LSWT 2026
3 June
14th Leipzig Semantic Web Day

UNIVERSITY OF TECHNOLOGY
IN THE EUROPEAN CAPITAL OF CULTURE
CHEMNITZ

Neuro-symbolic Consolidation of Crowdsourced Knowledge: The ORKG Use Case

Sandra Schaftner, Martin Gaedke

https://vsr.informatik.tu-chemnitz.de/research/

across
EUROPEAN UNIVERSITY
FOR CROSS-BORDER
KNOWLEDGE SHARING

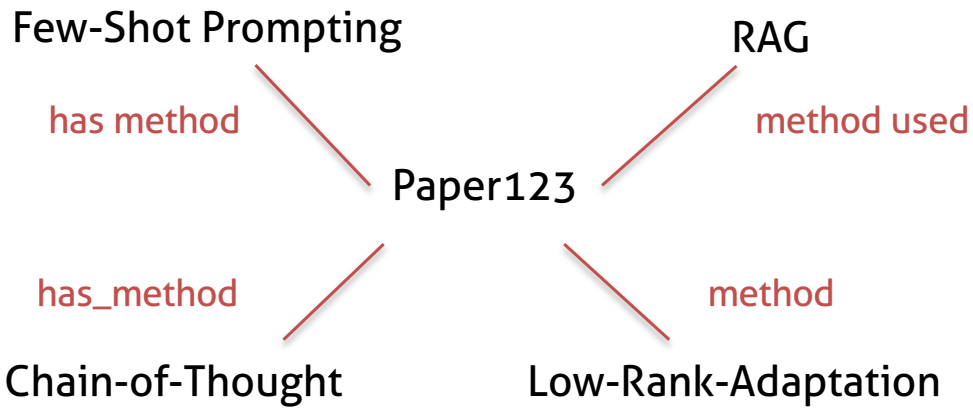
ENFIELD

Semantically FAIR Data

~~Knowledge Graphs make data FAIR~~

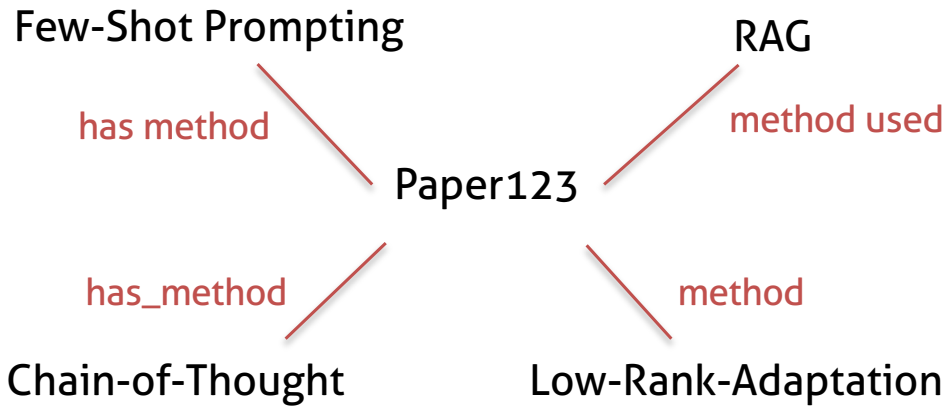
Semantically FAIR Data

~~Knowledge Graphs make data FAIR~~

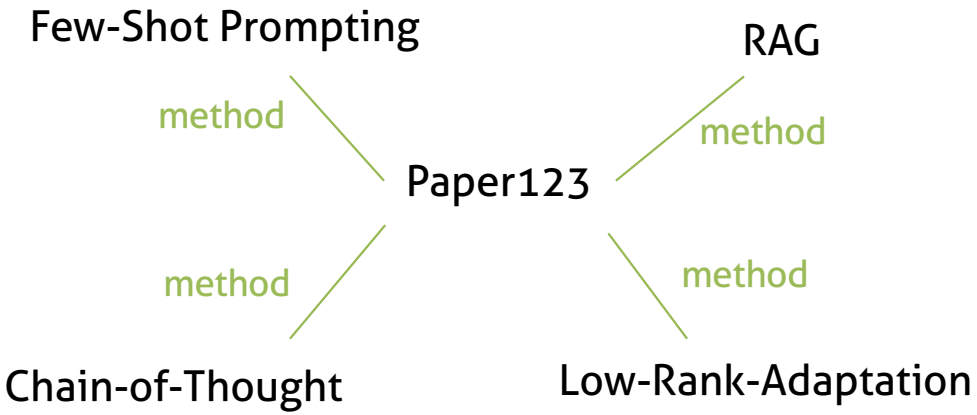


Semantically FAIR Data

~~Knowledge Graphs make data FAIR~~



Shared semantics make data FAIR



2 FAIRness Gap

Deep Semantic Scientific Knowledge Graphs

Publication metadata + Content-based information

Information

Authors
Venue
Publication date
Citations

Methods
Datasets
Benchmarks
Metrics

Deep Semantic Scientific Knowledge Graphs

Publication metadata + Content-based information

Dcterms
FaBiO
CiTO
Schema.org
Wikidata

Ontologies/
Vocabularies

?

