K-Flares: A K-Adapter based approach for the FLARES challenge

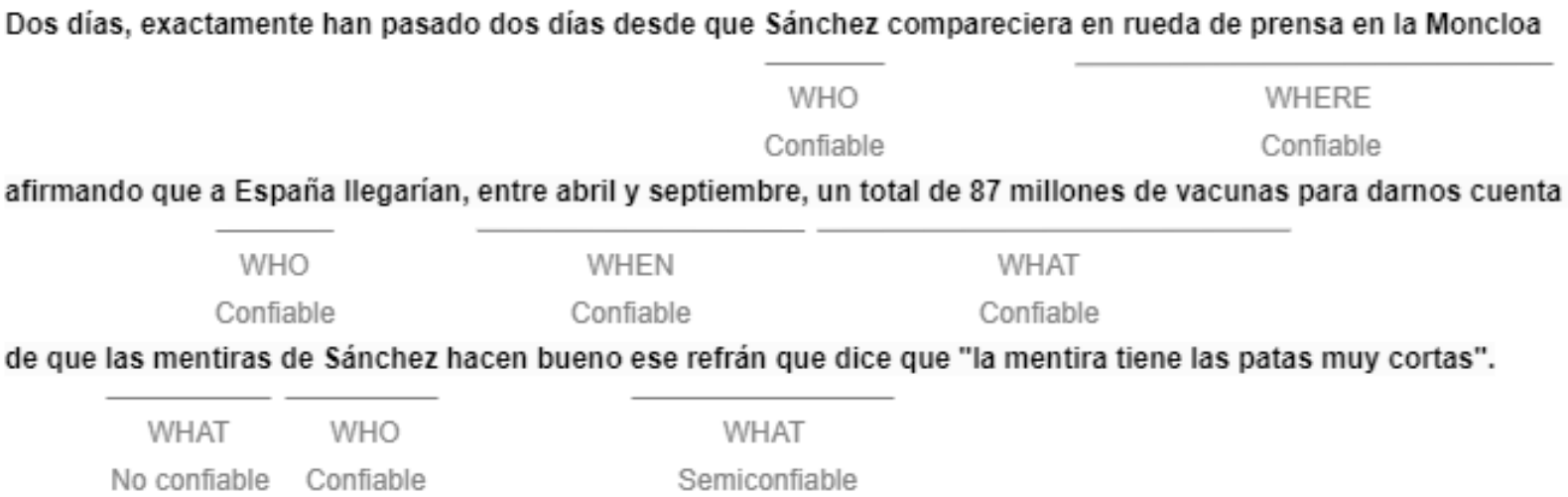
Joel Pardo[1], JiaYun[1], Virginia Ramón[1], Elvira Amador[1,2] y Pablo Calleja[1]
[1]*Ontology Engineering Group, Departamento de Inteligencia Artificial, ETSI Informáticos,*
[2]*Departamento de Sistemas Informáticos, ETSI Sistemas Informáticos,*
Universidad Politécnica de Madrid
{joel.pardof, jiayun.liu, virginia.ramon, elvira.amador, p.calleja}@upm.es

24/09/24

SEPLN 2024, Valladolid, Spain

FLARES dataset:

- Spanish news extracts annotated with **WHAT, WHO, WHERE, WHEN, WHY, and HOW** and "reliable", "semi-reliable" or "unreliable" (**"confiable", "semiconfiable" or "no confiable"** in Spanish)
- ~ **9,034 5W1H annotations** across 190 news articles
- 70% for training (6,934 annotations) and 30% for testing (2,100 annotations)



FLARES annotation example

Sociedad Española para el Procesamiento del Lenguaje Natural

R. Sepúlveda-Torres, A. Bonet-Jover, I. Diab, I. Guillén-Pacho, I. Cabrera-de Castro, C. BadenesOlmedo, E. Saquete, M. T. Martín-Valdivia, P. Martínez-Barco, L. A. Ureña-López, ***Overview of FLARES at IberLEF 2024: Fine-Grained Language-based Reliability Detection in Spanish News***, Procesamiento del Lenguaje Natural 73 (2024).

Subtask 1 → Detection of 5W1H

Dos días, exactamente han pasado dos días desde que Sánchez compareciera en rueda de prensa en la Moncloa

WHO WHERE

afirmando que a España llegarían, entre abril y septiembre, un total de 87 millones de vacunas para darnos cuenta

WHO WHEN WHAT

de que las mentiras de Sánchez hacen bueno ese refrán que dice que "la mentira tiene las patas muy cortas".

