RAKI

Projektnummer: 01MD19012C

Gefördert durch:

Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

Rapide Erklärbare Künstliche Intelligenz für Industrieanlagen

AP 7: Anwendungsfälle

SIEMENS

Overview

- Siemens Use Cases
  - Skill Description Learning
  - Cybersecurity
- Use Case 1: **Skill Description Learning**
  - Data
  - Machine Learning Evaluation
  - Verbalization Evaluation
- Use Case 2: **Cybersecurity**
  - Data
  - Machine Learning Evaluation
  - Verbalization Evaluation
- Conclusion and Outlook

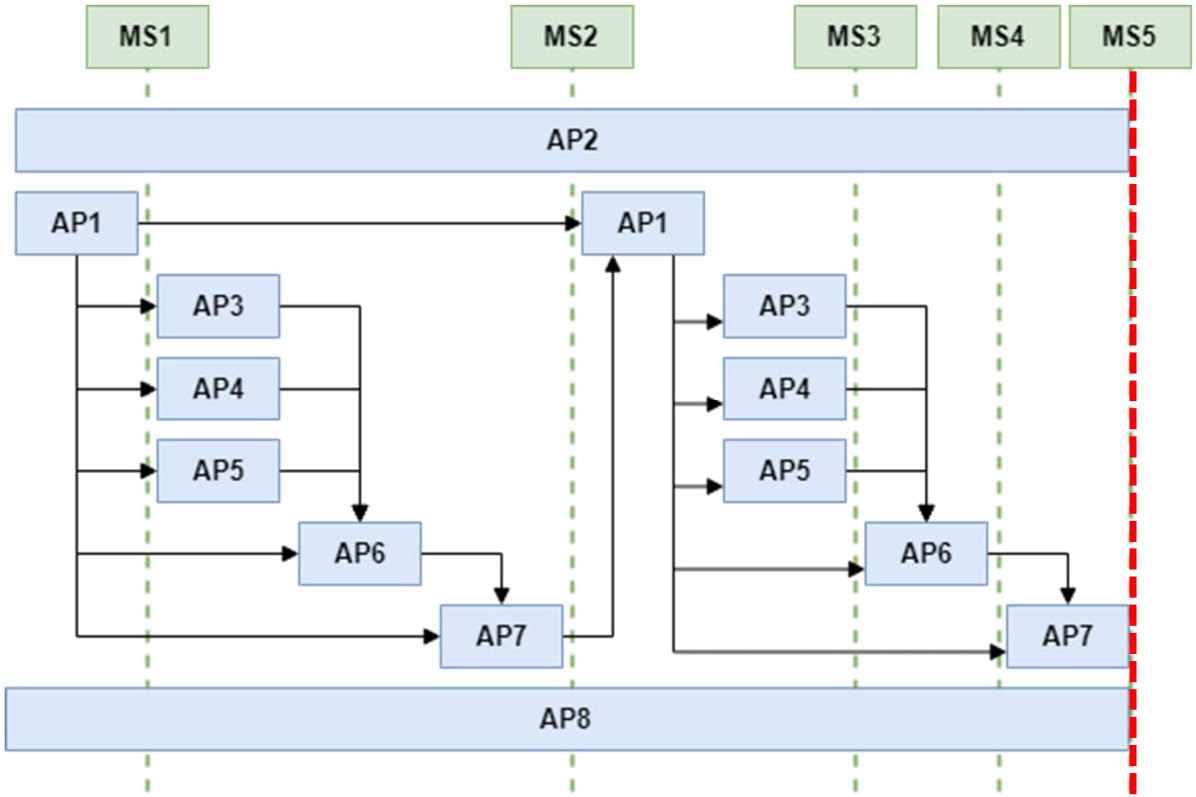
Current State

RAKI SIEMENS

Use case development:

- Skill description learning
- Skill matching
- Turbine package classification
- Turbine defect classification
- Experimental production plant
- Component selection tool
- Cybersecurity

Evaluation of first prototype on skill description learning use case



Use Case 1: Skill Description Learning

RAKI SIEMENS

Motivation:

- Cyber-physical systems for more flexibility, adaptability, and transparency in production
- Skill matching assigns operations in a production process to machines
- Need skill descriptions of the machines and skill requirements of the operations
- In some cases, skill descriptions might not be available at all, e.g., in the case of a legacy module
- Defining and digitizing skill descriptions of a production module are typically done manually by a domain expert

