HEINZ NIXDORF INSTITUT
UNIVERSITÄT PADERBORN

No Need to Be a Know-It-All: Fact Checking with Shallow Knowledge

Umair Qudus, Neha Pokharel, Michael Röder, and Axel-Cyrille Ngonga Ngomo

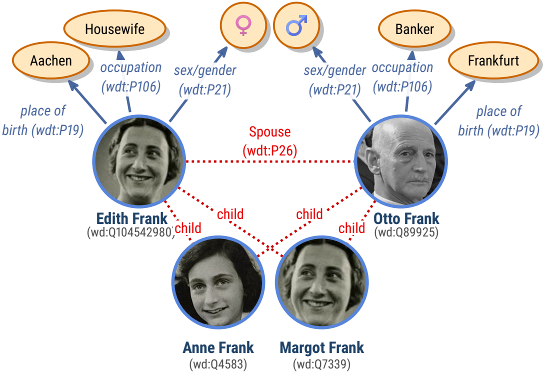
DICE

Data Science Group
Paderborn University

May 13, 2026

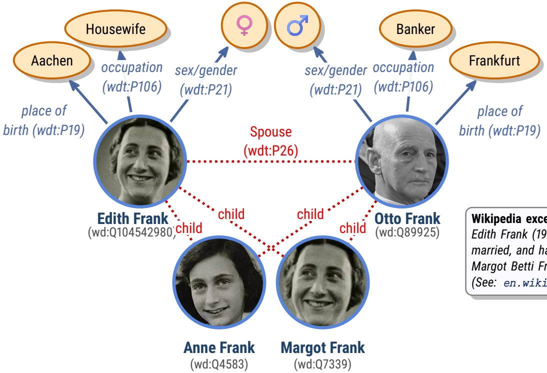
Introduction

Motivation



Introduction

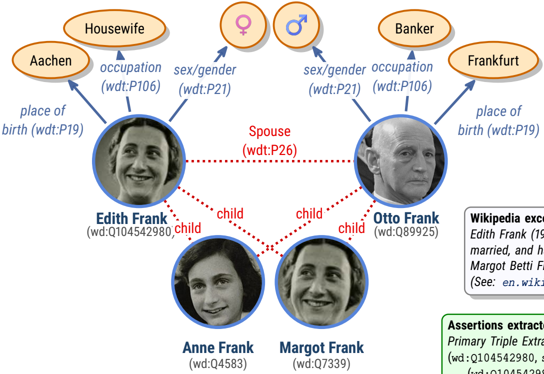
Motivation



Wikipedia excerpt:
Edith Frank (1900–1945) and Otto Frank (1889–1980) were married, and had two daughters, Anne Frank (1929–1945) and Margot Betti Frank (1926–1945).
(See: `en.wikipedia.org/wiki/Edith_Frank`)

Introduction

Motivation



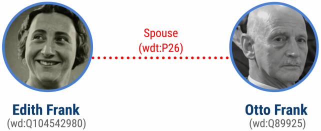
Wikipedia excerpt:
Edith Frank (1900–1945) and Otto Frank (1889–1980) were married, and had two daughters, Anne Frank (1929–1945) and Margot Betti Frank (1926–1945).
(See: `en.wikipedia.org/wiki/Edith_Frank`)

Assertions extracted by ShallKnow:
Primary Triple Extraction:
(wd:Q104542980, spouse (wdt:P26), wd:Q89925)
(wd:Q104542980, child (wdt:P40), wd:Q4583)
(wd:Q104542980, child (wdt:P40), wd:Q7339)
(wd:Q89925, child (wdt:P40), wd:Q4583)
(wd:Q89925, child (wdt:P40), wd:Q7339)

Secondary Triple Extraction:
(only triggered if the previous step extracted less than θ assertions)
(wd:Q4583, :P-born-in, wd:Q1794)
(wd:Q104542980, :P-married-to, wd:Q89925)...