WHAT WHO WHAT

Subtask 2 → Reliability classification

Dos días, exactamente han pasado dos días desde que Sánchez compareciera en rueda de prensa en la Moncloa

WHO WHERE

Confiable Confiable

afirmando que a España llegarían, entre abril y septiembre, un total de 87 millones de vacunas para darnos cuenta

WHO WHEN WHAT

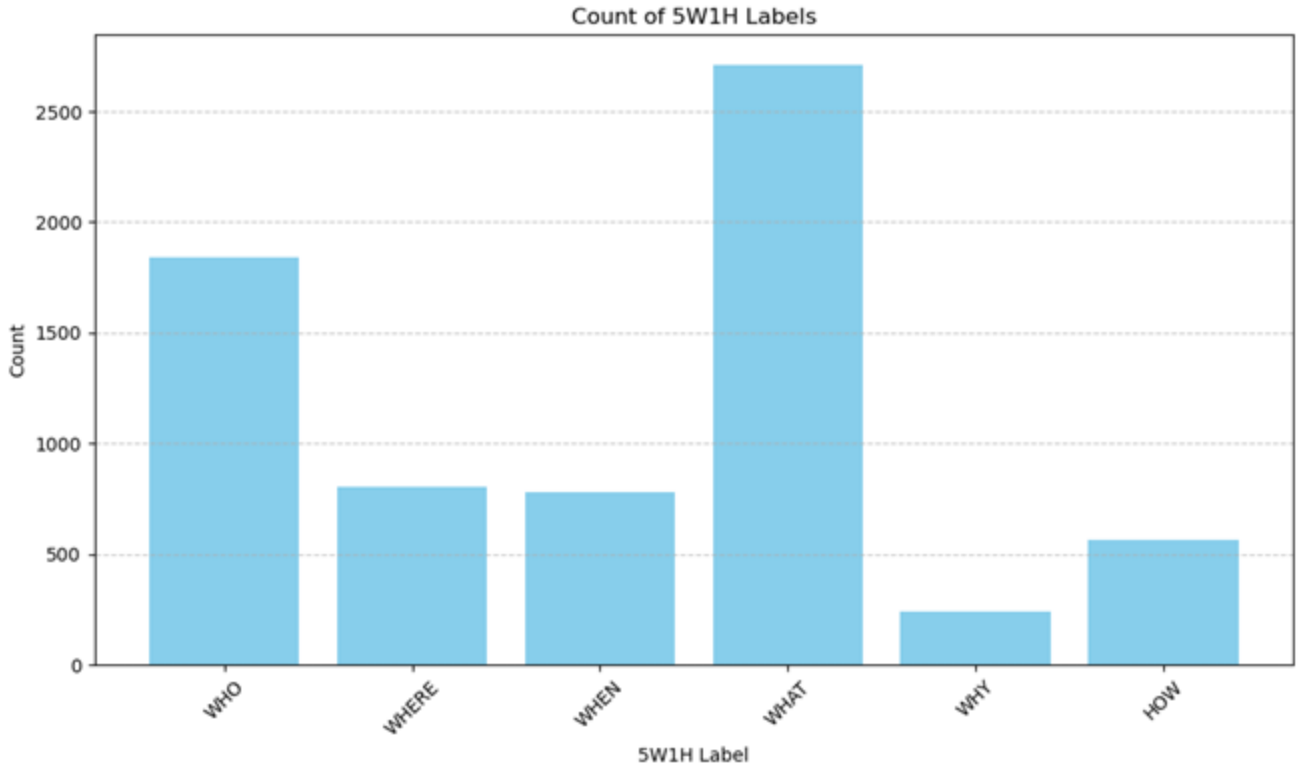
Confiable Confiable Confiable

de que las mentiras de Sánchez hacen bueno ese refrán que dice que "la mentira tiene las patas muy cortas".