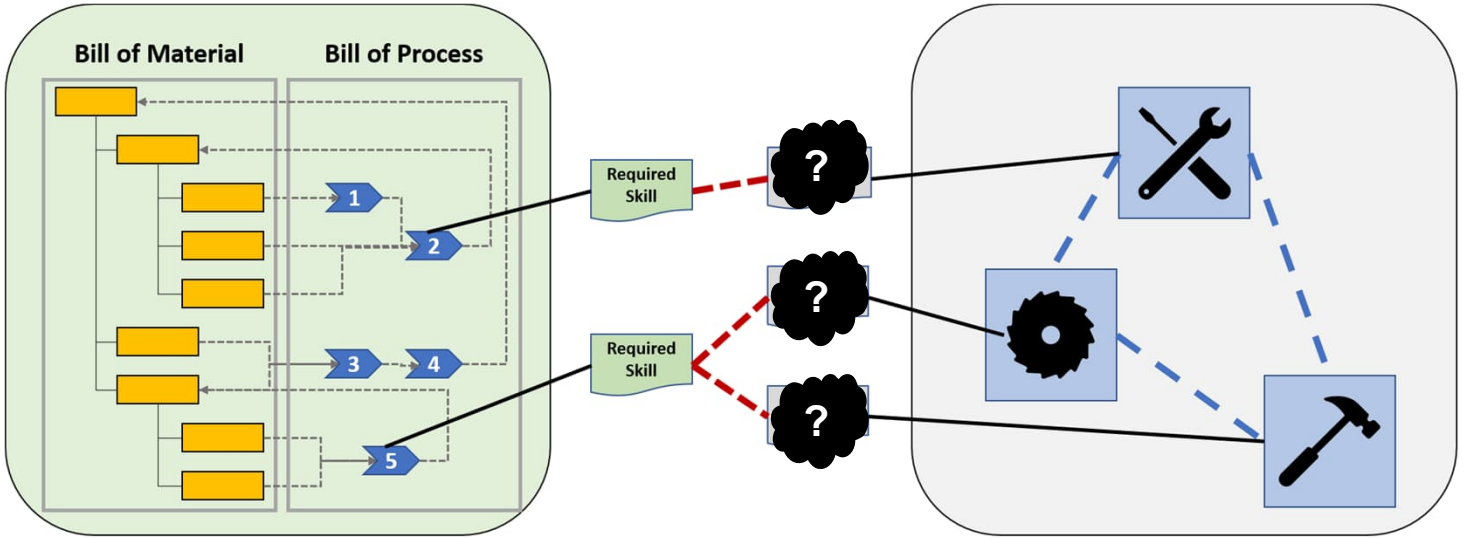
➢ Equip machines with explicit digitized skill descriptions, detailing their capabilities

Automatic skill description learning would minimize the labor time and domain expertise needed to equip production modules with their descriptions.

Use Case 1: Skill Description Learning

Given:

- Production log data – instance data
- Production ontology – background data



Desired: Skill descriptions of machines

Use Case 2: Cybersecurity

RAKI SIEMENS

Motivation:

- Continuous increase in cyber attacks has given rise to a growing demand for modern intrusion detection systems (IDS) that leverage machine learning to detect cyber attacks
- Many relevant examples of the application of deep learning and similar techniques for IDS already exist
- However, their drawbacks include alarm flooding problems and a lack of explainability