Introduction

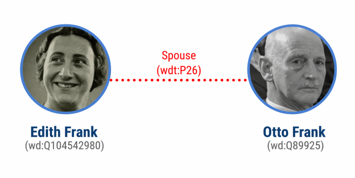
Families of approaches & limitations



- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Introduction

Families of approaches & limitations

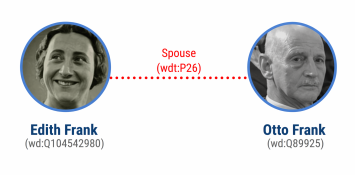


- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Families	Key Limitation

Introduction

Families of approaches & limitations

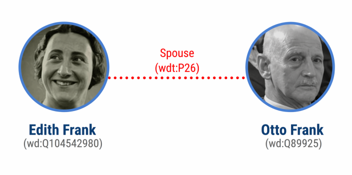


- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Families		Key Limitation
1. Text-based	$\longrightarrow$	Manual feature engineering

Introduction

Families of approaches & limitations

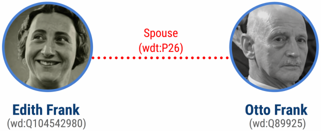


- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Families		Key Limitation
1. Text-based	$\longrightarrow$	Manual feature engineering
2. Graph-based	$\longrightarrow$	KG incompleteness & missing evidential paths

Introduction

Families of approaches & limitations

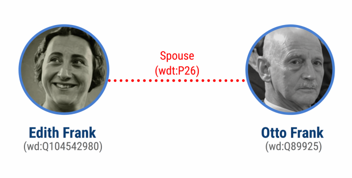


- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Families		**Key Limitation**
1. Text-based	$\longrightarrow$	Manual feature engineering
2. Graph-based	$\longrightarrow$	KG incompleteness & missing evidential paths
3. Embedding-based	$\longrightarrow$	Complex training & lack of interpretability

Introduction

Families of approaches & limitations

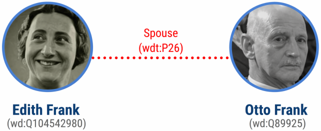


- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Families		Key Limitation
1. Text-based	$\longrightarrow$	Manual feature engineering
2. Graph-based	$\longrightarrow$	KG incompleteness & missing evidential paths
3. Embedding-based	$\longrightarrow$	Complex training & lack of interpretability
4. Hybrid	$\longrightarrow$	High system complexity

Introduction

Families of approaches & limitations



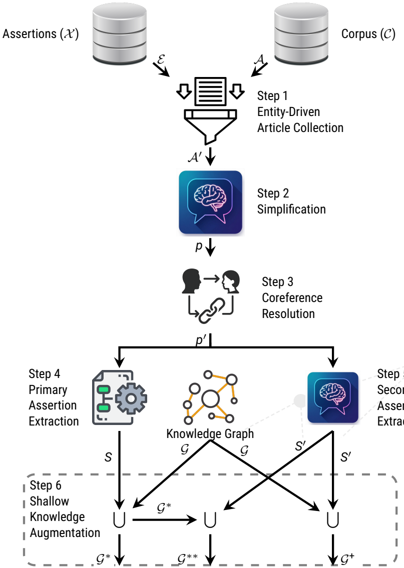
- Given: Knowledge graph $\mathbb{G}$ and assertion $t = (s, p, o)$
- Goal: Compute $P(t)$

Families		Key Limitation
2. Graph-based	$\longrightarrow$	KG incompleteness & missing evidential paths

Bridging the Gap

ShallKnowFramework

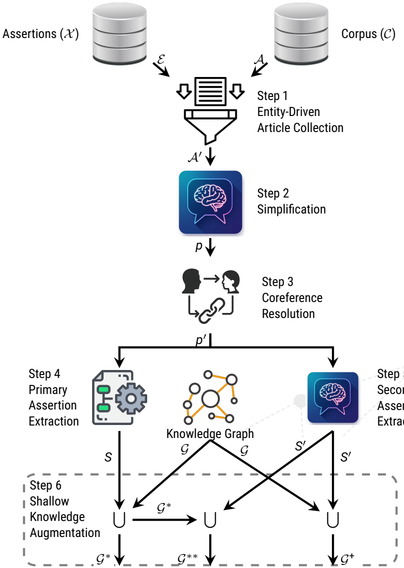
- **Goal:**
  Fix KG incompleteness for path-based fact-checking