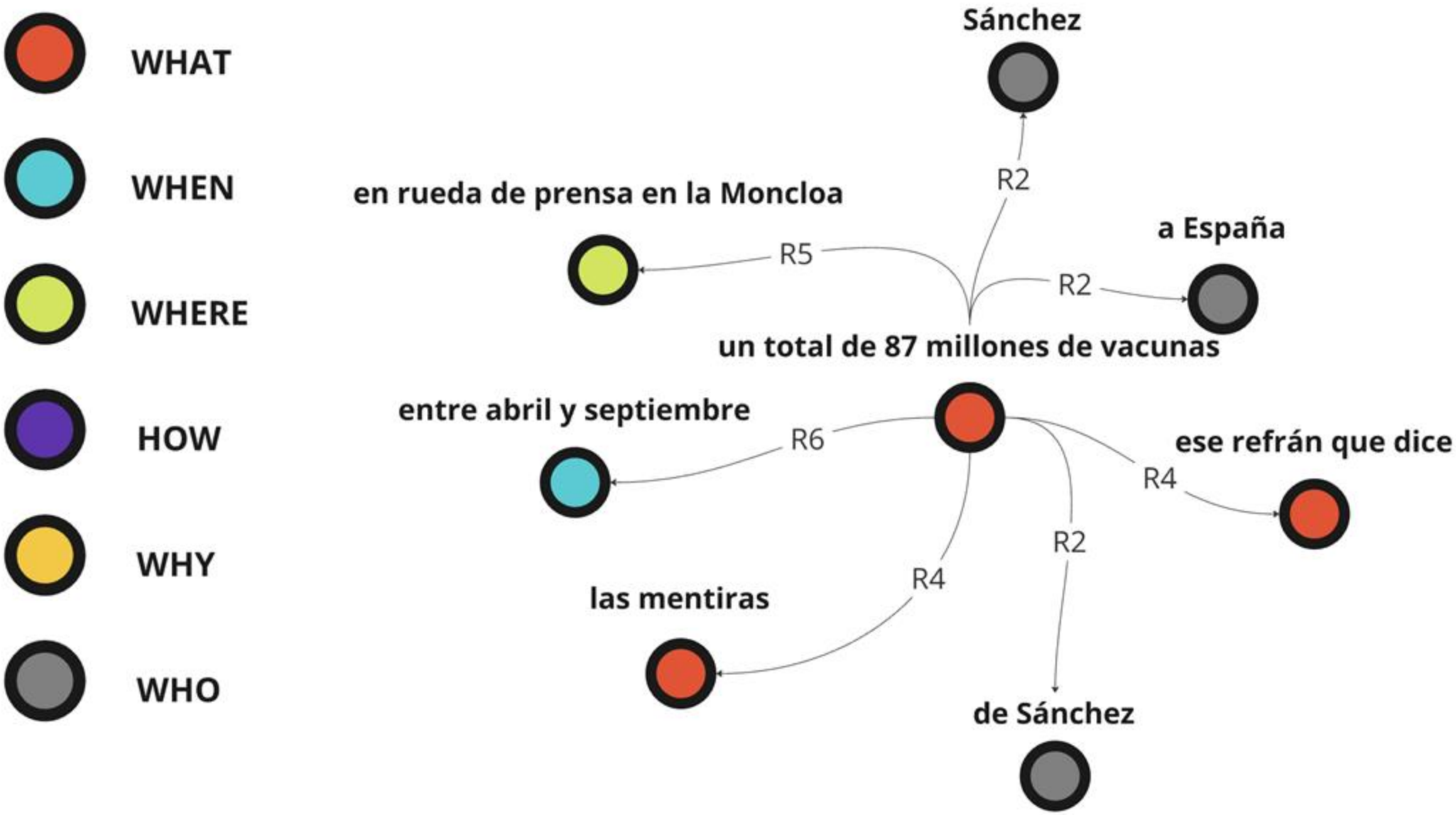
WHAT WHO WHAT

No confiable Confiable Semiconfiable

- No syntactical consistency
- Regular interconnection in the 5W1H instances in each individual text
  - Graph-like structure → inject information to the model