➢ Detect and explain cybersecurity anomalies

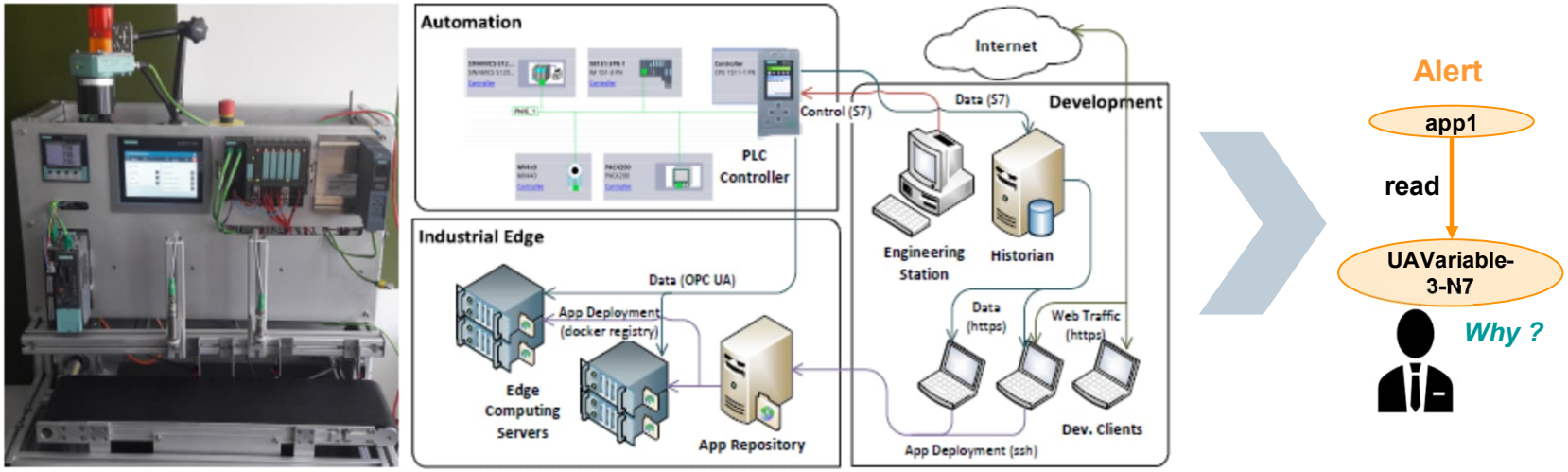
Generate trust and user support through explanation – explainable AI such that users can understand as suspicious flagged data events. Domain experts can make better triage decisions when they understand the security alerts given by the AI.

Use Case 2: Cybersecurity

RAKI SIEMENS

Given:

- Network log data – instance data
- Cybersecurity ontology – background data



Desired: Characterization of anomalies

Skill Description Learning: Data

RAKI SIEMENS

Knowledge base – ontologies:

- Process: Ontology modeling all operations that can be carried out at the production plant
- Production: Ontology modeling the equipment of the production plant, especially the machines
- Product: Ontology modeling the building blocks of the products

Setting:

- Number of concepts: 83
- Number of individuals: 79
- Number of data properties: 18
- Number of object properties: 6
- Number of learning problems: 23

Skill Description Learning: Example

Positive examples: Operation instances of the skill “Inserting by Module 1”
(**Negative examples**: Operation instances of other skills than “Inserting by Module 1”)

Background knowledge: domain knowledge regarding operations and corresponding instances e.g., material used in instances

Desired class expression examples:

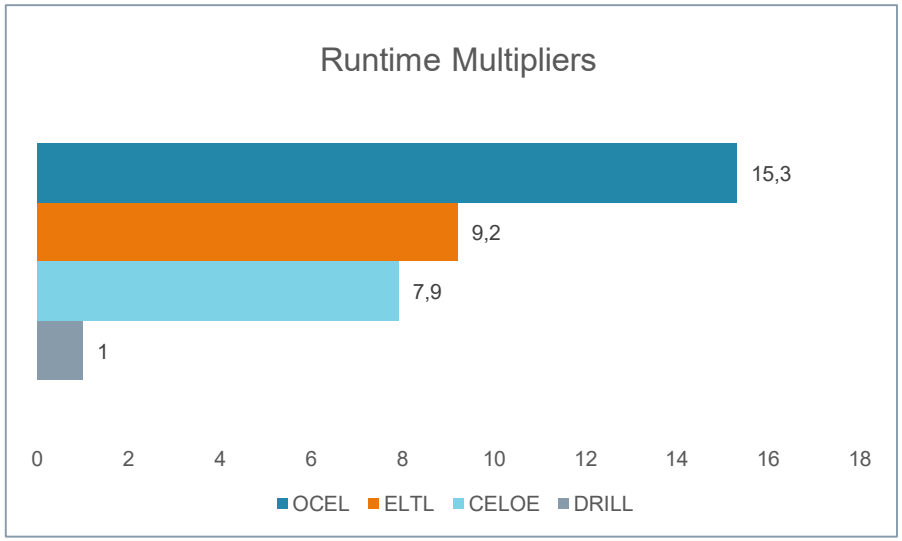
- Involves initial material roof 1 and base 2
- Has resulting material base 1 connected with roof 1
- Resulting material has orientation value 180
- Has optimization parameter value “quality“
- …

We want to arrive at a complete and precise description for the skill.

