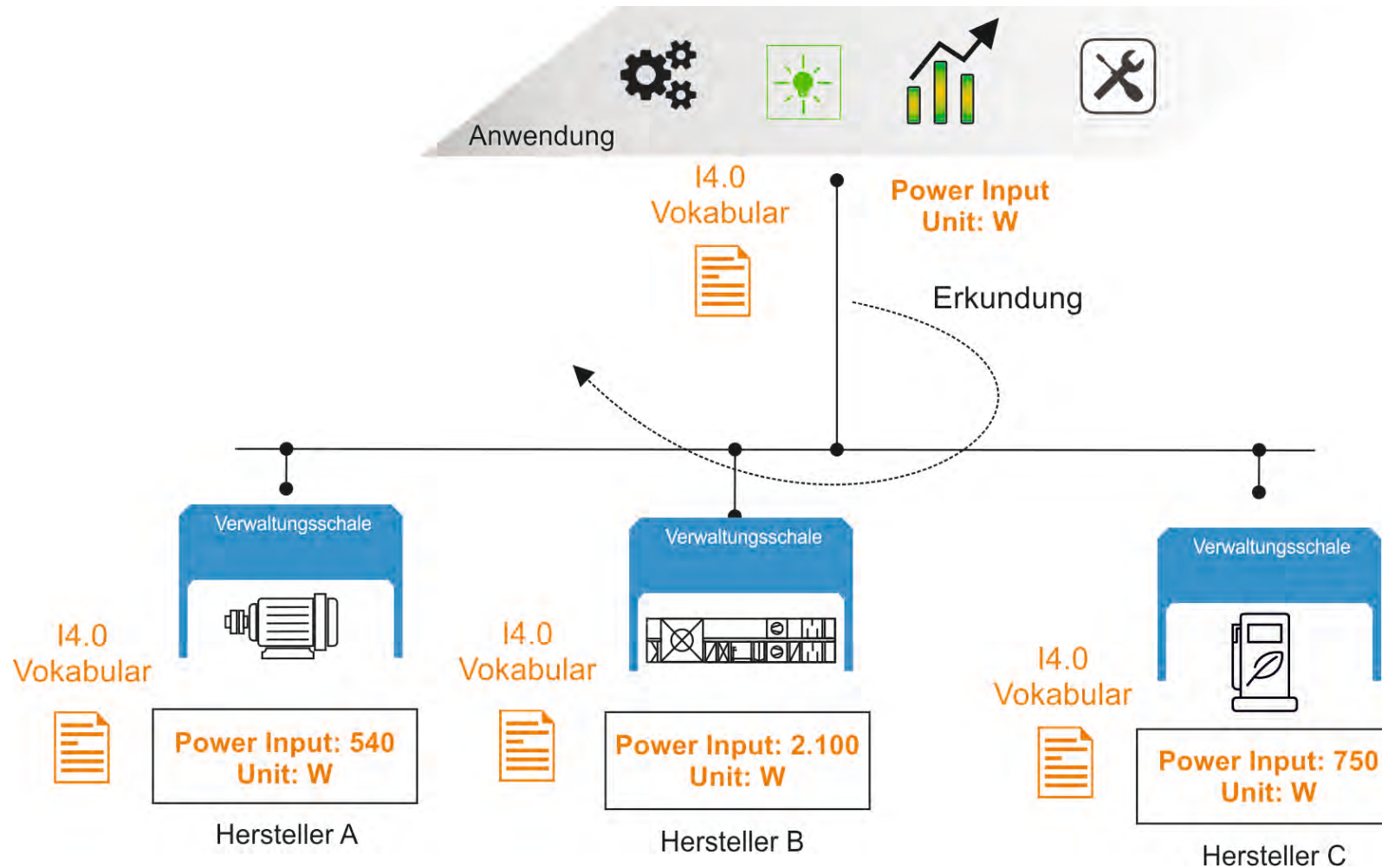