Example of graph

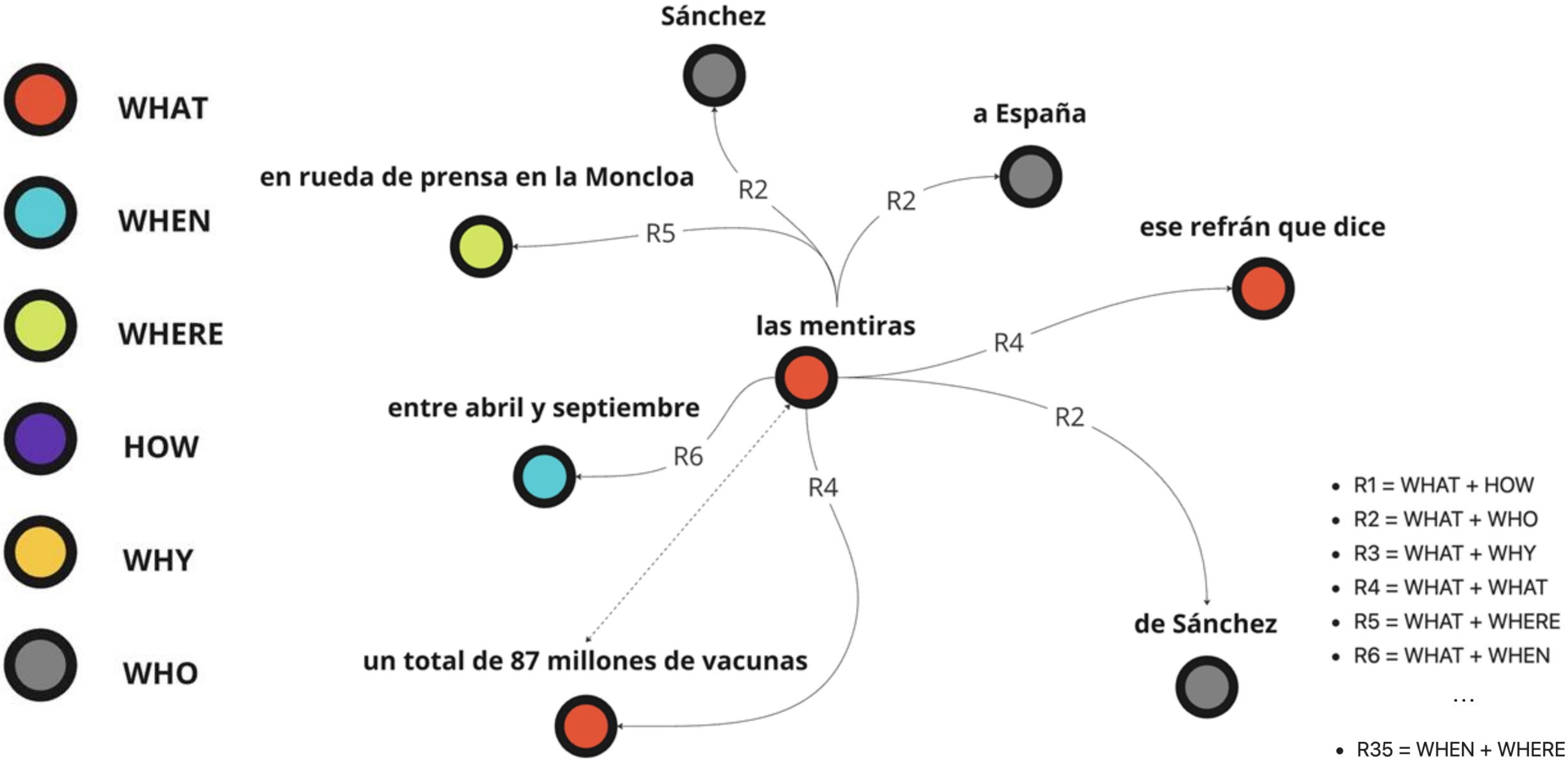


- R1 = WHAT + HOW
- R2 = WHAT + WHO
- R3 = WHAT + WHY
- R4 = WHAT + WHAT
- R5 = WHAT + WHERE
- R6 = WHAT + WHEN

…

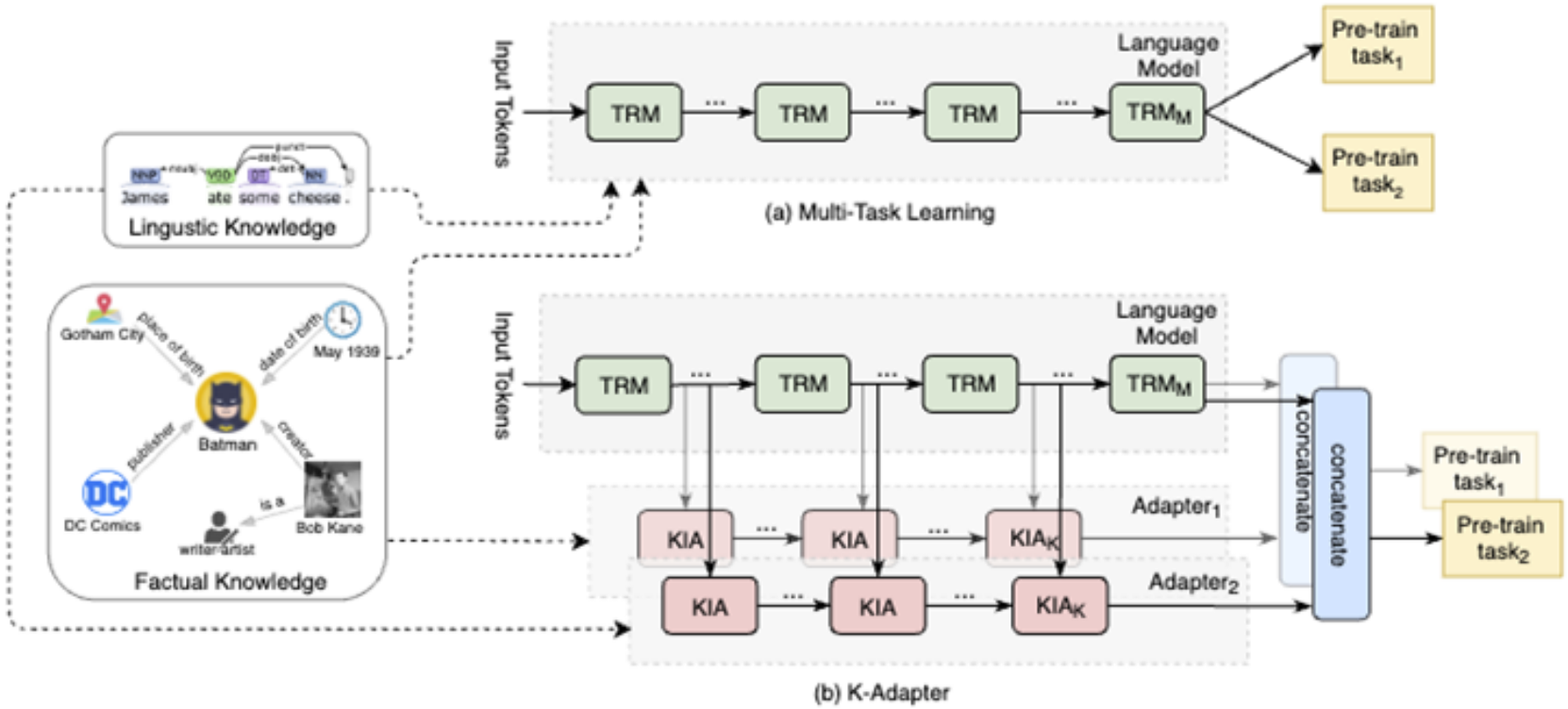
- R35 = WHEN + WHERE
- R36 = WHEN + WHEN

Graph extracted from data



Purpose: Knowledge Injection for pretrained language models to improve their performance without fine-tuning the entire model.

Adapters: neural network to encode knowledge from external sources. Added to the layers of a pre-trained language