ORKG Properties Ontology

Crowdsourcing approach

↓

> 13,000 properties

+

+ Rich source of knowledge
+ Wide adoption and integration

-

- Semantic inconsistencies
- Duplicates

ORKG Properties Ontology

Crowdsourcing approach

↓

> 13,000 properties

+

+ Rich source of knowledge
+ Wide adoption and integration

-

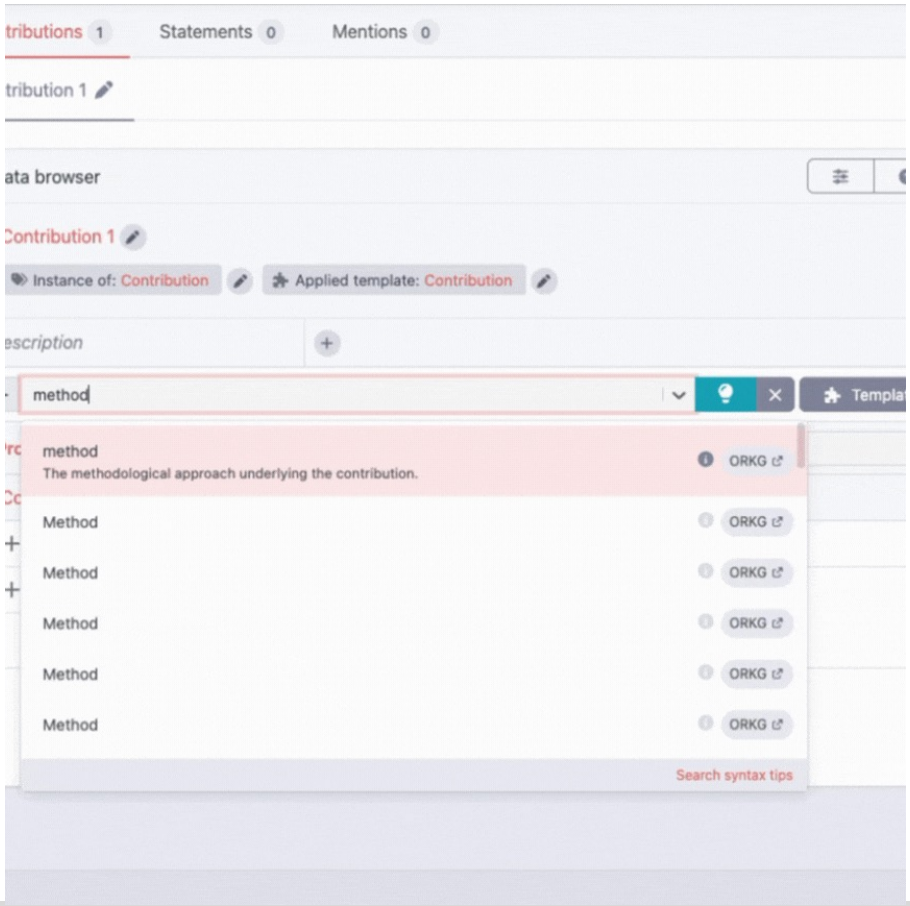
- Semantic inconsistencies
- Duplicates

↓

Neuro-symbolic consolidation

ORKG Properties Ontology

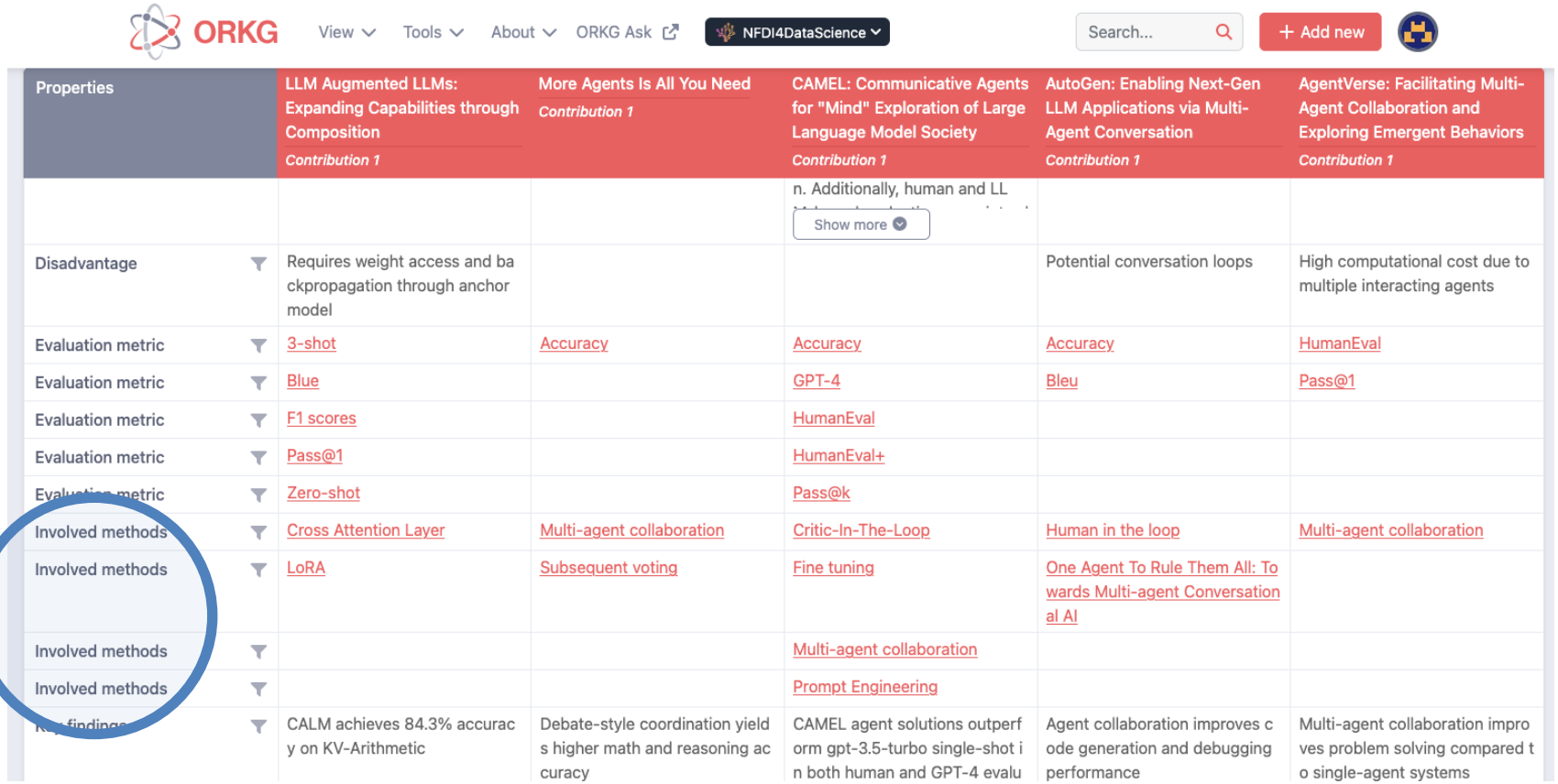
Limitations:
User Experience &
Machine Actionability