Bridging the Gap

SHALLKNOWFramework

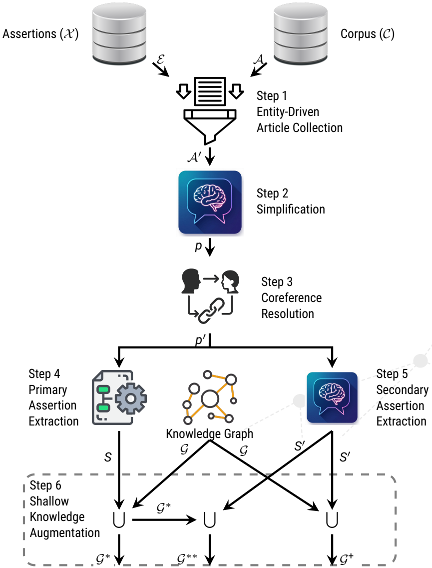
- **Goal:**
  Fix KG incompleteness for path-based fact-checking
- **The Solution: SHALLKNOW**
  Augments KGs with **Shallow Knowledge**



Bridging the Gap

ShallKnowFramework

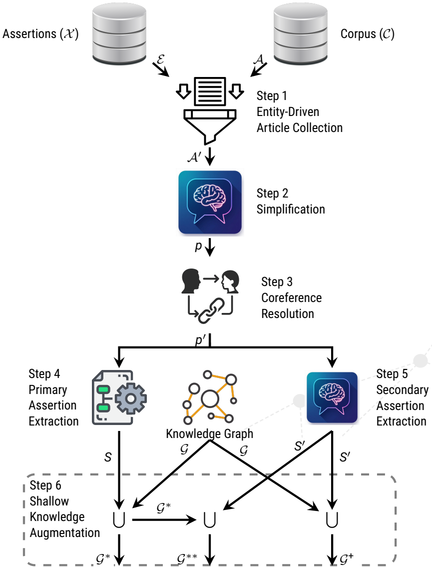
- **Goal:**
  Fix KG incompleteness for path-based fact-checking
- **The Solution: ShallKnow**
  Augments KGs with **Shallow Knowledge**
- **Shallow Knowledge**
  - RDF assertions mined from raw text.
  - Fills missing paths; no strict ontology needed.



Bridging the Gap

SHALLKNOWFramework

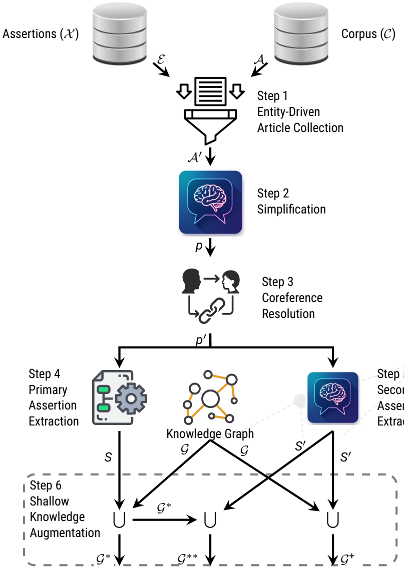
- **Goal:**
  Fix KG incompleteness for path-based fact-checking
- **The Solution: SHALLKNOW**
  Augments KGs with **Shallow Knowledge**
- **Shallow Knowledge**
  - RDF assertions mined from raw text.
  - Fills missing paths; no strict ontology needed.
- **Outcome**
  Boost in graph completeness



Methodology

Contextualizing Entities (Steps 1–3)

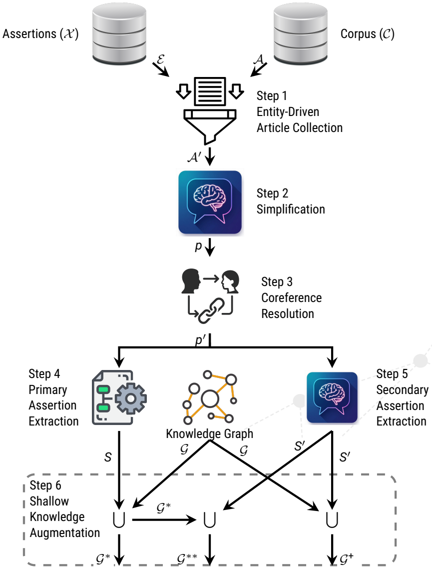
- Step 1: Article Collection
  - Retrieve reference corpus articles ($\mathcal{A}$)
  - Add 1-hop linked articles ($\mathcal{A}'$)