model. The adapters are trained separately on the knowledge sources. The adapter modules are designed to be plug-and-play



Wang, R., Tang, D., Duan, N., Wei, Z., Huang, X., Cao, G., & Zhou, M. (2020). K-adapter: Infusing knowledge into pre-trained models with adapters. *arXiv preprint arXiv:2002.01808*.

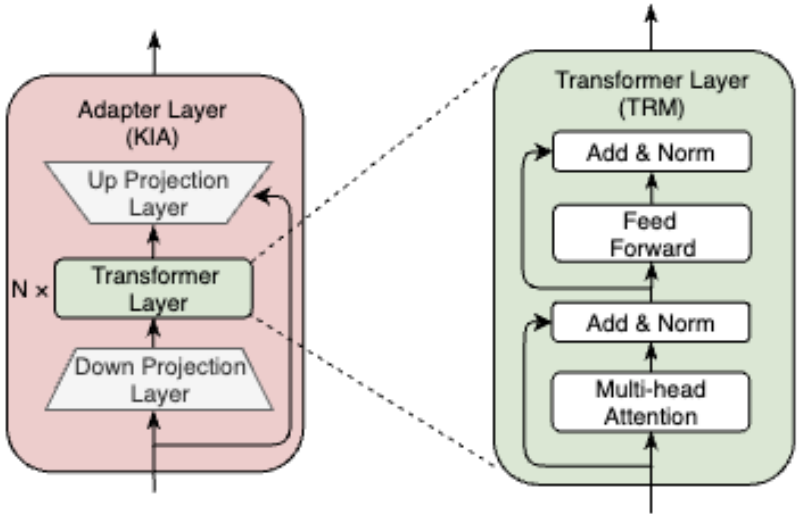


Figure 2: Structure of the adapter layer (left). The adapter layer consists of two projection layers and N=2 transformer layers, and a skip-connection between two projection layers.

