LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

SHARIF UNIVERSITY OF TECHNOLOGY

VOLKSWAGEN AKTIENGESELLSCHAFT DATA:LAB MUNICH

Hengam: An Adversarially Trained Transformer for Persian Temporal Tagging

Sajad Mirzababaei*, Amir Hossein Kargaran*, Hinrich Schütze, and Ehsaneddin Asgari

AACL

Why Temporal Tagging?

Temporal Tagging Applications:

- Text summarization
- Question answering
- Relation Extraction
- Information retrieval tasks requiring to classify information in a chronological order

1/11 (Mirzababaei*, Kargaran*) et al. - Hengam: An Adversarially Trained Transformer for Persian Temporal Tagging

Why Temporal Tagging for Persian?

Persian native speakers: 70 million (110 million total speakers)

Why Temporal Tagging for Persian?

Temporal Tagging in top High-resource languages, e.g., English

- Renowned rule-based systems, e.g., *HeidelTime* (2010), *SUTime* (2013)
- High-quality datasets
- Learning-based approaches, e.g., "*BERT got a Date*" (2021)

Temporal Tagging in Persian

- Rule-based systems, one attempt, *Parstime* (2018), lack of documentation to run
- No high-quality datasets available
- Learning-based approaches, trained on low number of sentences in NER task (2019-2021)

Rule-Based Limitations

Rule-based system limitations:

- Inability to handle ambiguities in the language
- Incapability to deal with a wide range of temporal terms
- Failing to generalize

Why Temporal Tagging for Persian?

Temporal Tagging in top High-resource languages, e.g., English

- Renowned rule-based systems, e.g., *HeidelTime* (2010), *SUTime* (2013)
- High-quality datasets
- Learning-based approaches, e.g., “*BERT got a Date*” (2021)

Temporal Tagging in Persian

- Rule-based systems, one attempt, *Parstime* (2018), lack of documentation to run
- No high-quality datasets available
- Learning-based approaches, trained on low number of sentences in NER task (2019-2021)

Persian Challenges

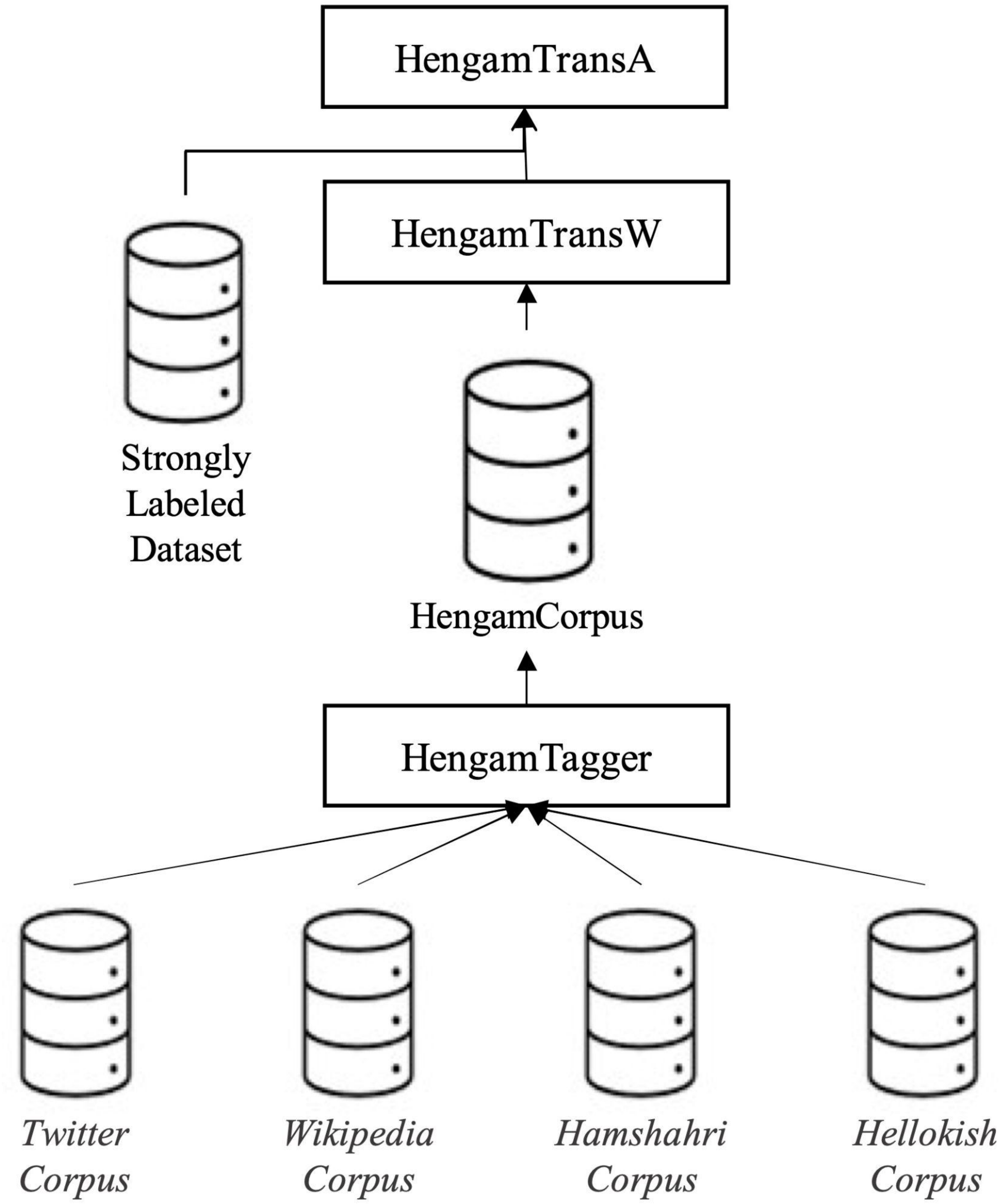
Language-Specific Challenges for Persian:

- Difference between formal and informal writing styles
- Lexical ambiguity (homographs)
- Use of three calendar systems in Persian: the Gregorian, Hijri, and Jalali calendars

Solution

Hengam Workflow

- HengamTagger: our rule-based system
- HengamCorpus: labeled corpus by HengamTagger
- HengamTransW: trained transformer model over HengamCorpus
- HengamTransA: fine-tuned HengamTransW over strongly labeled data