Methodology

Contextualizing Entities (Steps 1–3)

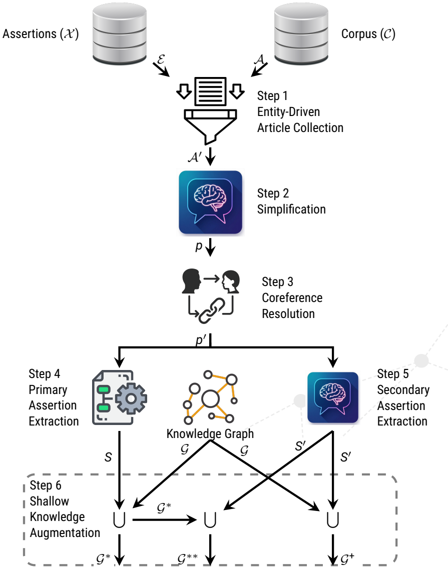
- **Step 1: Article Collection**
  - Retrieve reference corpus articles ($\mathcal{A}$)
  - Add 1-hop linked articles ($\mathcal{A}'$)
- **Step 2: Simplification**
  - LLM removes markups and noise
  - Generates summary paragraph (p)



Methodology

Contextualizing Entities (Steps 1–3)

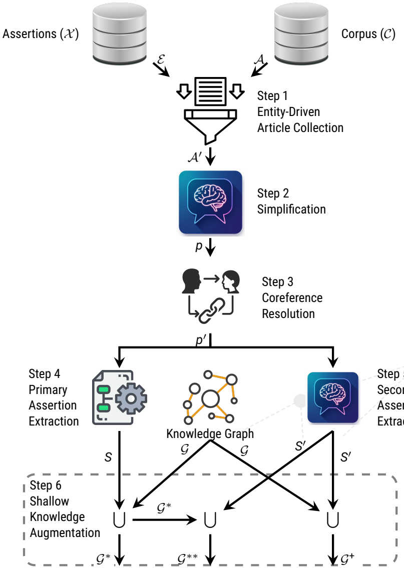
- Step 1: Article Collection
  - Retrieve reference corpus articles ($\mathcal{A}$)
  - Add 1-hop linked articles ($\mathcal{A}'$)
- Step 2: Simplification
  - LLM removes markups and noise
  - Generates summary paragraph (p)
- Step 3: Coreference Resolution
  - Replace pronouns with entity names
  - **Result:** clean snippet (p')



Methodology

The Two-Tier Extraction Engine

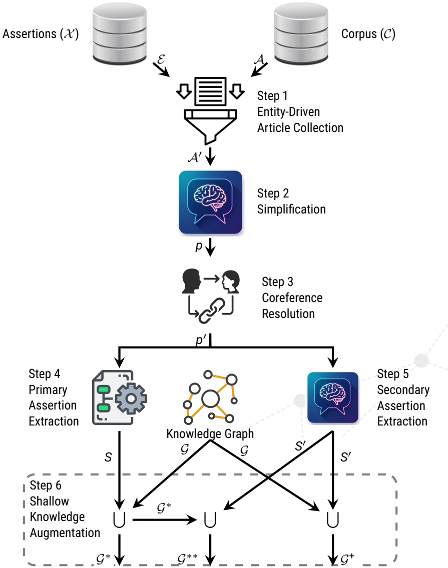
- Step 4: Primary Extraction
  - RE tools extraction
  - KG properties ($P_{\mathcal{G}}$)
  - Default execution
  - Fast, structured, schema-aligned extraction



Methodology

The Two-Tier Extraction Engine

- **Step 4: Primary Extraction**
  - RE tools extraction
  - KG properties ($P_{\mathcal{G}}$)
  - Default execution
  - Fast, structured, schema-aligned extraction
- **Step 5: Secondary Extraction**
  - LLM-based extraction
  - Novel properties ($P_{\mathcal{C}}$)
  - Trigger if: $|\mathcal{S}| < \theta$
  - Improves recall and fill missing gaps