https://orkg.org/papers/create

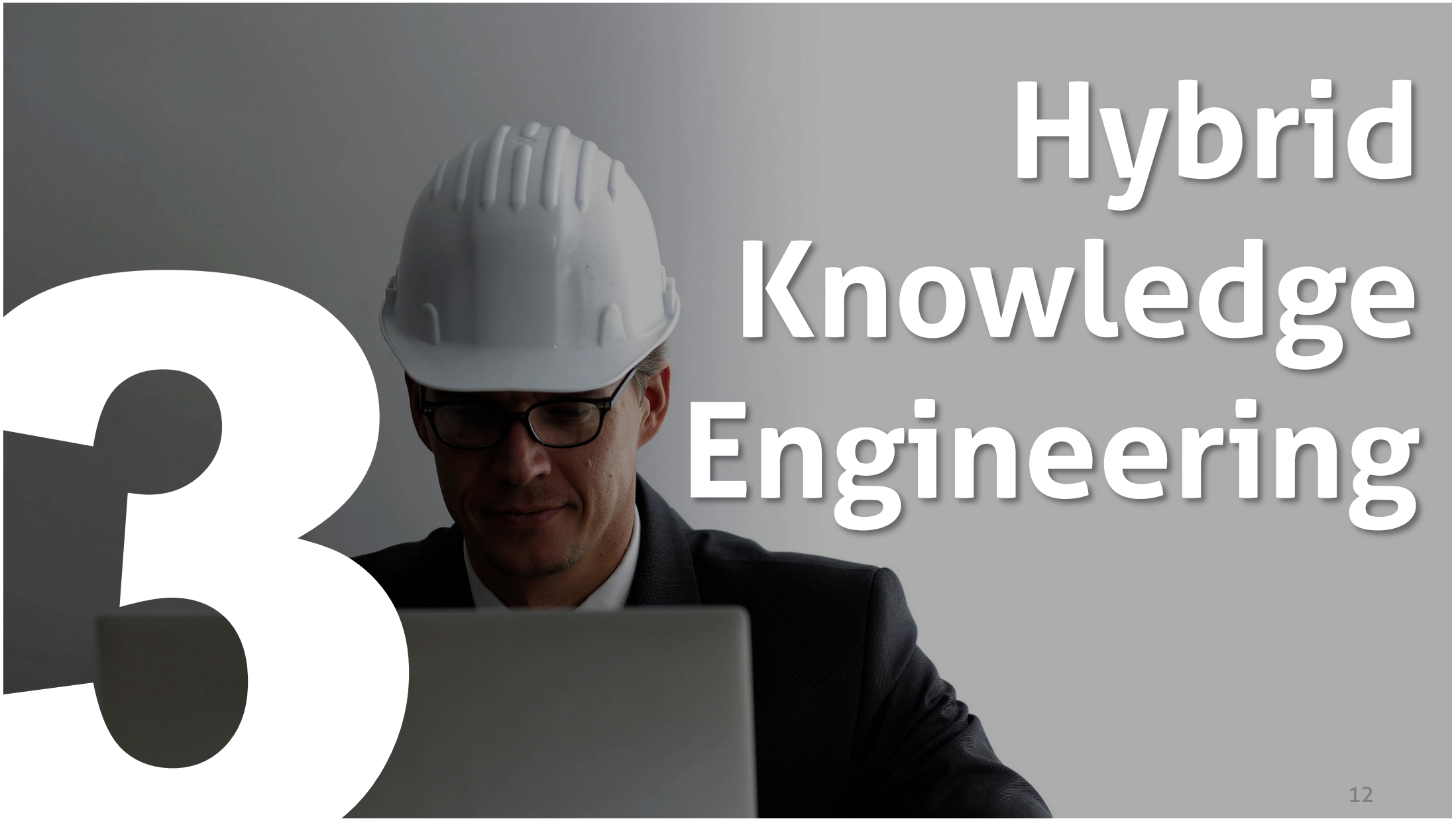
ORKG Properties Ontology

Limitations: Structured comparisons

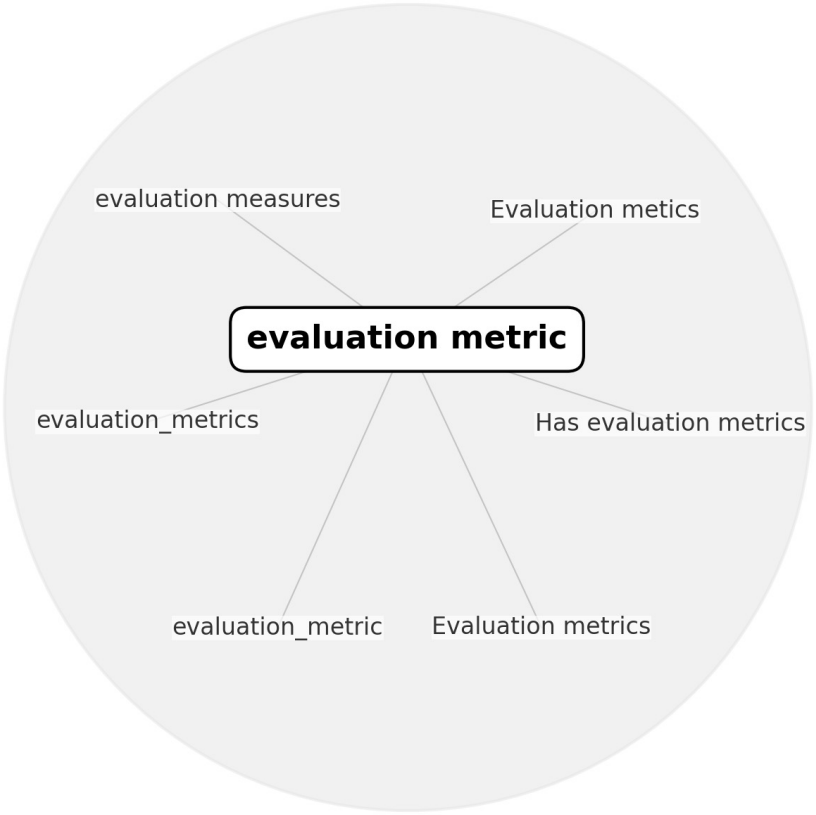


Properties	LLM Augmented LLMs: Expanding Capabilities through Composition *Contribution 1*	More Agents Is All You Need *Contribution 1*	CAMEL: Communicative Agents for "Mind" Exploration of Large Language Model Society *Contribution 1*	AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation *Contribution 1*	AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors *Contribution 1*
			n. Additionally, human and LL Show more		
Disadvantage	Requires weight access and backpropagation through anchor model			Potential conversation loops	High computational cost due to multiple interacting agents
Evaluation metric	3-shot	Accuracy	Accuracy	Accuracy	HumanEval
Evaluation metric	Blue		GPT-4	Bleu	Pass@1
Evaluation metric	F1 scores		HumanEval		
Evaluation metric	Pass@1		HumanEval+		
Evaluation metric	Zero-shot		Pass@k		
Involved methods	Cross Attention Layer	Multi-agent collaboration	Critic-In-The-Loop	Human in the loop	Multi-agent collaboration
Involved methods	LoRA	Subsequent voting	Fine tuning	One Agent To Rule Them All: Towards Multi-agent Conversational AI	
Involved methods			Multi-agent collaboration		
Involved methods			Prompt Engineering		
Key findings	CALM achieves 84.3% accuracy on KV-Arithmetic	Debate-style coordination yields higher math and reasoning accuracy	CAMEL agent solutions outperform gpt-3.5-turbo single-shot in both human and GPT-4 evalu	Agent collaboration improves code generation and debugging performance	Multi-agent collaboration improves problem solving compared to single-agent systems