Similar to Encoders Layers

ORIGINAL WORK

Language Model:

Roberta-large language model

Adapters:

1- *Factual*: T-REX dataset. Trained for relation classification

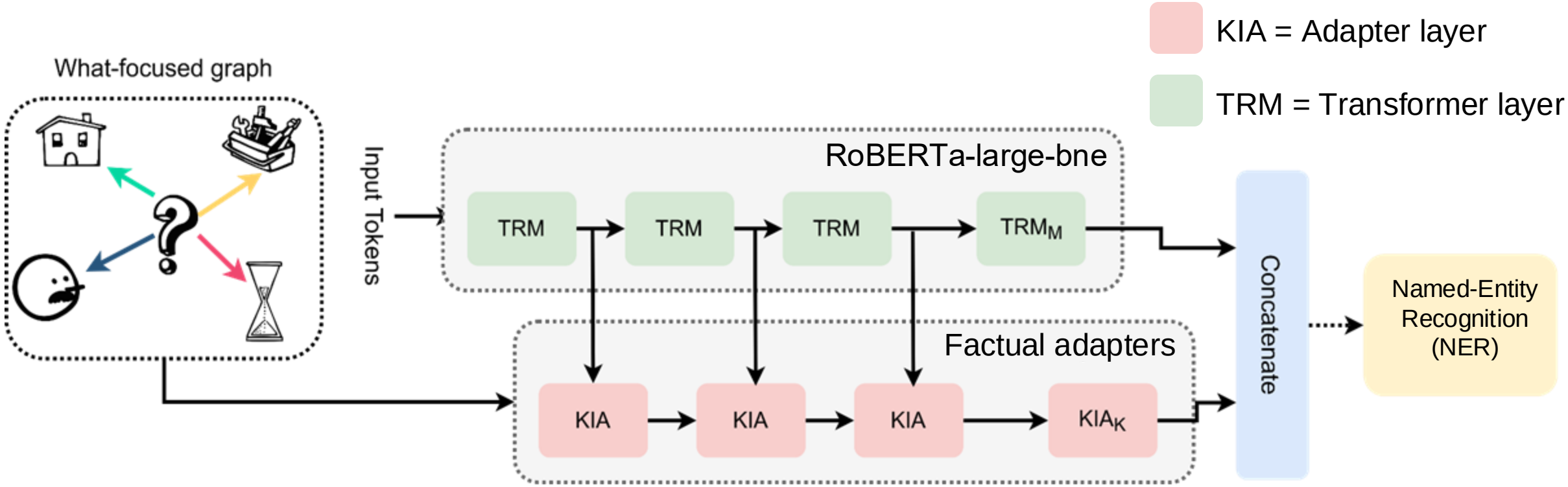
2- *Linguistic*: Dataset from dependency parser from Book Corpus. Trained for dependency relation prediction (predict the head index of each token in the sentence)

Tasks for evaluation:

- Entity Typing (Type of NER/Classification)
- Question Answering
- Relation extraction