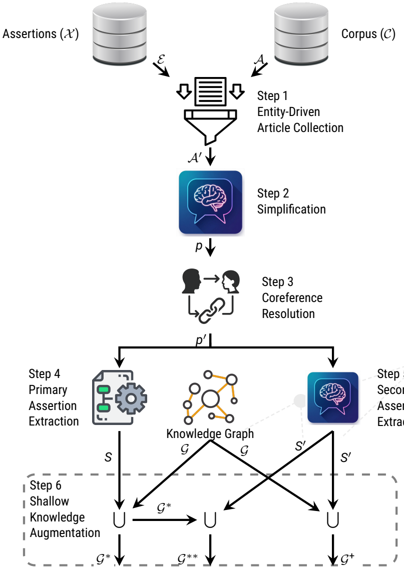


Methodology

The Two-Tier Extraction Engine

- Step 4: Primary Extraction
  - RE tools extraction
  - KG properties ($P_{\mathcal{G}}$)
  - Default execution
  - Fast, structured, schema-aligned extraction
- Step 5: Secondary Extraction
  - LLM-based extraction
  - Novel properties (P_{C})
  - Trigger if: $|\mathcal{S}| < \theta$
  - Improves recall and fill missing gaps

This threshold-based mechanism (θ) balances extraction depth with computational efficiency.



Methodology

Shallow Knowledge Augmentation Strategies

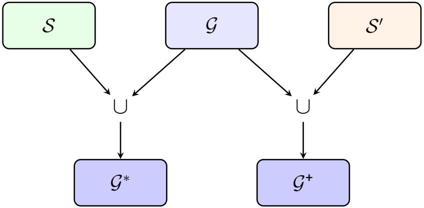
- **Given:**
  - $\mathcal{G}$: Original KG
  - $\mathcal{S}$: Primary Assertions *(existing schema)*
  - $\mathcal{S}'$: Secondary Assertions *(novel properties)*

$\mathcal{S}$	$\mathcal{G}$	$\mathcal{S}'$

Methodology

Shallow Knowledge Augmentation Strategies

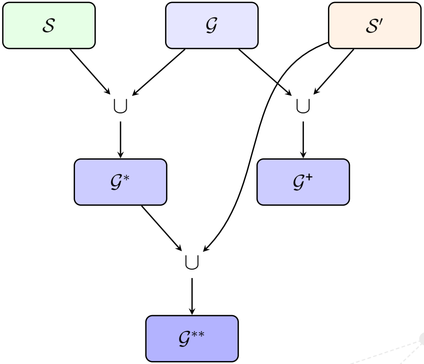
- **Given:**
  - $\mathcal{G}$: Original KG
  - $\mathcal{S}$: Primary Assertions *(existing schema)*
  - $\mathcal{S}'$: Secondary Assertions *(novel properties)*
- **The Augmented KGs:**
  - $\mathcal{G}^*$: Enriches existing links
  - $\mathcal{G}^+$: Introduces novel paths
  - $\mathcal{G}^{**}$: The full augmentation



Methodology

Shallow Knowledge Augmentation Strategies

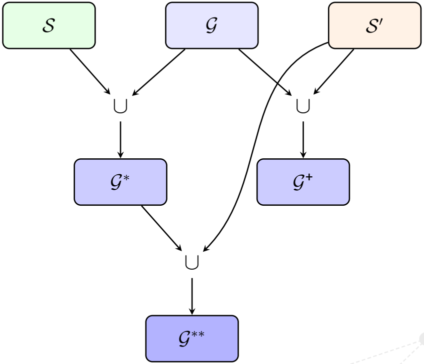
- **Given:**
  - $\mathcal{G}$: Original KG
  - $\mathcal{S}$: Primary Assertions *(existing schema)*
  - $\mathcal{S}'$: Secondary Assertions *(novel properties)*
- **The Augmented KGs:**
  - $\mathcal{G}^*$: Enriches existing links
  - $\mathcal{G}^+$: Introduces novel paths
  - $\mathcal{G}^{**}$: The full augmentation



Methodology

Shallow Knowledge Augmentation Strategies

- **Given:**
  - $\mathcal{G}$: Original KG
  - $\mathcal{S}$: Primary Assertions *(existing schema)*
  - $\mathcal{S}'$: Secondary Assertions *(novel properties)*
- **The Augmented KGs:**
  - $\mathcal{G}^*$: Enriches existing links
  - $\mathcal{G}^+$: Introduces novel paths
  - $\mathcal{G}^{**}$: The full augmentation
- **Outcome:** Enriched graphs for downstream path-based fact-checking task.