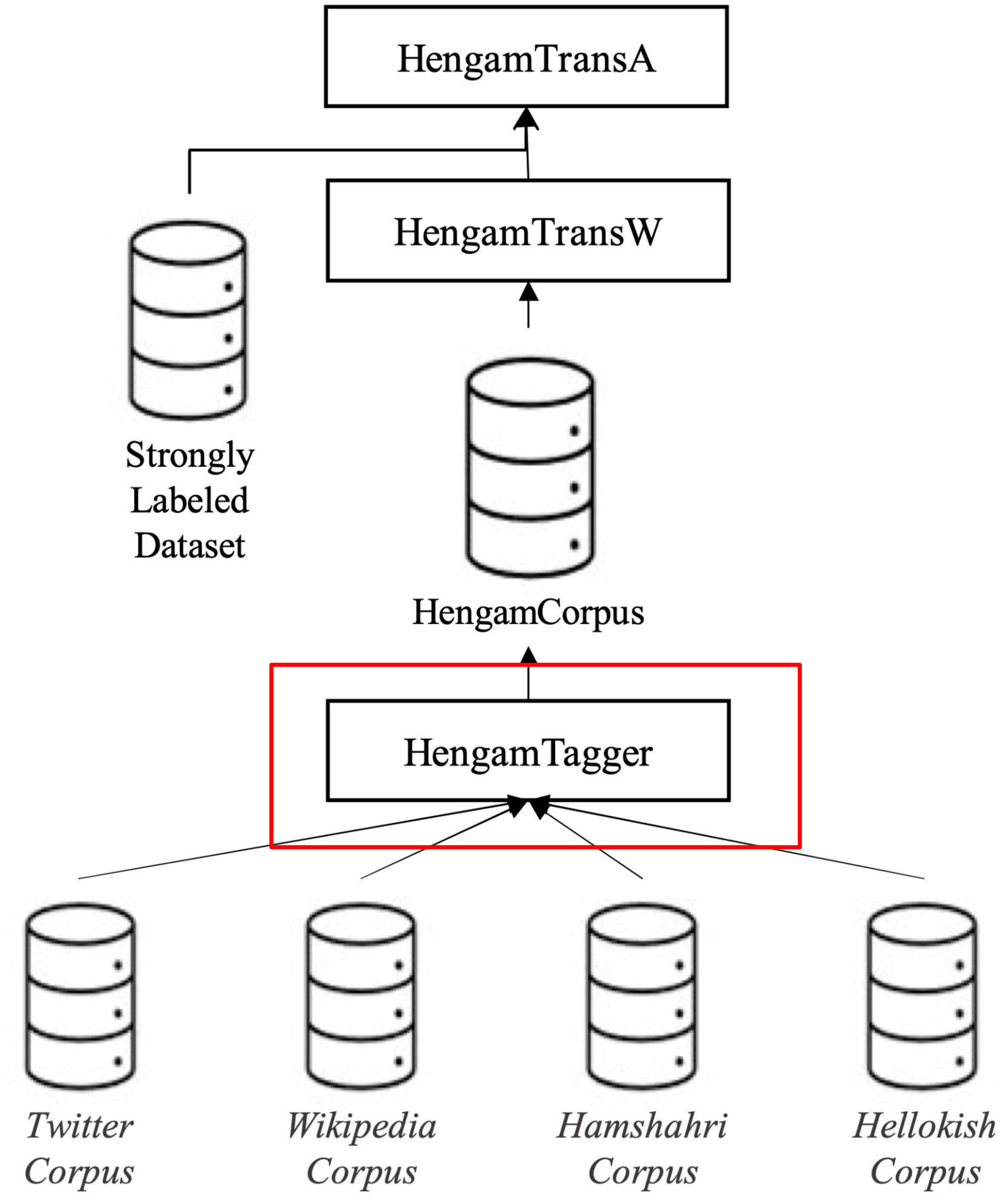


Rule-Based Temporal Tagging for Persian

One attempt, *Parstime* (2018), lack of documentation to run

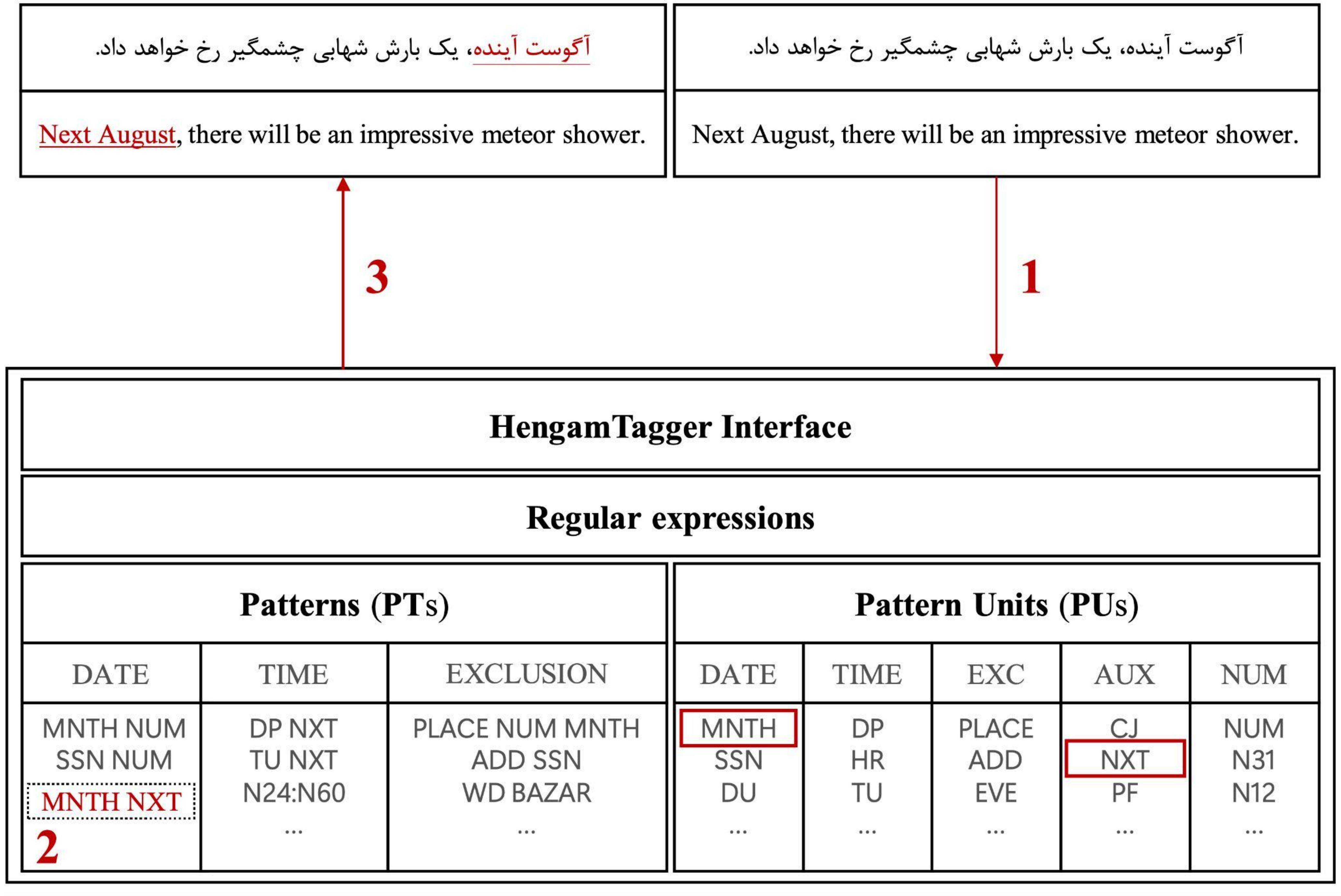
Rule-Based Temporal Tagging for Persian

HengamTagger



Rule-Based Temporal Tagging for Persian

HengamTagger



آگوست آینده، یک بارش شهابی چشمگیر رخ خواهد داد.

