DataMind at AKBC Shared Task 2026: Relation-Specific Multiple Prompting for Closed-Book Knowledge Base Completion

Tanaree Patcharatrakul[1] **Peerapon Vateekul**[2]*

[1]International School of Engineering, Faculty of Engineering, Chulalongkorn University
[2]Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University

6738108421@student.chula.ac.th **peerapon.v@chula.ac.th** *Corresponding author

Acknowledgments

We thank Dr. Piyawat Lertvittayakumjorn for his advice throughout this work.

TASK

List every correct object for a (subject, relation) pair from model parametric memory, using no retrieval, fine-tuning, and under a 32B parameter limit.

3 Relation Types

Numeric `hasArea` `hasCapacity`

Single-object `personHasCityOfDeath`

Set-valued `countryLandBordersCountry` `companyTradesAtStockExchange` `awardWonBy`

METHODOLOGY

We used **Qwen3.6-27B** and **Gemma-3-4B-it**. All relations use Qwen3.6-27B. For hasArea, Gemma-3-4B-it is also used to identify the prompt language.

(subject, relation) → **Relation-specific prompts**: Few-shot prompts written for each relation, often in several variants → **Self-consistency**: For numeric relations, use 5 samples per prompt, rising to 9 when they disagree → **Cross-prompt ensemble**: Several prompt variants predict independently and use agreement-based ensemble → **Final answer**

Error-driven refinement: In some prompts the model reviews the errors and rewrites the prompts (labeled as "meta_")

Numeric

hasArea

Nested ensemble of 6 runs. For each run, confidence-escalated self-consistency (5→9) is used.

Run 1 | Run 2 | Run 3 | Run 4 | Run 5 | Run 6* → Output as the cluster-median of 6 runs

***Run 6: native-language prompt**

Qwen names the local language & confidence → Confidence < 0.7? Ask Gemma-3-4B instead → Use the translated prompt only if ≥ 0.7

base prompt (runs 1–5)

```
"What is the total area of {subject}
in square kilometers? For countries
use total area (land + inland water)."
```

hasCapacity

Single prompt with confidence-escalated self-consistency (5→9).

prompt

```
"…what type and level of venue is it? …
Estimate capacity using regional norms
for that venue type."
```

Single-object

personHasCityOfDeath

Four prompt variants run independently; a city is output only if all four return the same city.

recent_aware
Considers whether the person may have died recently before naming a city

```
"If {subject} was prominent in the
1980s-2010s and was likely in their 70s-
90s, they may have passed away recently"
```

recent_precise
The same recency reasoning plus city-naming rules

```
"Do not return [] just because you are
uncertain -- a best guess is better than
nothing."
```

city_precision
City-naming rules only

```
"Brooklyn, Queens, Manhattan, Bronx,
Staten Island → 'New York City'"
```

meta_antidefault
From error-driven refinement: corrects real death place vs. birth or career city confusion

```
"Output the city ONLY if you have a
direct, specific memory of where they
died."
```

Set-valued

countryLandBordersCountry

One prompt which goes through N/S/E/W borders

prompt

```
"Check all directions: north, south,
east, west. Only land borders count, not
maritime."
```

companyTradesAtStockExchange

Five prompt variants. Output kept only if at least three agree.

simple
Baseline few-shot query

listed_check
Confirm company is listed before naming the exchange.

meta_precision
Recalls actual listing; no inferring from name or sector.

meta_verify
Verify before answering: Publicly traded company? Actual listing recalled?

anti_confusion
Separate common mix-ups, e.g., NYSE / Nasdaq

Keep if ≥ 3 of 5 agree

awardWonBy

Two prompts sweep by going through A to Z and year-by-year. A name is kept only if ≥ 2 of 5 samples agree. Results are unioned.