Experiments

Setup details

- Benchmark Datasets (size in #assertions)

Dataset	Train (T/F/Total)	Test (T/F/Total)	DP
BPDP 22 [3]	100/100/200	103/103/206	2
FactBench Mix 22 [9]	633/486/1119	637/492/1129	9
FAVEL-DS [7]	380/385/765	163/164/327	11

Experiments

Setup details

- Benchmark Datasets (size in #assertions)

Dataset	Train (T/F/Total)	Test (T/F/Total)	DP
BPDP 22 [3]	100/100/200	103/103/206	2
FactBench Mix 22 [9]	633/486/1119	637/492/1129	9
FAVEL-DS [7]	380/385/765	163/164/327	11

- Reference Corpus for Extraction: Plain text snippets from 6.7 million English Wikipedia articles ($\approx$ 300M sentences)

Experiments

Setup details

- Benchmark Datasets (size in #assertions)

Dataset	Train (T/F/Total)	Test (T/F/Total)	DP
BPDP 22 [3]	100/100/200	103/103/206	2
FactBench Mix 22 [9]	633/486/1119	637/492/1129	9
FAVEL-DS [7]	380/385/765	163/164/327	11

- Reference Corpus for Extraction: Plain text snippets from 6.7 million English Wikipedia articles ($\approx$ 300M sentences)
- Evaluated Approaches: 10 state-of-the-art path-based fact-checking approaches

Experiments

Setup details

- Benchmark Datasets (size in #assertions)

Dataset	Train (T/F/Total)	Test (T/F/Total)	DP
BPDP 22 [3]	100/100/200	103/103/206	2
FactBench Mix 22 [9]	633/486/1119	637/492/1129	9
FAVEL-DS [7]	380/385/765	163/164/327	11

- Reference Corpus for Extraction: Plain text snippets from 6.7 million English Wikipedia articles ($\approx$ 300M sentences)
- Evaluated Approaches: 10 state-of-the-art path-based fact-checking approaches
- Tools:
  - DeepSeek-r1-14B (open-source LLM) for text simplification and assertion generation
  - REBEL and Relik as assertion extraction tools
  - spaCy for Named Entity Recognition (NER)

Experiments

Setup details

- Benchmark Datasets (size in #assertions)

Dataset	Train (T/F/Total)	Test (T/F/Total)	DP
BPDP 22 [3]	100/100/200	103/103/206	2
FactBench Mix 22 [9]	633/486/1119	637/492/1129	9
FAVEL-DS [7]	380/385/765	163/164/327	11

- Reference Corpus for Extraction: Plain text snippets from 6.7 million English Wikipedia articles ($\approx$ 300M sentences)
- Evaluated Approaches: 10 state-of-the-art path-based fact-checking approaches
- Tools:
  - DeepSeek-r1-14B (open-source LLM) for text simplification and assertion generation
  - REBEL and Relik as assertion extraction tools
  - spaCy for Named Entity Recognition (NER)
- Evaluation Metric: AUC-ROC