Next August, there will be an impressive meteor shower.

آگوست آینده، یک بارش شهابی چشمگیر رخ خواهد داد.

Next August, there will be an impressive meteor shower.

HengamTagger Interface

Regular expressions

Patterns (PTs)

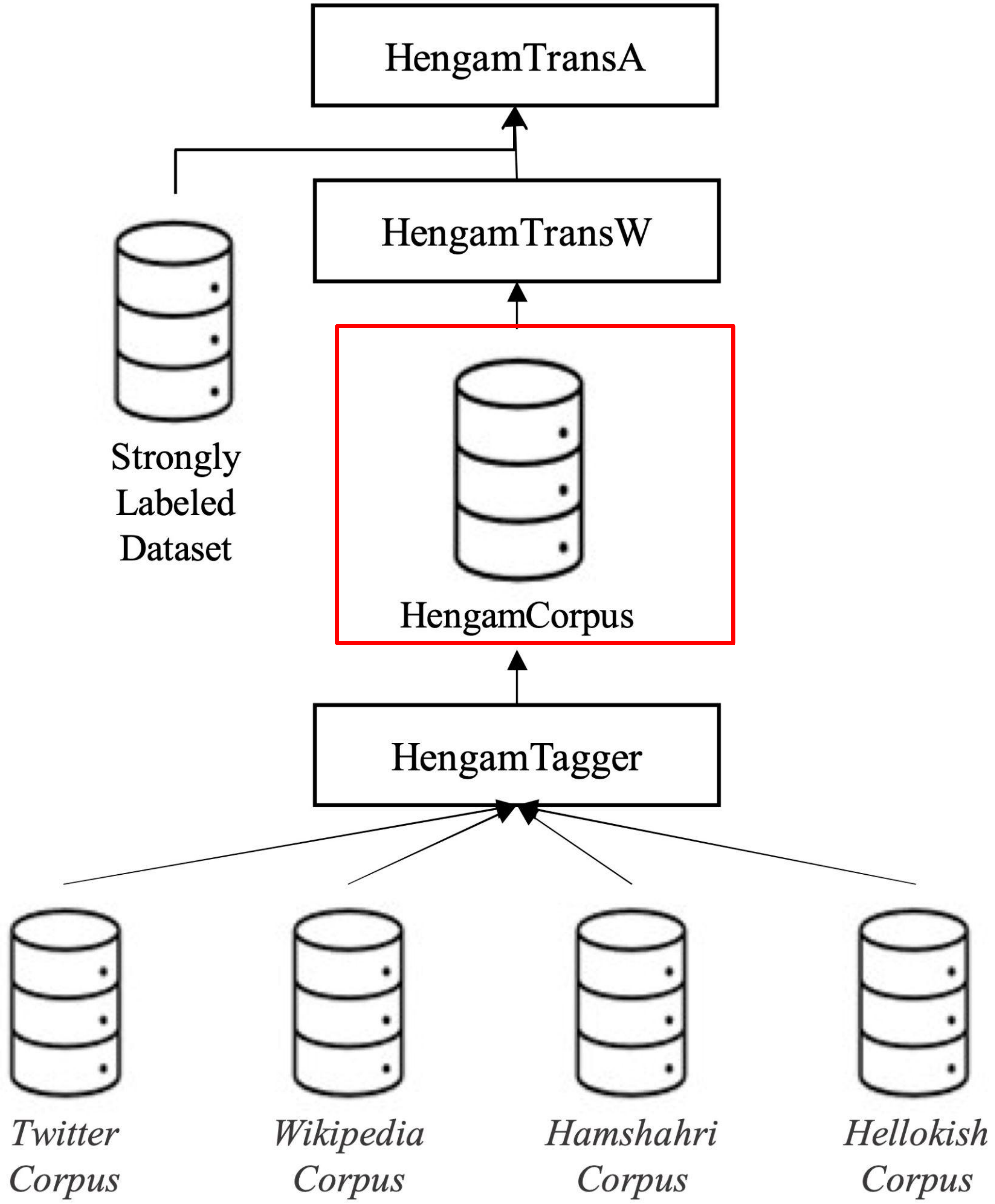
DATE	TIME	EXCLUSION
MNTH NUM SSN NUM MNTH NXT	DP NXT TU NXT N24:N60 ...	PLACE NUM MNTH ADD SSN WD BAZAR ...