DRILL Experiments Results Skill Description: Quantitative

	F-measure (mean)	F-measure (std)	Accuracy (mean)	Accuracy (std)
DRILL	**1**	0.02	**1**	0
CELOE	**1**	0	**1**	0
OCEL	-0.01	0	0.96	0.21
ELTL	0.91	0.24	0.97	0.11

	NumClassTested (mean)	NumClassTested (std)	Runtime (mean)	Runtime (std)
DRILL	285.83	846.34	**0.39**	0.59
CELOE	223.70	521.92	3.07	0.23
OCEL	7520.48	3154.19	5.98	0.04
ELTL	-1	0	3.57	1.14



- The mean and standard deviation are calculated using all 23 learning problems
- DRILL and CELOE both have perfect F-measure and accuracy

DRILL Experiments Results Skill Description: Qualitative

Expected Outcome	DRILL	CELOE	OCEL	ELTL
hasPositionParam_0 ⊓ hasOptimizationParamQuality ⊓ hasOrientationParam_0 ⊓ hasPunchingText_ ⊓ involvesInitialMaterialBase2	**hasPositionParam_0 ⊓ involvesInitialMaterialBase2**	ChargingByEmptyModule1	ChargingByEmpty Module1	ChargingByEmpty Module1
hasOptimizationParamQuality ⊓ hasOrientationParam_0 ⊓ hasPunchingText_ ⊓ involvesInitialMaterialPlate1 ⊓ ((hasPositionParam_4 ⊓ involvesInitialMaterialBase2) ⊔ hasPositionParam_0))	**involvesInitialMaterialBase2 ⊓ involvesInitialMaterialPlate1**	involvesInitialMaterialPlate1 ⊓ (PunchingByIndustrialRobot1 ⊔ (InsertingByIndustrialRobot1 ⊓ involvesInitialMaterialBase2))	null	hasPunchingText_ ⊓ involvesInitialMaterial Plate1

- For DRILL, we can include an “exclude concept” functionality to avoid trivial concepts. This functionality improves the qualitative results while there is a quantitative results trade-off.
- For CELOE, 18 out of 23 target expressions are highly trivial; same for OCEL, with another two null results. For ELTL, at least 16 out of 23 predictions are equally trivial, while for DRILL, there are no obviously trivial results.

Verbalization Results: Examples

DRILL Class Expression:

(hasPositionParam_0 ⊓ involvesInitialMaterialBase2)

Verbalization:

Every charging by empty module 1 has position parameter 0 whose an involves initial material base 2.

DRILL Class Expression:

(hasPositionParam_1 ⊓ hasPunchingText_)

Verbalization:

Every inserting by assembly module 2 is a has position parameter 1 whose a has no punching text.

DRILL Class Expression:

(hasPositionParam_4 ⊓ (involvesInitialMaterialBase2 ⊓ (involvesInitialMaterialBlock4 ⊔ involvesInitialMaterialPlate1)))

Verbalization:

Every industrial robot 1 is a has position parameter 4 whose an involves initial material base 2 whose an involves initial material block 4 or an involves initial material plate 1.

Cybersecurity: Data

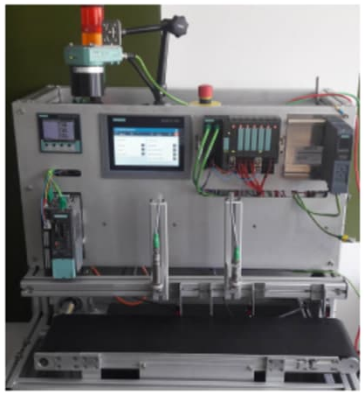
Knowledge Base – Ontology based on industrial demonstrator system

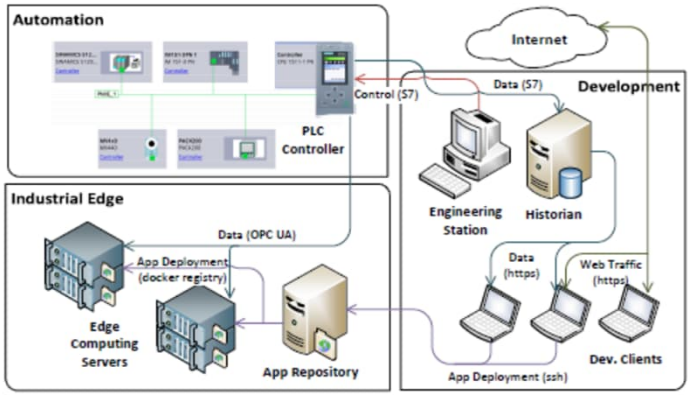
The demonstrator system for e.g. measuring the height of objects for quality control amongst other capabilities follows the design of modern industrial systems integrating IT and OT elements.