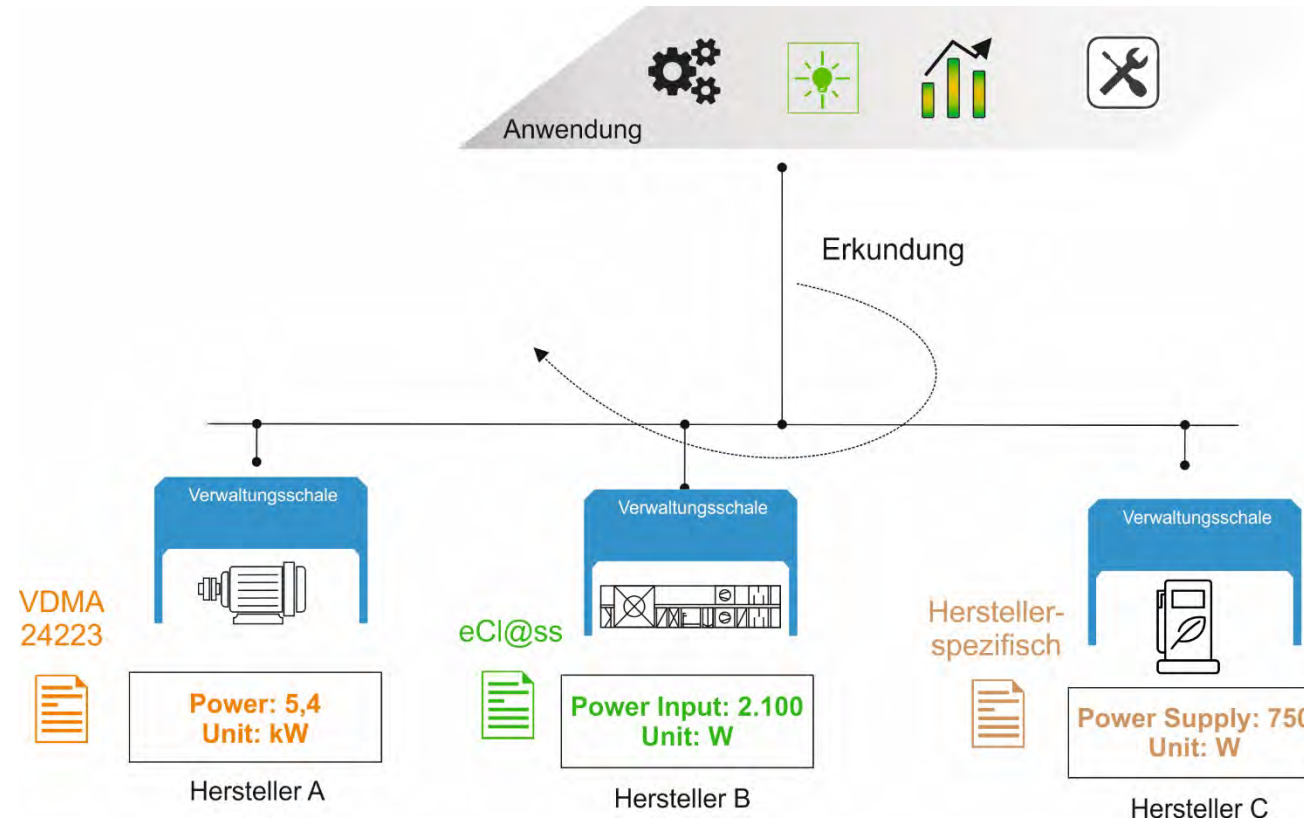