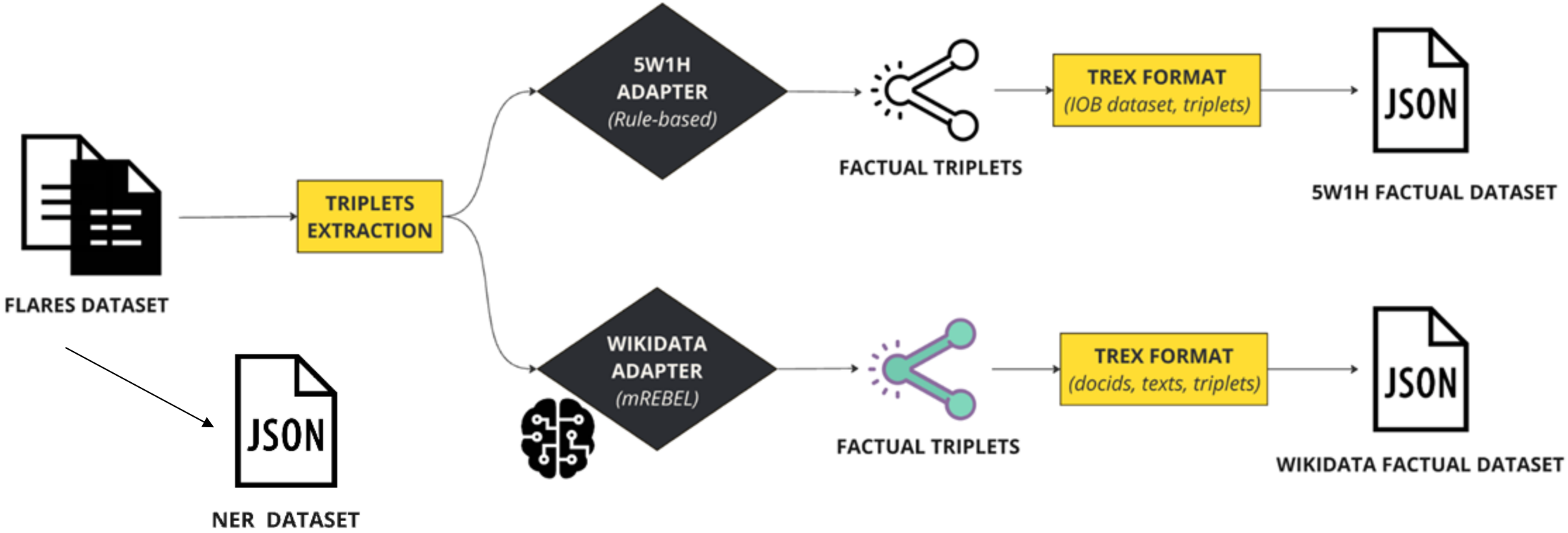
OUR WORK

- **Language Model:** Roberta MarIA bsc-bne (Spanish)
- **Adapters:** Two factual adapters (5W1H labels, Wikidata labels) with the same training dataset
- **Task for evaluation:** Named Entity Recognition (Token Sequence Classification)



The factual adapters for the challenge

Rule-based:
Mix of the labels between subject and object labels



mREBEL:
Given a text input, directly output triplets

5W1H TRIPLETS:

WIKIDATA TRIPLETS:

Using mREBEL

1- REBEL: relation extraction architecture trained with Wikidata + Wikipedia information

2- mREBEL: Multilingual training dataset

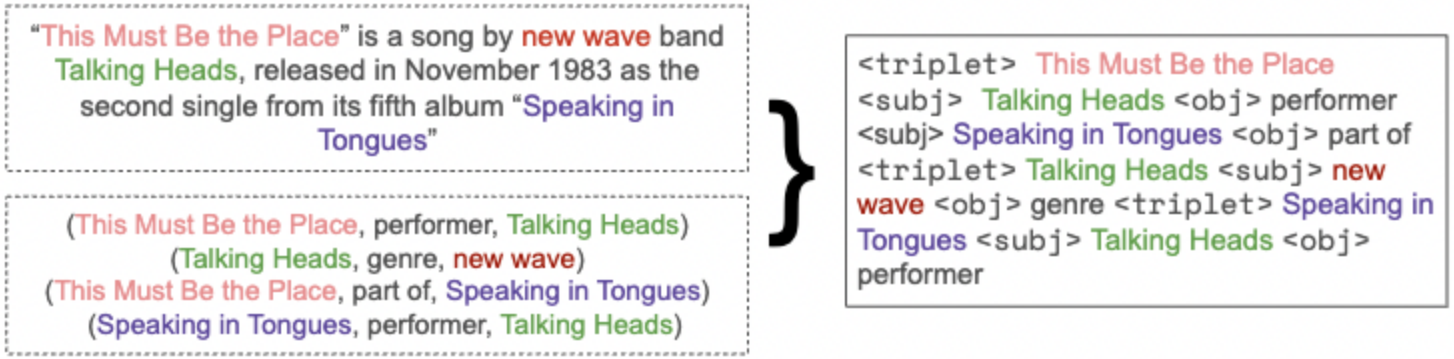


Figure 1: Example of the triplet linearization process for REBEL.