- **Automation part**: Summarizes the engineering design of the manufacturing prototype.
- **Edge part:** Contains app initialization events, data events and the applications.
- **Network part**: Contains network connections and their properties as subclasses, as well as IPs and their subdomains with local and global and automation, development and edge networks as subclasses.

Setting:

- Number of concepts: 3595
- Number of individuals: 19810
- Number of properties: 74
- Number of data properties: 0
- Number of object properties: 74





SIEMENS

Cybersecurity: Example

RAKI SIEMENS

Positive examples: Anomalies in a certain category, e.g. “Credential Use”
Negative examples: Normal data events in that category

Background knowledge: domain knowledge modelling a wide range of cybersecurity-relevant knowledge such as product information and are linked to domain-specific knowledge, coming from industrial automation system demonstrator

Desired class expression examples:

- Security Breach is something whose service is SSH
- Credential Use is something whose initial server is something that is part of a development network
- …

We want to get an explanation for a cyber security alert.

DRILL Experiments Results Cybersecurity: Quantitative

RAKI SIEMENS

- Number of parameters: 8393889
- KG preprocessing took : 178.35s
- DrillAverage
    - **F-measure: avg. 0.81 | std. 0.14**
    - **Accuracy: avg. 0.71 | std. 0.19**
    - **NumClassTested: avg. 6975.00 | std. 1001.26**
    - **Runtime: 301.02s | std.13.11**

➢ Only DRILL can produce results! Data too large for CELOE, OCEL, ELTL → **Time Out!**

DRILL Experiments Results Cybersecurity: Qualitative

DRILL	Verbalization	Context
\u00acinitiatedFrom_192.168.0.17_CL	**Every credential use anomaly is something that is not an initiated from 192 . 168 . 0 . 17**	Access to OPC-UA server from an IP address that corresponds to a development host. IP Address 192.168.0.17 corresponds to a edge host. Anomaly is that usually the access to OPC-UA server is initiated from an IP address that corresponds to a edge host, but here we see it is not!
\u00acRead_UAVariable-currentL2_CL	**Every variables access anomaly is something that is not a read ua variable - current l 2**	App changes the way it accesses some variables (e.g. writes instead or reads) or App accesses variables completely unrelated to those accessed usually. Current L2 has datatype REAL, instead of String or localized Text, which is an anomaly.

- As the DL-Learner cannot produce output, there is no qualitative comparison possible.
- Insightful explanations for 4 out of 9 learning problems are generated.

Conclusion and Outlook

RAKI SIEMENS

Skill Description Learning:

- Perfect F-Measure and Accuracy with DRILL & CELOE
- 8x faster than CELOE, 15x faster than OCEL
- All DRILL results are non-trivial as opposed to the other methods
- Verbalization makes the results usable for domain experts

Cybersecurity Use Case:

- Only possible with DRILL, conventional methods time out
- Avg. F-Measure of 81% and avg. accuracy of 71% for DRILL
- Insightful explanations for domain experts for roughly 50% of cyber attacks

Outlook:

- Skill description learning can be used for skill matching in an industrial automatization context
- With increasing need for XAI in the cybersecurity domain, potential to integrate this technology in a cybersecurity monitoring tool