Semantische Interoperabilität durch Natural Language Processing als Basis für Self-X-Fähigkeiten von Verwaltungsschalen in semantisch heterogenen Asset-Netzwerken

AUTOMATION 2021

Verwaltungsschalen als Basis für Interoperabilität – Homogene Semantik

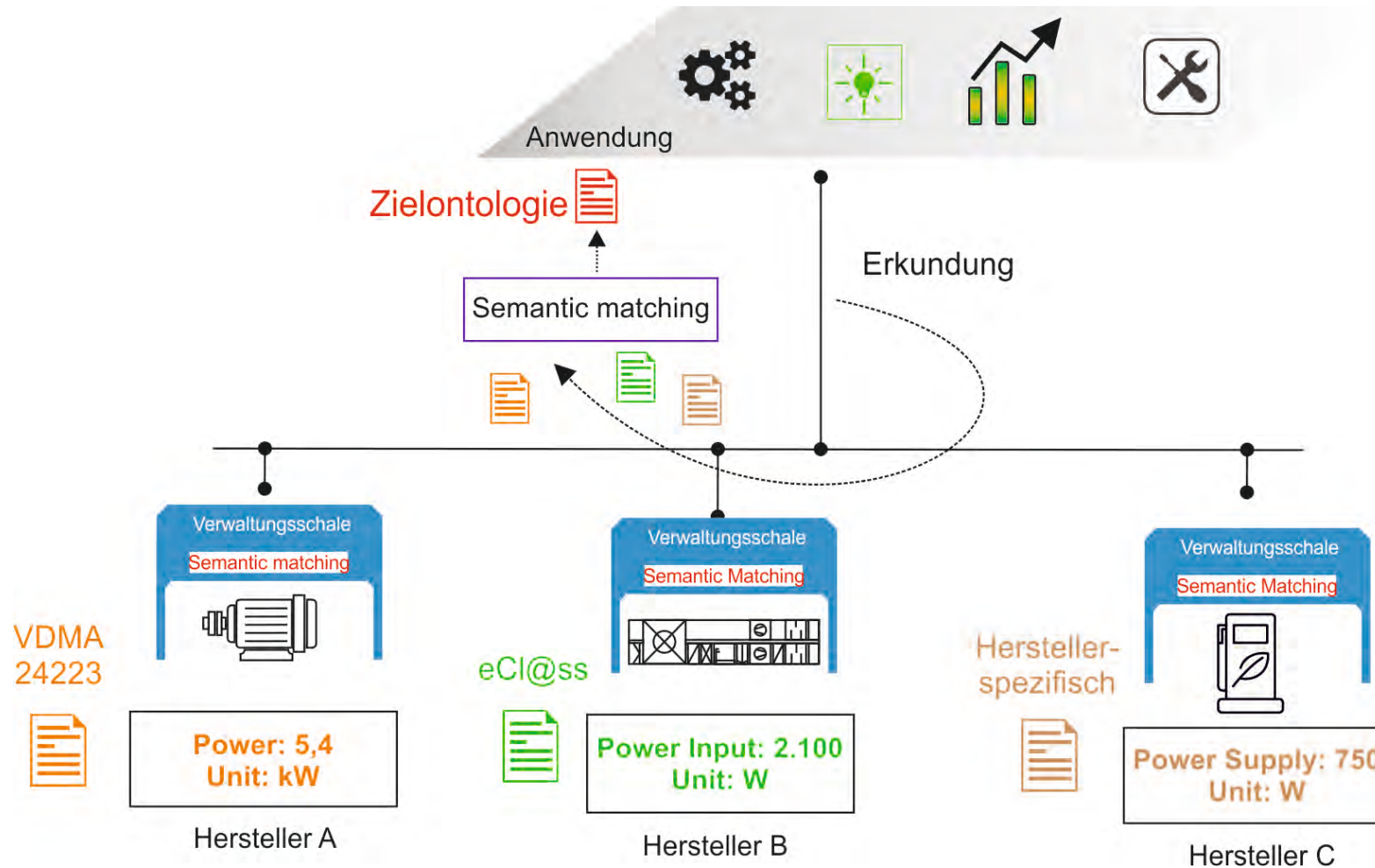


Verwaltungsschalen als Basis für Interoperabilität – Heterogene Semantik

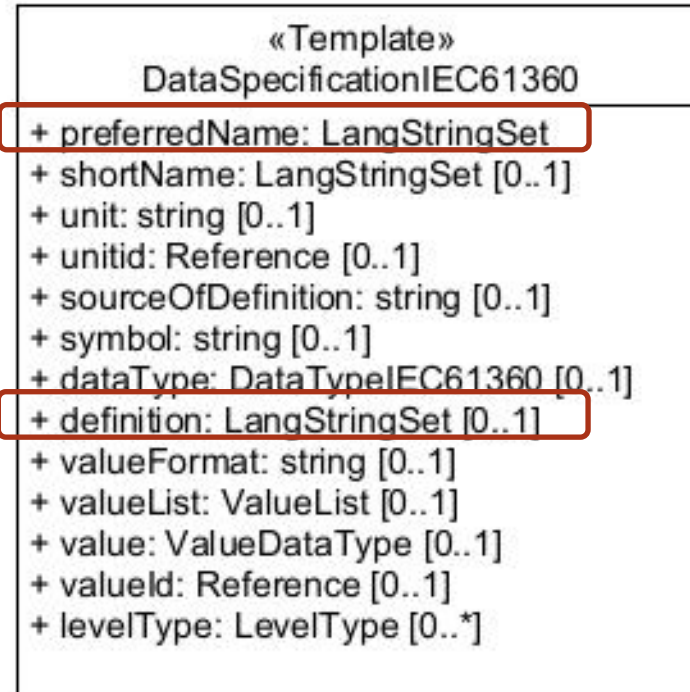


In case a concept description is not just a copy or refinement of an external standard, the provider of the AAS using this concept description shall be aware that no interoperability with other AAS can be ensured. [AAS in Detail]

