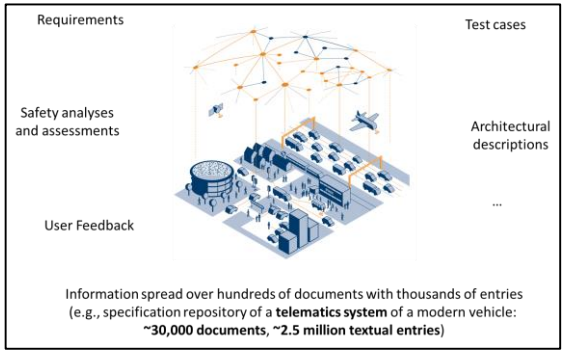
Explainability



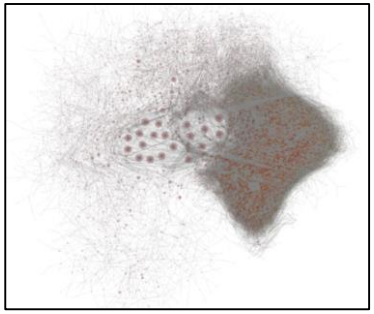
Winkler, Vogelsang: "What does my Classifier Learn? A Visual Approach to Understanding Natural Language Text Classifiers", *NLDB'17*

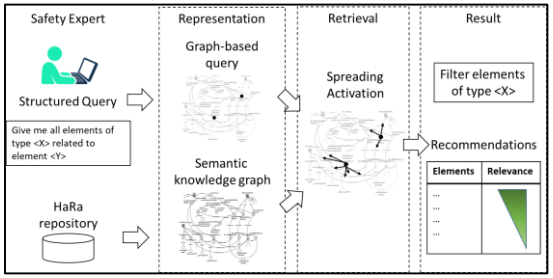
Summary

Natural Language in CPS

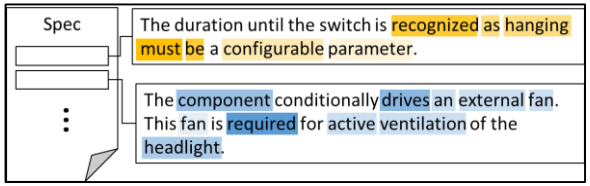


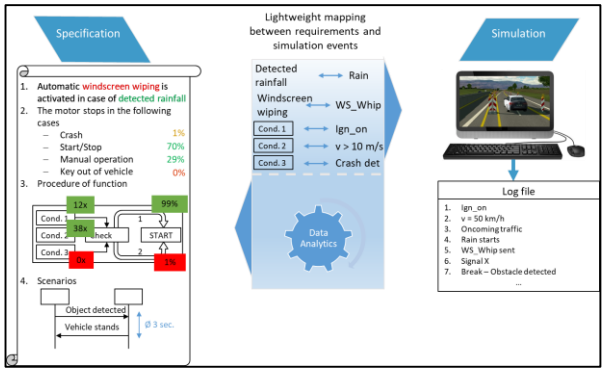
Past research





Future research





We **research and develop** technologies to **support system engineers** and **automate** time-consuming or error-prone tasks and process steps.

@andivogelsang
andreas.vogelsang@tu-berlin.de