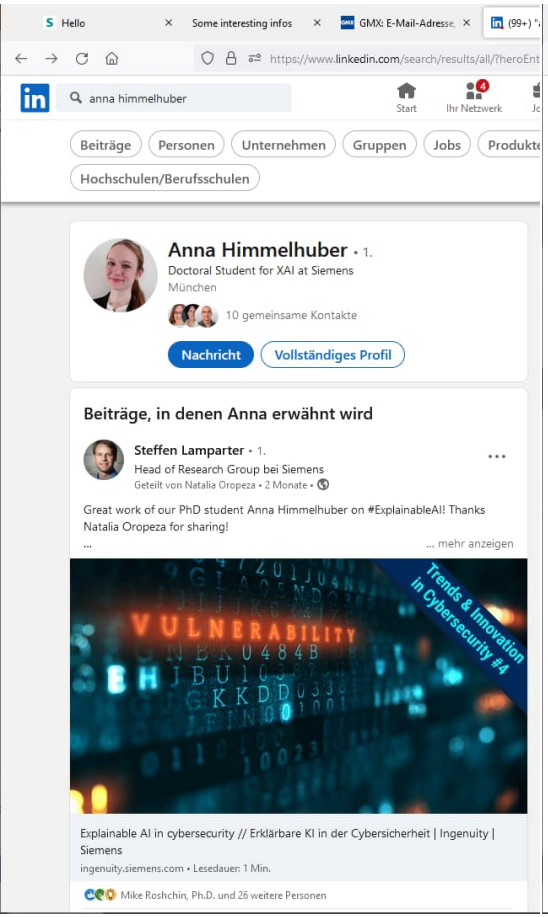
Selected Publications

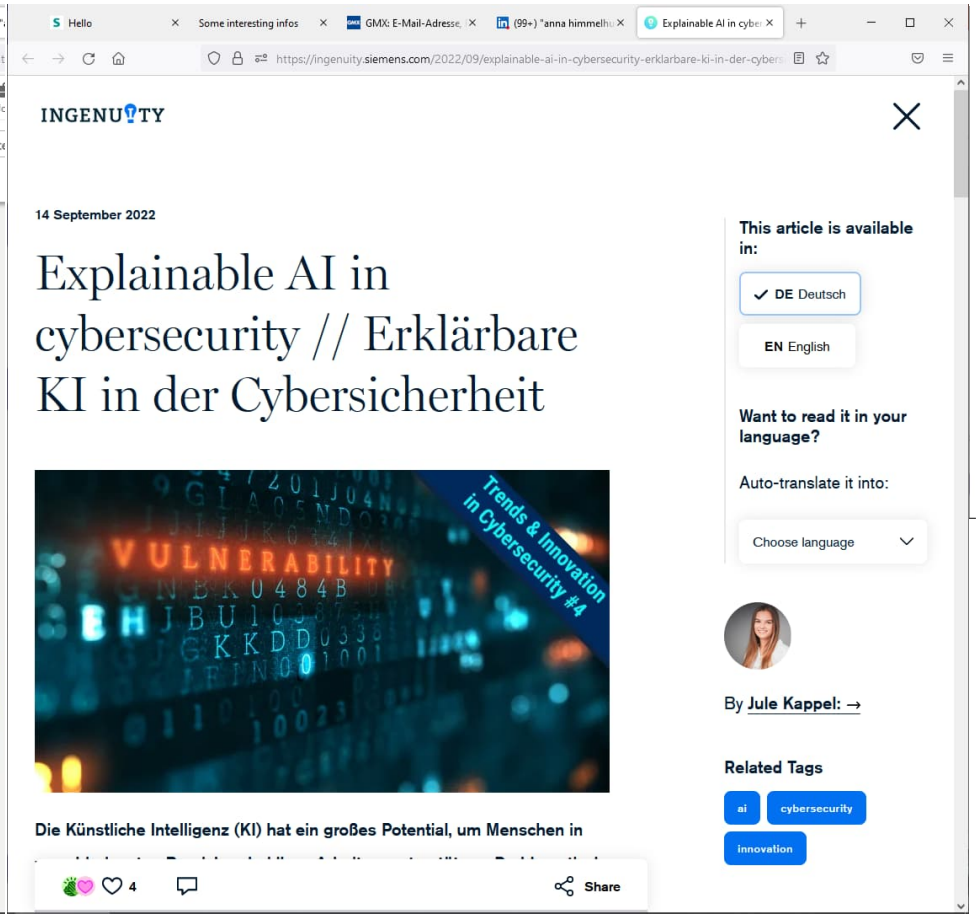
RAKI SIEMENS

- "Ontology-based Skill Description Learning for Flexible Production Systems", Himmelhuber, A., Grimm, S., Runkler, T., Zillner, S., ETFA 2020, Vienna
- "Skill Description Learning: Wissen über Maschinen rekonstruieren", Pressemeldung, see https://raki-projekt.de/news/2020-26-05-Skill-Description-Learning
- "Neural Multi-Hop Reasoning With Logical Rules on Biomedical Knowledge Graphs", Liu, Y., Hildebrandt, M., Joblin, M., Ringsquandl, M., Raissouni, R., Tresp, V., ESWC 2021
- "A New Concept for Explaining Graph Neural Networks ", Himmelhuber, A., Grimm, S., Zillner, S., Ringsquandl, M., Joblin, M. and Runkler, T., International Workshop on Neural-Symbolic Learning and Reasoning 2021
- "Combining Sub-symbolic and Symbolic Methods for Explainability ", Himmelhuber, A., Grimm, S., Zillner, S., Joblin, M., Ringsquandl, M. and Runkler, T., International Joint Conference on Rules and Reasoning 2021
- "TLogic: Temporal Logical Rules for Explainable Link Forecasting on Temporal Knowledge Graphs", Liu, Y., Ma, Y., Hildebrandt, M., Joblin, M., Tresp, V., AAAI 2022
- "Detection, Explanation and Filtering of Cyber Attacks Combining Symbolic and Sub-Symbolic Methods“, Himmelhuber, A., Dold, D., Grimm, S., Zillner, S., and Runkler, T., IEEE Symposium Series On Computational Intelligence 2022
- "Rapid Explainability for Skill Description Learning ", Demir C., Himmelhuber, Liu Y., Bigerl A., Moussallem D., Ngomo A., ISWC 2022