Pattern Units (PUs)

DATE	TIME	EXC	AUX	NUM
MNTH SSN DU ...	DP HR TU ...	PLACE ADD EVE ...	CJ NXT PF ...	NUM N31 N12 ...

https://github.com/kargaranamir/parstdex

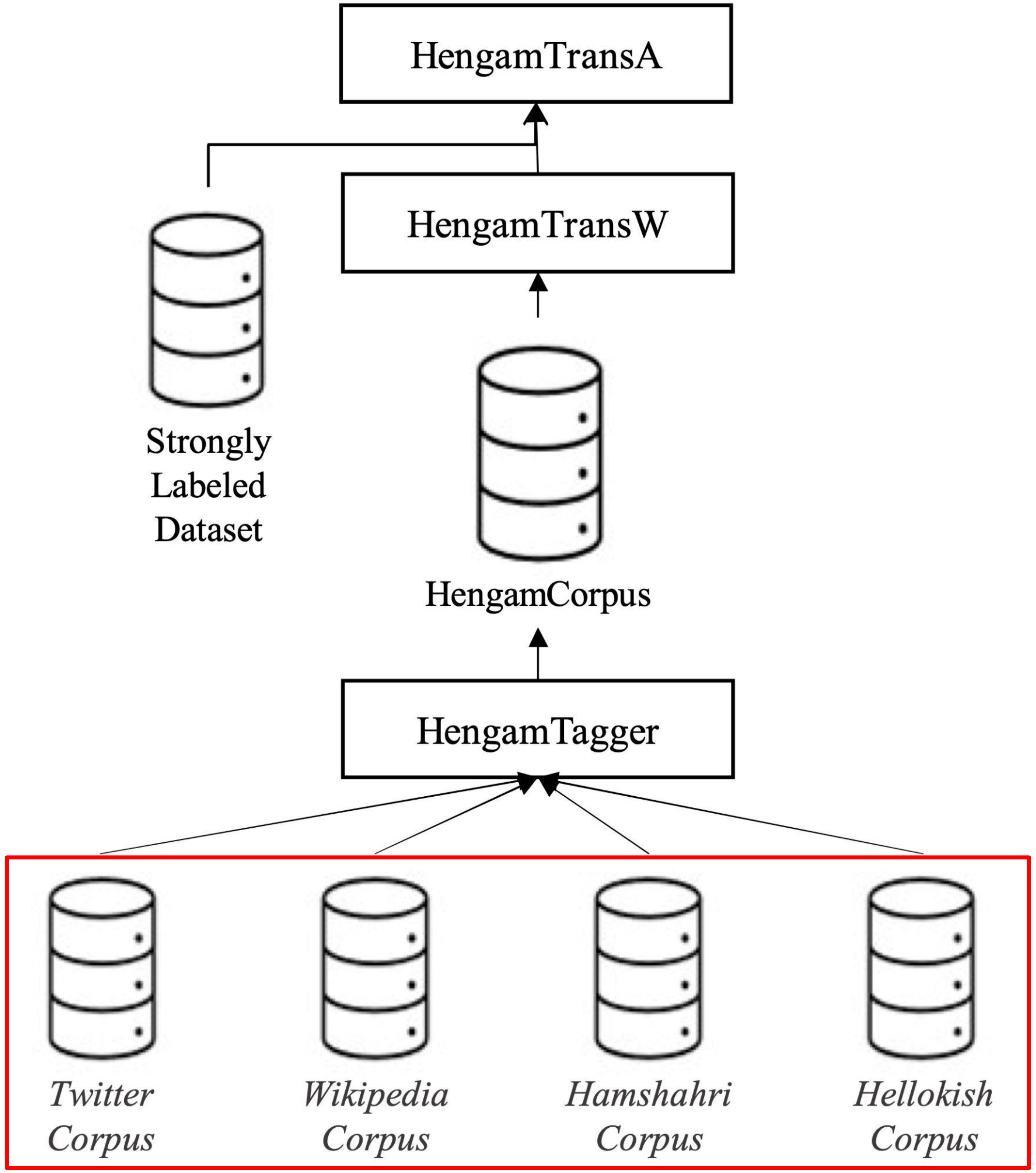
HengamCorpus



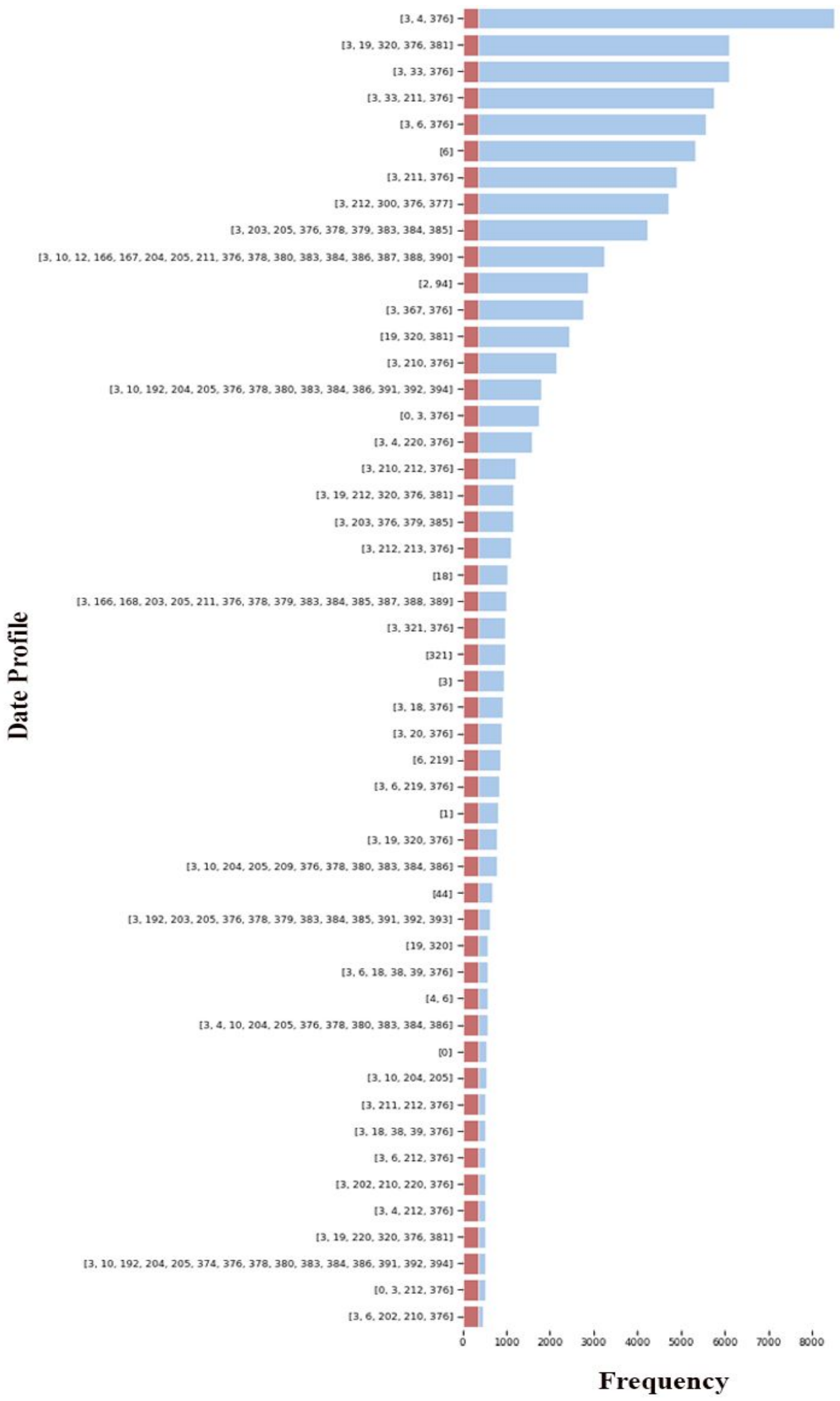