Verwaltungsschalen als Basis für Interoperabilität – Heterogene Semantik mit Semantic Matching

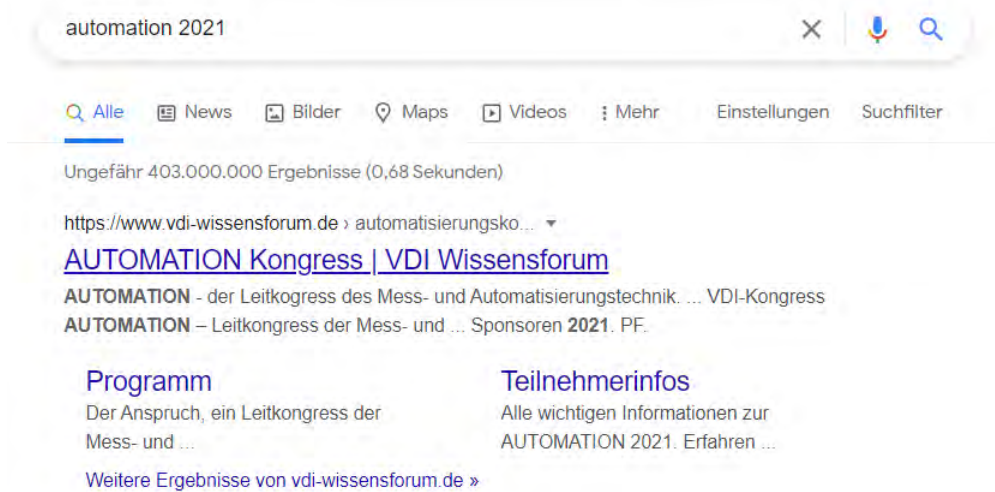


Data Specification Template IEC 61360



➤ Einsatz von Methoden aus dem Bereich Natural Language Processing

Natural Language Processing bei Google

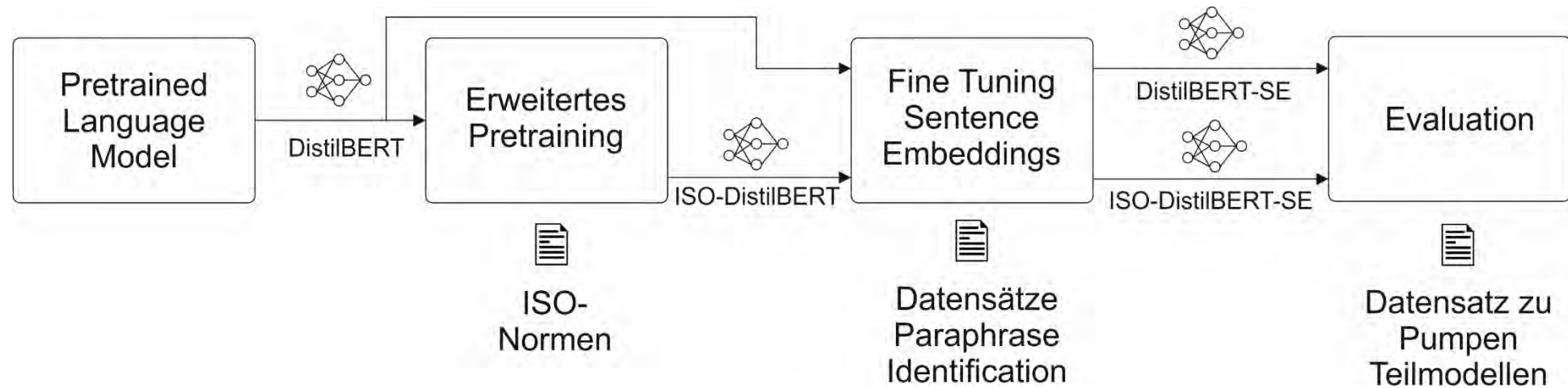


- NLP: Computer sind in der Lage nützliche Aufgaben in Zusammenhang mit menschlichen Sprache auszuführen
- Einsatz von Language Models (BERT, RoBERTa, GPT)