Selected Publications

RAKI SIEMENS





Die Künstliche Intelligenz (KI) hat ein großes Potential, um Menschen in verschiedensten Bereichen bei ihrer Arbeit zu unterstützen. Problematisch in der Anwendung ist jedoch häufig, dass die Entscheidungen der KI für den Menschen nicht nachvollziehbar sind und daher wenig vertrauenswürdig erscheinen. Die Arbeitsweise einer KI sowie ihre Resultate für die AnwenderInnen so verständlich wie möglich zu machen, ist das Ziel der so genannten „Erklärbaren Künstlichen Intelligenz". Anna Himmelhuber ist PhD-Studentin bei Siemens und forscht zu Erklärbarer KI. Im Interview erklärt Anna, wie diese Technologie in der Cybersecurity hilfreich sein kann und welche Chance sie in der KI für die Zukunft sieht.

4 Share

Du hast angesprochen, dass die KI neben Chancen auch Risiken mit sich bringt. Wie kann man sich gegen diese schützen?

Anna Himmelhuber (T DAI SMR-DE)

Anna: Die ständige Überwachung der Modelle ist sehr wichtig, um sicherzugehen, dass diese robust bleiben. Eine Möglichkeit ist auch, selbst Attacken gegen die eigenen Modelle laufen zu lassen und zu beobachten, wie diese auf den Angriff reagieren. Anhand der Ergebnisse können die Modelle geupdated und somit sicherer gemacht werden. Und auch das Thema Erklärbarkeit spielt bei der Sicherheit der eigenen Systeme eine Rolle: Durch Erklärbare KI wird Transparenz darüber geschaffen, was in den KI-Modellen

Besten Dank für die Aufmerksamkeit!