HengamCorpus

Raw Persian Texts:

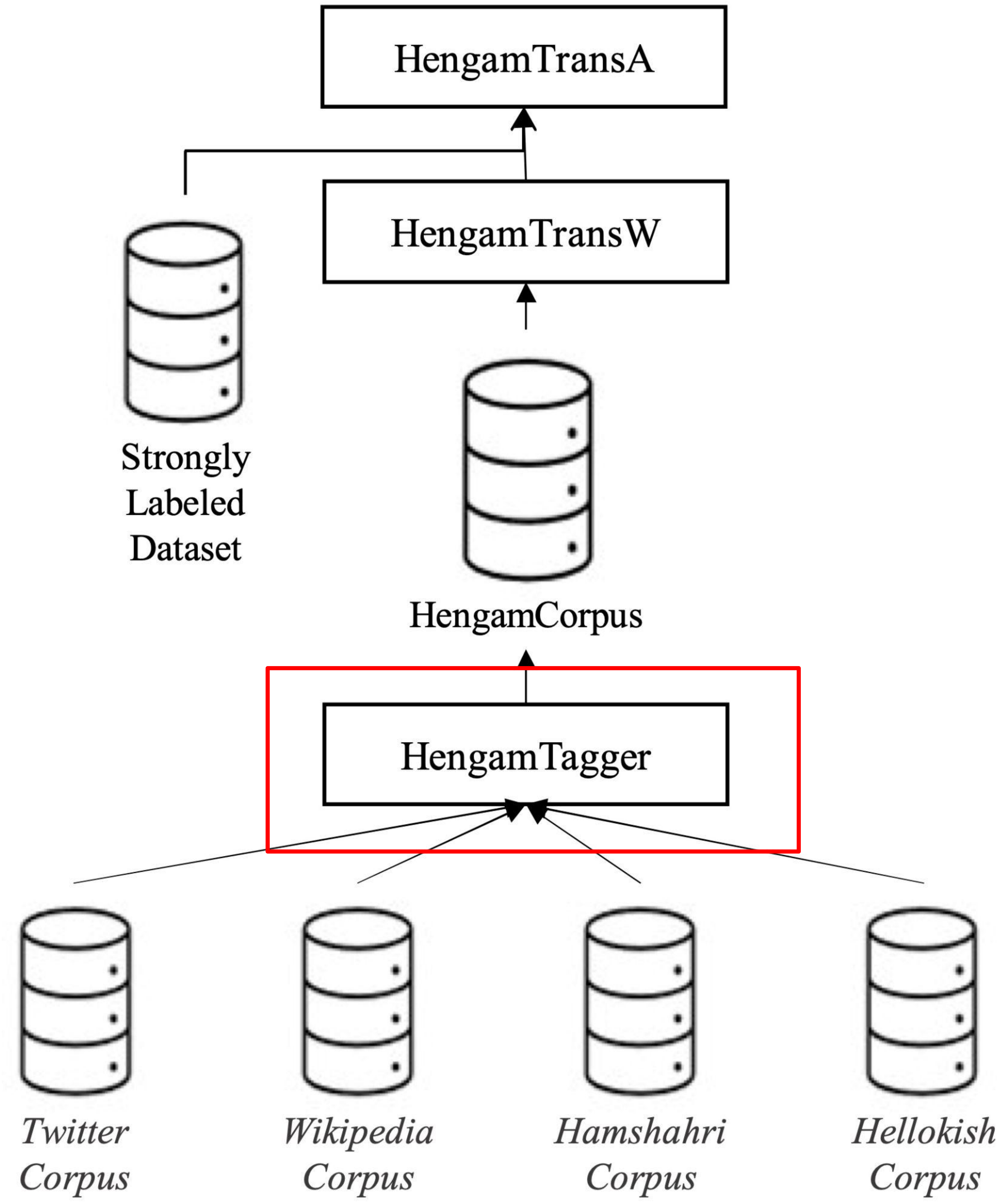
- Formal Texts
  - – Wikipedia - 3,858,609 sentences
  - – Hamshahri - 1,793,147 sentences
- Informal texts
  - – Twitter - 9,852,565 tweets
  - – Hellokish - 7,899 sentences