https://orkg.org/comparisons/R1587807


3
Hybrid
Knowledge
Engineering

Method

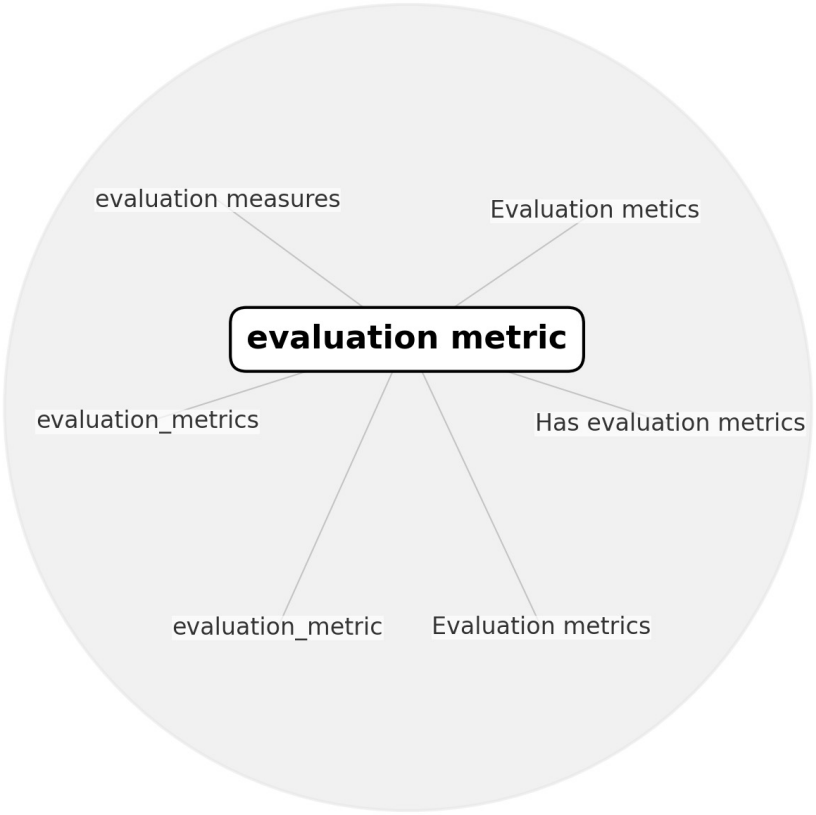


Clustering via **embedding model**

+

Cluster splitting and property selection via **LLM**

Method



Challenges

Clustering:
No universal "best" threshold size

Splitting:
Adequate few-shot examples for the LLM

Method

Few-shot examples for LLM cluster splitting decisions:

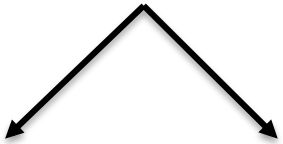
```
Cluster: ["Conclusion", "Summary", "has conclusion", "is conclusive"]
Decision: {"decision": "KEEP"}  <-- Synonyms

Cluster: ["published by", "published in", "has published"]
Decision: {"decision": "BREAK"} <-- Distinct semantic roles (Agent vs. Venue)

Cluster: ["evaluation method", "evaluation metric", "evaluation results"]
Decision: {"decision": "BREAK"} <-- Distinct research phases

Cluster: ["ontology based", "Ontology ID", "Ontology URL"]
Decision: {"decision": "BREAK"} <-- Attribute vs. Identifier vs. Locator
```

Challenges



Clustering:
No universal "best" threshold size

Splitting:
Adequate few-shot examples for the LLM

4 Significant Consolidation

Evaluation

Total count of properties:

ORKG Properties:

13,009 properties

OPO-Consolidated:

2865 properties

Evaluation

Ambiguity measurement:

ORKG Properties:

47 results

OPO-Consolidated:

8 results

Evaluation method: Average number of candidate properties returned for a specific property from the ORKG Gold Standard Dataset[3]

3 Vladyslav Nechakhin et al. 2024. ORKG Properties and LLM-Generated Research Dimensions Evaluation Dataset. doi:10.25835/6OYN9D1N

Evaluation

Backward compatibility:

100 %

```
orkg:P175087 a rdf:Property ;
    rdfs:label "research_method" ;
    owl:equivalentProperty orkg:P15068 ;
    skos:exactMatch orkg:P15068 .

orkg:P15068 a rdf:Property ;
    rdfs:label "research method"@en ;
    dcterms:created "2020-03-30T20:57:44.811000+02:00"^^xsd:dateTime ;
    dcterms:identifier "P15068" ;
    rdfs:comment ""@en ;
    rdfs:isDefinedBy <https://w3id.org/orkg-properties-ontology-consolidated> ;
    skos:prefLabel "research method"@en .
```

https://w3id.org/orkg-properties-ontology-consolidated

Code, Paper & Ontology

https://sandraschaftner.github.io

Discussion & Future Work

Further crowdsourced ontologies to be consolidated?

Ideas for improving the consolidation approach?

Contact & Acknowledgment

sandra.schaftner@
informatik.tu-chemnitz.de

TECHNISCHE UNIVERSITÄT
IN DER KULTURHAUPTSTADT EUROPAS
CHEMNITZ

VSR

across
EUROPEAN UNIVERSITY
FOR CROSS-BORDER
KNOWLEDGE SHARING

ENFIELD

Co-funded by
the European Union