DistilBERT-NLI-STSB-Quora

- BERT wird bei fast jeder englischen Suchanfrage bei Google eingesetzt [6] Google Search On 2020
- Pretrained Language Model auf BooksCorpus und englischem Wikipedia
 - 3,3 Milliarden Wörter
- Trainieren von Sentence Embeddings für den Anwendungsfall Paraphrase Identification:
 - NLI/MultiNLI: Entailment, Contradiction, Neutral von Premise und Hypothesis (433.000 Paare)
 - STS: Score zwischen 1 und 5 wie ähnlich Sätze sind (8.628 Paare)
 - Quora Questions Pairs: Identifizierung ob Fragen Duplikate voneinander sind (400.000 Fragen)
- Einsatz von DistilBERT-NLI-STSB-Quora für den Anwendungsfall semantische Interoperabilität

Pretrained Language Models



- Ziel: Erweitertes Pretraining (**semantische Inhalte**) und Fine Tuning auf den drei Datensätzen (**Adaption der Aufgabe**) ermöglichen Zero-Shot Learning
- **Vorteil:** Keine Trainingsdatensätze notwendig, die Merkmale und deren Definitionen enthalten

Der Datensatz & Ergebnis

Property aus der Zielontologie		Beispiel Paraphrase	
Name	Definition	Name	Definition
Name of product	Product designation under which the device is marketed	Product title	Title under which the product is known

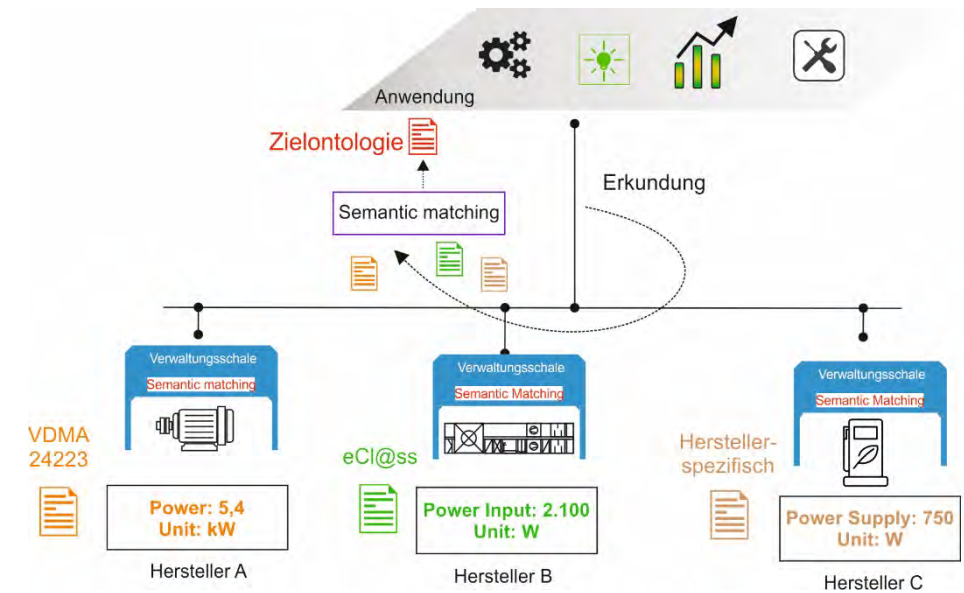
- Zielontologie: Teilmodelle für Pumpen, 250 Merkmale mit Definitionen
- Unbekannte Paraphrasen: 650 Merkmale mit Definitionen

➤ Ergebnis:

- DistilBERT-SE: 69,89 %
- ISO-DistilBERT-SE: 74,26 %

Fazit

- Erweiterung des I4.0-Forschungsansatzes zur Interaktion von semantisch homogenen auf semantisch heterogenen Verwaltungsschalen
- NLP-Methoden können eingesetzt werden um heterogene Semantik aufeinander abzubilden
- Erreichte Genauigkeit des Modells reicht jedoch noch nicht aus, um tatsächlich eingesetzt zu werden
 - Pretraining auf mehr domänenspezifischen Texten
 - Verwendung anderer Attribute der Submodel Elements der Verwaltungsschale (Unit, Datentyp etc.) um Modell zu verbessern



Vielen Dank für die Aufmerksamkeit



Technology
Arts Sciences
TH Köln

Weiteres aus dem Bereich Lehre, Forschung und Entwicklung finden Sie auf Instagram, YouTube und unserem Web-Auftritt.