RAKI

Projektnummer: 01MD19012C

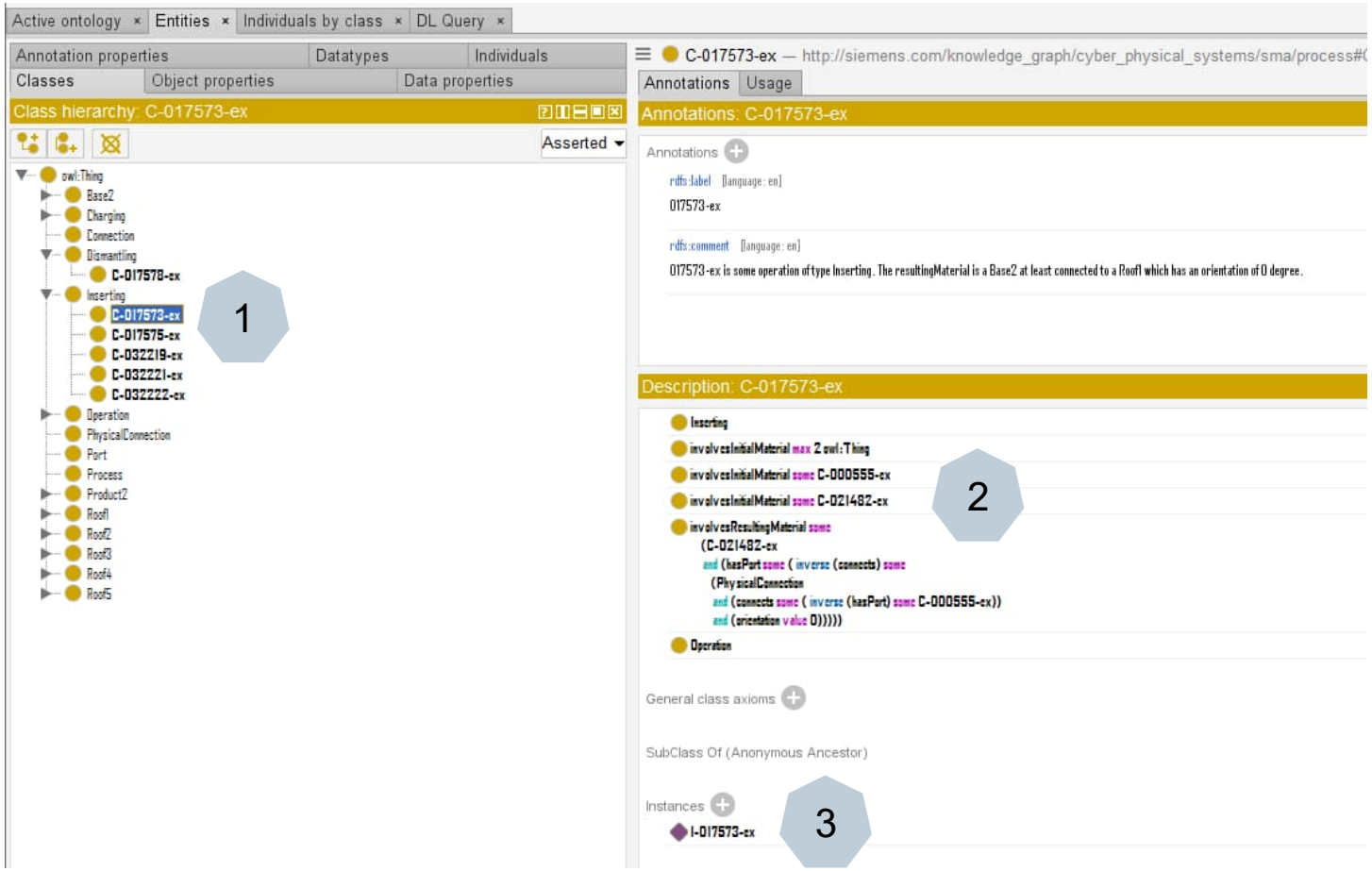
Gefördert durch:

Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

SIEMENS

Skill Description Learning: Ontology



1 Operation of the skill “Inserting by Module 1”

2 Background knowledge regarding the operation, e.g., material involved

3 Instance of the operation

DRILL Experiments Results: Quantitative

RAKI SIEMENS

	F-measure (mean)	F-measure (std)	Accuracy (mean)	Accuracy (std)	NumClassTested (mean)	NumClassTested (std)	Runtime (mean)	Runtime (std)
DRILL	0.77	0.23	0.96	0.04	6326.04	5293.18	3.21	1.96
CELOE	1	0	1	0	223.7	521.92	3.87	0.35
OCEL	-0.01	0	0.91	0.28	7192.78	3024.48	6.76	0.26
ELTL	0.91	0.24	0.97	0.11	-1	0	4.32	1.13

- The mean and standard deviation are calculated using all 23 learning problems.
- DRILL has lower F-measure than the DL-Learner algorithms, while the accuracy performance is comparable. This could be due to the fact that DRILL excludes certain concepts in the class expressions, which makes the learning more difficult, while the DL-Learner algorithms could provide trivial solutions with better performance.
- DRILL tests a significantly higher number of classes than CELOE (around 30 times more) and slightly fewer classes than OCEL.
- DRILL only needs around the same time as CELOE and half the runtime as OCEL, showing the scalability of the approach.