Results (Part 1): BPDP & FactBench

DS	Approach	w/o	$\mathcal{G}^*$	Δ	$\mathcal{G}^+$	Δ	$\mathcal{G}^{**}$	Δ
BPDP 22	FAVEL [7]	0.6701	0.6822	†+0.0121	0.7091	†**+0.0390**	0.7232	†+0.0410
	COPAAL [10]	0.5314	0.5497	†+0.0183	0.5370	†+0.0190	0.5515	†+0.0201
	PredPath [8]	0.6311	0.6522	†**+0.0211**	0.6680	†+0.0369	0.6748	†**+0.0437**
	Pathent [11]	0.4985	0.4985	0.0000	0.4988	†+0.0005	0.4990	†+0.0005
	DP [2]	0.5000	0.5000	0.0000	0.5000	0.0000	0.5000	0.0000
	PRA [5]	0.5801	0.5905	†+0.0104	0.6155	†+0.0354	0.6212	†+0.0411
	Jaccard [6]	0.5009	0.5011	†+0.0002	0.5015	†+0.0006	0.5016	†+0.0007
	Adamic [1]	0.4983	0.4983	0.0000	0.4983	0.0000	0.4983	0.0000
	SimRank [4]	0.4817	0.4856	†+0.0039	0.4920	†+0.0087	0.4945	†+0.0100
FactBench Mix 22	FAVEL [7]	0.8133	0.8656	†+0.0223	0.9052	†+0.0919	0.9121	†+0.0988
	COPAAL [10]	0.7903	0.8404	†**+0.0501**	0.8960	†**+0.1057**	0.9011	†**+0.1108**
	PredPath [8]	0.7355	0.7425	+0.0070	0.7766	†+0.0411	0.7895	†+0.0540
	Pathent	0.7054	0.7035	†+0.0019	0.7130	†+0.0076	0.7218	†+0.0164
	DP	0.4467	0.4569	†+0.0102	0.4575	†+0.0108	0.4607	†+0.0140
	PRA	0.5704	0.6033	†+0.0329	0.6170	†+0.0466	0.6221	†+0.0517
	Jaccard	0.6271	0.6344	†+0.0073	0.6537	†+0.0266	0.6621	†+0.0350
	Adamic Adar	0.6591	0.6793	†+0.0202	0.6980	†+0.0389	0.7328	†+0.0737
	SimRank	0.5881	0.6104	†+0.0223	0.6525	†+0.0644	0.6622	†+0.0741

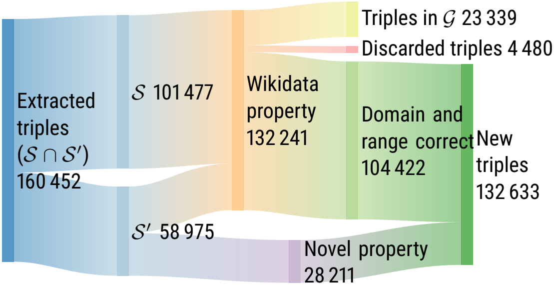
†We use a Wilcoxon signed rank test with a significance threshold α = 0.05.

Results (Part 2): FAVEL-DS

DS	Approach	w/o	$\mathcal{G}^*$	Δ	$\mathcal{G}^+$	Δ	$\mathcal{G}^{**}$	Δ
FAVEL-DS	FAVEL	0.7231	0.7532	†+0.0301	0.7831	†+0.0600	0.7921	†+0.0690
	COPAAL	0.5102	0.7222	†**+0.2120**	0.7400	†**+0.2298**	0.7493	†**+0.2391**
	PredPath	0.7189	0.7233	+0.0044	0.7482	†+0.0293	0.7623	†+0.0434
	Pathent	0.5733	0.6221	†+0.0588	0.7400	†+0.1667	0.7533	†+0.1800
	DP	0.5247	0.5247	0.0000	0.5350	†+0.0103	0.5379	†+0.0132
	PRA	0.5322	0.5466	†+0.0144	0.5530	†+0.0208	0.5567	†+0.0245
	Jaccard	0.5398	0.5554	†+0.0156	0.5840	†+0.0442	0.5958	†+0.0560
	Adamic	0.5467	0.5865	†+0.0398	0.6380	†+0.0913	0.6747	†+0.1280
	SimRank	0.5603	0.5778	+0.0175	0.6023	†+0.0420	0.6085	†+0.0428

†We use a Wilcoxon signed rank test with a significance threshold α = 0.05.

Quality Control: Can We Trust Shallow Assertions?



High inter-annotator agreement confirms the reliability of extracted assertions (up to 89% precision and κ = 0.8790).

Conclusion and Future Work

- Summary of Contributions
  - Addressed the core limitation of path-based fact-checking: the reliance on KG completeness.
  - Introduced **SHALLKNOW**, a model-agnostic framework that systematically augments incomplete KGs using shallow knowledge from unstructured text.
  - Demonstrated consistent and significant AUROC improvements across 3 benchmark datasets for state-of-the-art path-based methods.