A to Z name prompt
Year-by-year prompt
→ **Union**

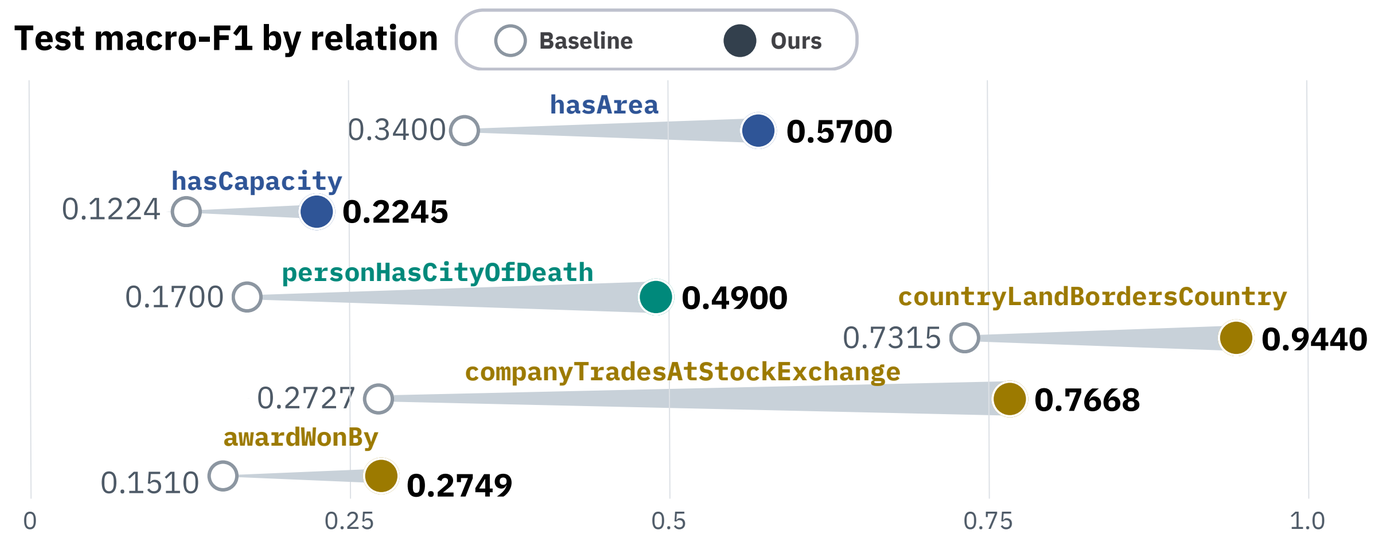
RESULTS

0.5577
Val macro-F1 · baseline 0.3131

0.5699
Test macro-F1 · baseline 0.2964

All scores (macro-F1)

Relation	Base val	Ours val	Base test	Ours test
hasArea	0.2900	**0.5800**	0.3400	**0.5700**
hasCapacity	0.1800	**0.2200**	0.1224	**0.2245**
personHasCityOfDeath	0.2100	**0.5000**	0.1700	**0.4900**
countryLandBordersCountry	0.6650	**0.9892**	0.7315	**0.9440**
companyTradesAtStockExchange	0.3543	**0.6752**	0.2727	**0.7668**
awardWonBy	0.1008	**0.1802**	0.1510	**0.2749**
Overall	**0.3131**	**0.5577**	**0.2964**	**0.5699**



LIMITATIONS

- Entity prominence and identification errors
  - hasArea and companyTradesAtStockExchange struggle with less prominent entities.
- Incomplete enumeration in awardWonBy
  - Some awards have hundreds of recipients. The model recalls some reliably but cannot list the full set.
- personHasCityOfDeath
  - Errors include no death location, or a career city or birthplace instead. Recency alone cannot explain these failures.

CONCLUSION

- Test macro-F1 of **0.5699**, within a 32B budget and without retrieval or fine-tuning.
- **Main strength is agreement.** Keeping only the answers that recur across prompts and samples draws out reliable knowledge and filters errors.
- Techniques are tailored to each relation.
- Errors remain in entity identification, enumeration, and factual prediction.

FURTHER READING

- Paper: https://aclanthology.org/2026.akbc-1.44/
- Code & prompts: https://github.com/tanaree-pat/akbc-2026

Paper **Code**