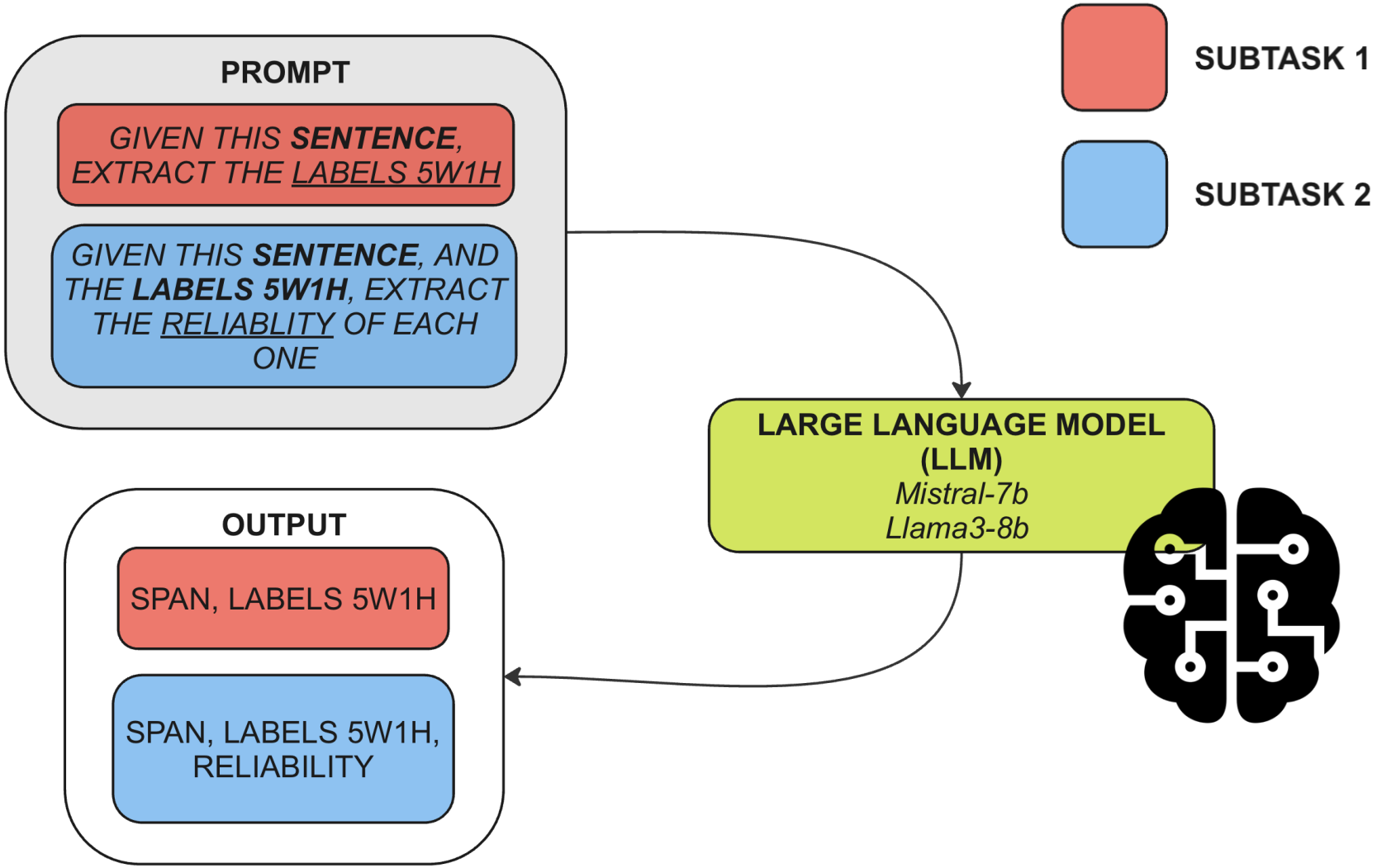
Cabot, P. L. H., Tedeschi, S., Ngomo, A. C. N., & Navigli, R. (2023). REDFM: a Filtered and Multilingual Relation Extraction Dataset *arXiv preprint arXiv:2306.09802*.

K-Adapter Best Hyperparameters (Subtask 1):

Table 1
Comparison of hyperparameters across 5W1H labels and Wikidata Factual Adapters, and Complete K-Adapter configurations.

Parameter	5W1H labels adapter	Wikidata adapter	Complete K-Adapter
epochs	5	10	8
model	roberta-large	Roberta-large	roberta-large
per_gpu_train_batch_size	64	32	8
per_gpu_eval_batch_size	64	8	8
max_seq_length	64	64	512
scheduler	WarmupLinearSchedule	WarmupLinearSchedule	WarmupLinearSchedule
optimizer	AdamW	AdamW	AdamW
learning_rate	5e-5	2e-5	5e-5
warmup_steps	1200	500	120
freeze_adapter	False	False	True
adapter_size	768	768	768
adapter_list	"0,11,22"	"0,11,22"	"0,11,22"
adapter_transformer_layers	2	2	-
fusion_mode	-	-	add

Experiments with Prompt Engineering



Subtask 1

Submitted (Ranking)	Adapters	30% split test set	70% split test set	Training epochs
K-adapter KG (1st)	5W1H	0.65957	0.66544	8
K-adapter KG 2A (3rd)	5W1H + Wikidata	0.39649	0.40070	8
K-adapter KG+ (3rd)	5W1H	0.39578	0.39572	12

Subtask 2

Submitted (Ranking)	30% split test set	70% split test set
Mistral-7B Fine-tuning (4rd)	0.30585	0.29063

K-Flares: A K-Adapter based approach for the FLARES challenge

Joel Pardo[1], JiaYun[1], Virginia Ramón[1], Elvira Amador[1,2] y Pablo Calleja[1]
[1]*Ontology Engineering Group, Departamento de Inteligencia Artificial, ETSI Informáticos,*
[2]*Departamento de Sistemas Informáticos, ETSI Sistemas Informáticos,*
Universidad Politécnica de Madrid
{joel.pardof, jiayun.liu, virginia.ramon, elvira.amador, p.calleja}@upm.es

24/09/24

SEPLN 2024, Valladolid, Spain