HengamCorpus



Skewness of date/time profile distributions



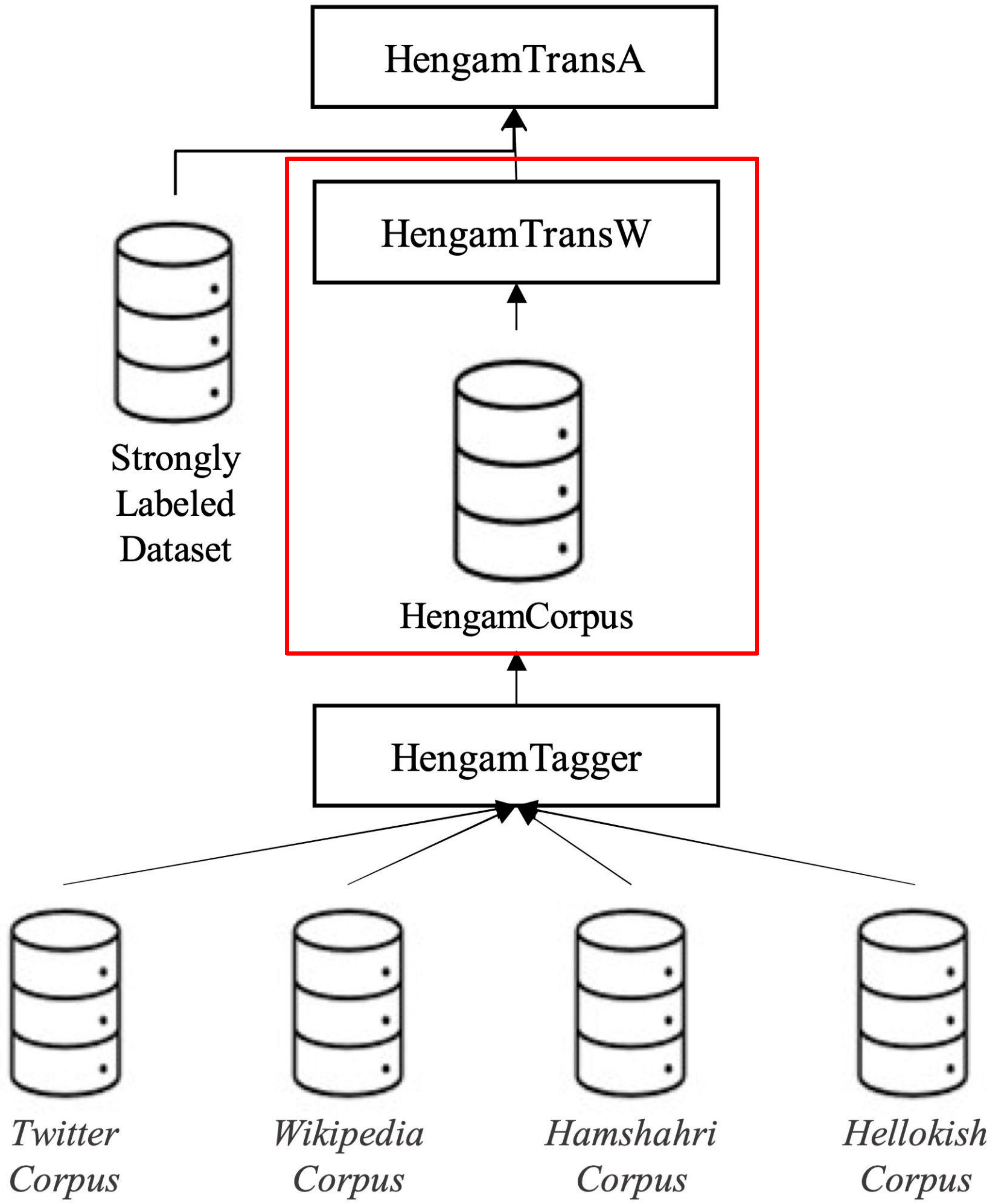
HengamTansformer

HengamTransformer architecture:

- XLM-RobBERTa transformer model
- Linear-chain CRF layer

Fine-tuning:

- Dataset: HengamCorpus
  - Train: 75%
  - Test: 10%
  - Validation: 15%



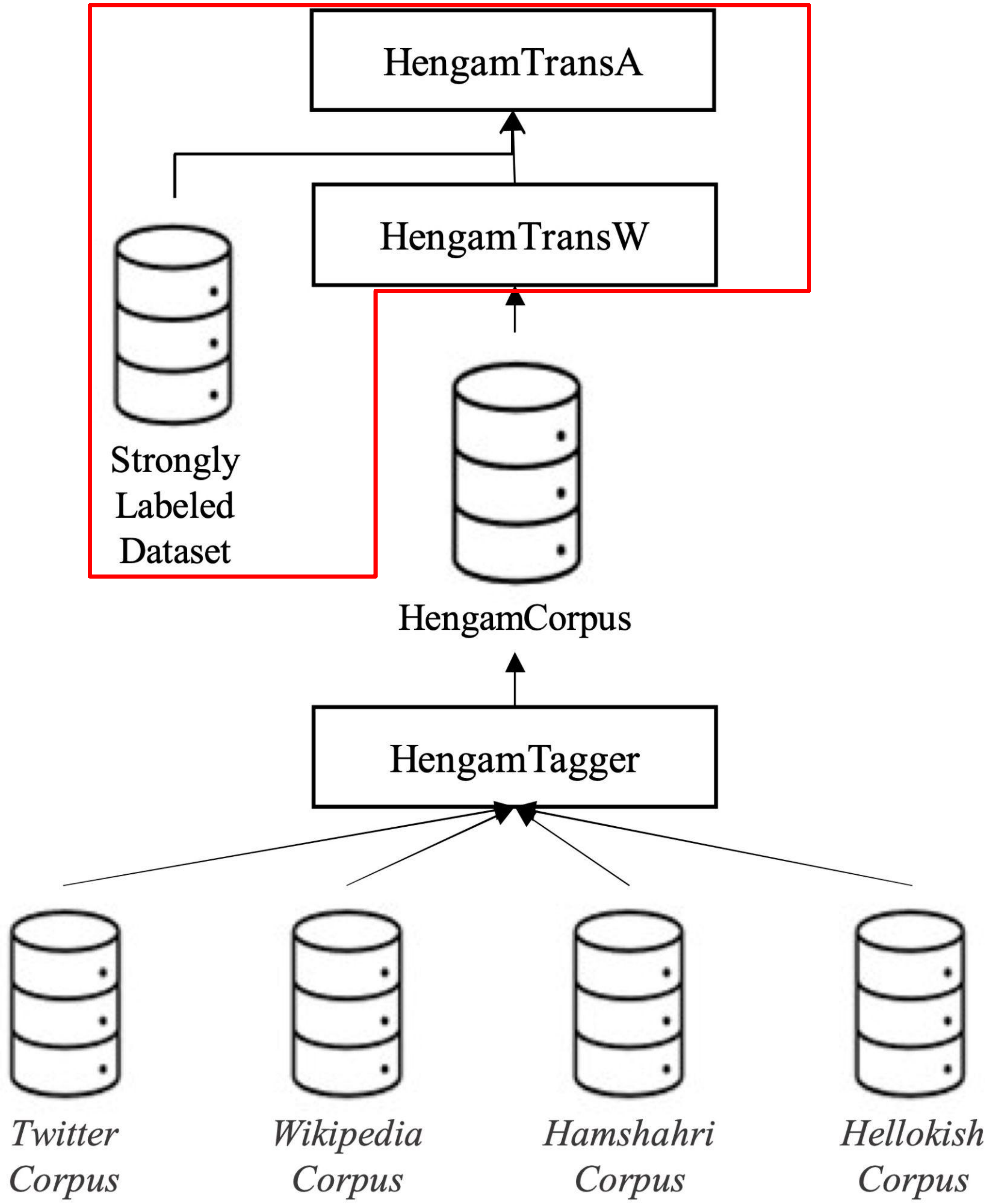
HengamTansformer

Strongly labeled dataset:

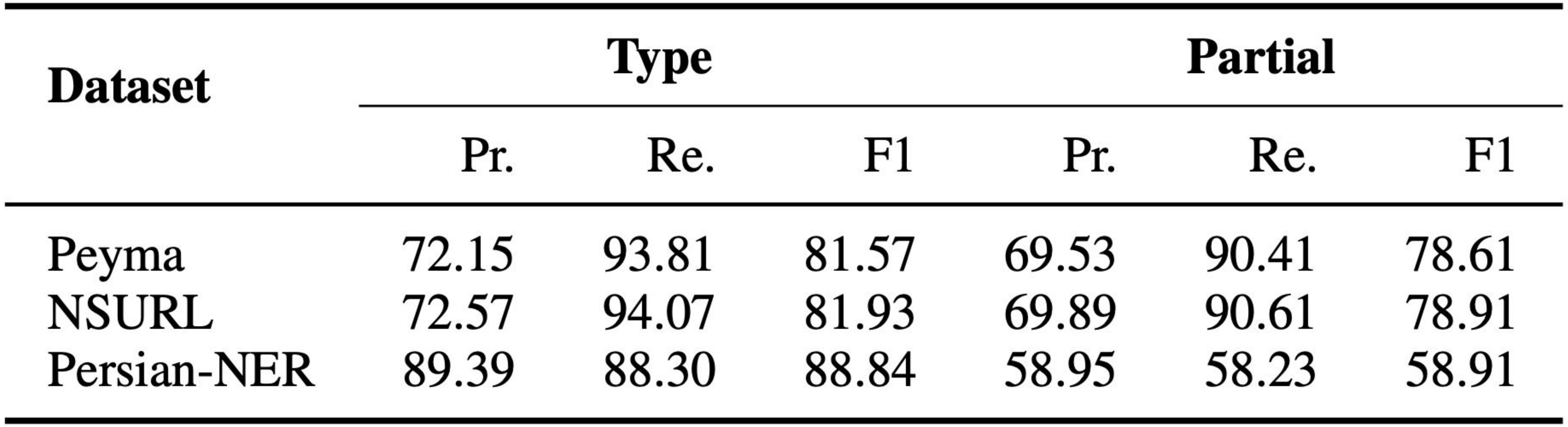
- 1500 sentences (0.5% ||HengamCorpus||)
- Kappa agreement score of 0.95

Adversarial training:

- Minimizing maximum risk of adversarial perturbations
- Method: K-projected gradient descent (K-PGD)



Exploring NER datasets using HengamTagger



Dataset	Type			Partial		
	Pr.	Re.	F1	Pr.	Re.	F1
Peyma	72.15	93.81	81.57	69.53	90.41	78.61
NSURL	72.57	94.07	81.93	69.89	90.61	78.91
Persian-NER	89.39	88.30	88.84	58.95	58.23	58.91

The performance of HengamTagger (Precision, Recall, and F1 scores) on Persian NER datasets containing temporal labels