Conclusion and Future Work

- Summary of Contributions
  - Addressed the core limitation of path-based fact-checking: the reliance on KG completeness.
  - Introduced **ShallKnow**, a model-agnostic framework that systematically augments incomplete KGs using shallow knowledge from unstructured text.
  - Demonstrated consistent and significant AUROC improvements across 3 benchmark datasets for state-of-the-art path-based methods.

- Future Directions
  - **Scalability:** Improve computational efficiency and scalability of the assertion extraction process.
  - **Refinement:** Normalize semantically equivalent properties and better manage rare relations.

Summary

That's all folks!

SHALLKNOW boosts path-based fact-checking by augmenting incomplete knowledge graphs with shallow knowledge from text.

Web: dice-research.org

Code: github.com/dice-group/shallknow

Twitter: @DiceResearch

The work has been supported by the German Federal Ministry of Education and Research (BMBF) within the project NEBULA (13N16364), the Ministry of Culture and Science of NRW (MKW NRW) within the project SAIL (NW21-059D), and the German Federal Ministry of Research, Technology and Space (BMFTR) within the project KI-Akademie OWL(16IS24057B) and Learn2RAG (01MK250104).

NEBULA
Nutzerzentrierte KI-basierte Erkennung von Fake News und Fehlinformationen

SAIL
SUSTAINABLE LIFE-CYCLE OF INTELLIGENT SOCIO-TECHNICAL SYSTEMS

KI AKADEMIE OWL

LEARN2RAG

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Bundesministerium für Forschung, Technologie und Raumfahrt

Ministry of Culture and Science of the State of North Rhine-Westphalia

References

[1] Adamic, L.A., Adar, E.: Friends and neighbors on the web. Social Networks **25**(3), 211–230 (2003). https://doi.org/https://doi.org/10.1016/S0378-8733(03)00009-1

[2] Ciampaglia, G.L., Shiralkar, P., Rocha, L.M., Bollen, J., Menczer, F., Flammini, A.: Computational fact checking from knowledge networks. PloS one **10**(6), e0128193 (2015)

[3] Gerber, D., Esteves, D., Lehmann, J., Bühmann, L., Usbeck, R., Ngonga Ngomo, A.C., Speck, R.: Defacto-temporal and multilingual deep fact validation. Web Semant. **35**(P2), 85–101 (Dec 2015). https://doi.org/10.1016/j.websem.2015.08.001, https://doi.org/10.1016/j.websem.2015.08.001

[4] Jeh, G., Widom, J.: Simrank: A measure of structural-context similarity. In: Proceedings of the Eighth ACM SIGKDD Intern. Conf. on Knowledge Discovery and Data Mining (2002)

[5] Lao, N., Cohen, W.: Relational retrieval using a combination of path-constrained random walks. Machine Learning **81**, 53–67 (10

Expert Inter-Annotator Agreement verified factual accuracy.

89% Precision

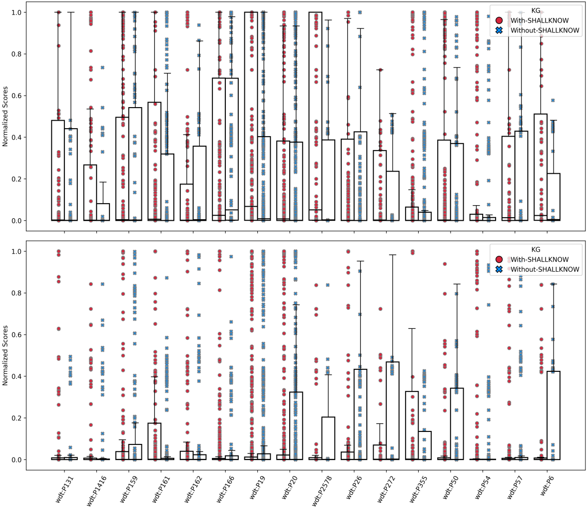
for novel predicates

86% Precision

for schema-aligned predicates

More Results

Predicate-wise Score Distribution



Predicate-wise predicted scores for **true** (top, higher is better) and **false** (bottom, lower is better) assertions with and without SHALLKNOW.