Hengam Evaluation

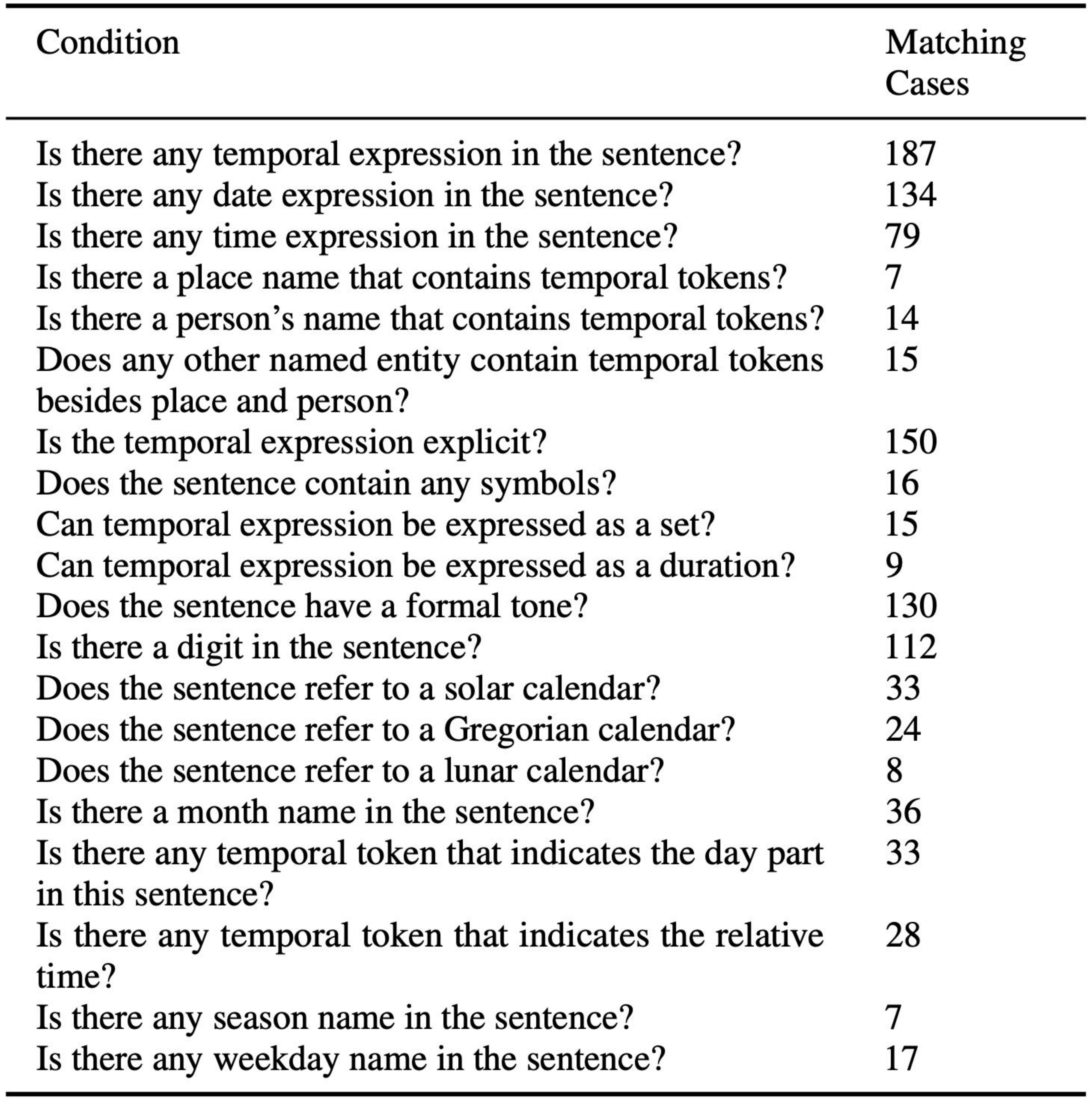
HengamGold

Creating an evaluation dataset:

- There is no evaluation dataset that covers most temporal patterns
- The existing datasets are noisy

HengamGold:

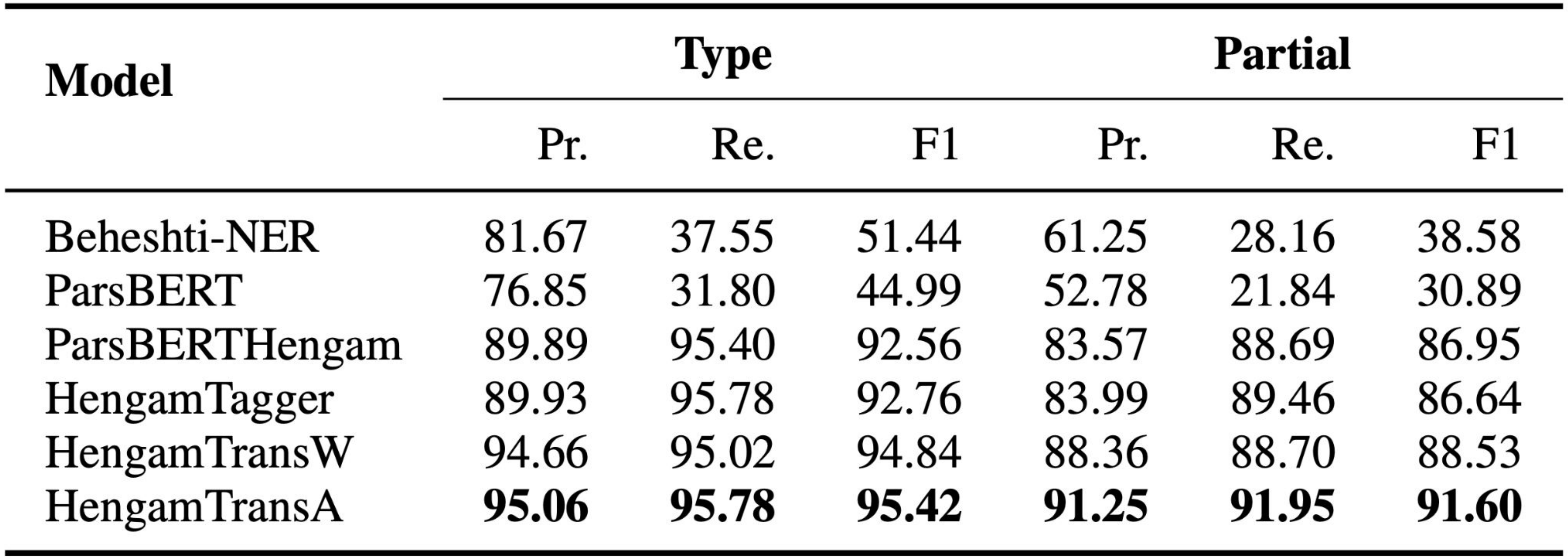
- 200 sentences
- Designed with 20 parameters
- Kappa agreement score of 0.97



Condition	Matching Cases
Is there any temporal expression in the sentence?	187
Is there any date expression in the sentence?	134
Is there any time expression in the sentence?	79
Is there a place name that contains temporal tokens?	7
Is there a person's name that contains temporal tokens?	14
Does any other named entity contain temporal tokens besides place and person?	15
Is the temporal expression explicit?	150
Does the sentence contain any symbols?	16
Can temporal expression be expressed as a set?	15
Can temporal expression be expressed as a duration?	9
Does the sentence have a formal tone?	130
Is there a digit in the sentence?	112
Does the sentence refer to a solar calendar?	33
Does the sentence refer to a Gregorian calendar?	24
Does the sentence refer to a lunar calendar?	8
Is there a month name in the sentence?	36
Is there any temporal token that indicates the day part in this sentence?	33
Is there any temporal token that indicates the relative time?	28
Is there any season name in the sentence?	7
Is there any weekday name in the sentence?	17

Parameters used in the creation of HengamGold

Hengam Evaluation



Model	Type			Partial		
	Pr.	Re.	F1	Pr.	Re.	F1
Beheshti-NER	81.67	37.55	51.44	61.25	28.16	38.58
ParsBERT	76.85	31.80	44.99	52.78	21.84	30.89
ParsBERTHengam	89.89	95.40	92.56	83.57	88.69	86.95
HengamTagger	89.93	95.78	92.76	83.99	89.46	86.64
HengamTransW	94.66	95.02	94.84	88.36	88.70	88.53
HengamTransA	**95.06**	**95.78**	**95.42**	**91.25**	**91.95**	**91.60**

Comparison of different variations of Hengam temporal detectors.
The Hengam models are compared with the Beheshti-NER (Taher et al., 2020), and ParsBERT (Farahani et al., 2021) as well.

Key Takeaways

- **HengamTagger**: an efficient and extensible **rule-based** temporal expression identification tool. It can be easily extended in supporting languages other than Persian.

- **HengamCorpus**: a sizeable unbiased **dataset** covering the majority of formal and informal temporal expressions in Persian.

- **HengamTransformer**: a state-of-the-art adversarial **transformer-based** temporal tagger that not only achieves the best performance but also successfully deals with language ambiguities and incorrect spellings.

Code, Model, Data, Interface: https://github.com/kargaranamir/hengam