

# ZS4IE: A toolkit for Zero-Shot Information Extraction with Simple Verbalizations

Oscar Sainz<sup>1,\*</sup>, Haoling Qiu<sup>2,\*</sup>, Oier Lopez de Lacalle<sup>1</sup>, Eneko Agirre<sup>1</sup> and Bonan Min<sup>2</sup>

<sup>1</sup>HiTZ Basque Center for Language Technologies - Ixa NLP Group

University of the Basque Country (UPV/EHU)

<sup>2</sup>Raytheon BBN Technologies

## Schema development

### NER VERBALIZATIONS

{X} is a person. → PERSON  
{X} is a date. → DATE

### EVENT VERBALIZATIONS

{E} refers to a death. → LIFE.DIE

### RELATION VERBALIZATIONS

{X} is employed by {Y}. → EMPLOYEEOF

### EVENT ARGUMENT VERBALIZATIONS

{X} died. → VICTIM-ARG  
Someone {E} on {X}. → PLACE-ARG



- The current **define-then-annotate-and-train workflow** unfortunately **requires starting from scratch** for each new domain and schema.
- We present an alternative **verbalize-while-defining** workflow where the analyst **defines the schema interactively**.
- The user makes use of **simple natural language verbalizations** to define new entity, event, relation or arguments type.

## Inference

### (1) CANDIDATE GENERATION



### (2) LABEL VERBALIZATION

for candidate in candidates:

PERSON  
T1 John Smith is a person.  
ORGANIZATION  
T2 John Smith is an organization.  
DATE  
T3 John Smith is a date.  
...

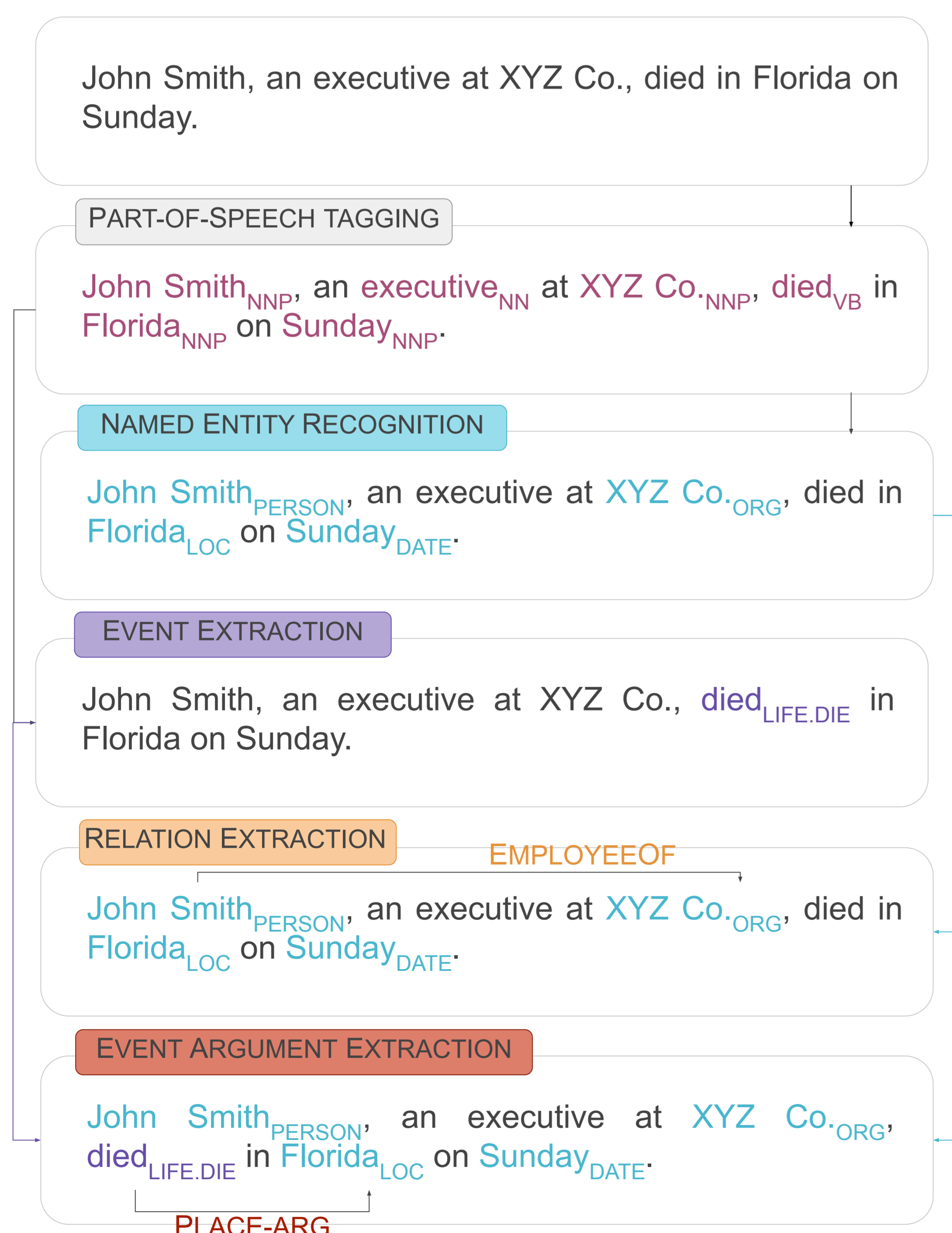
### (3) INFERENCE

P( Ent | CONTEXT John Smith is a person. ) = ██████  
P( Ent | CONTEXT John Smith is a date. ) = █  
P( Ent | CONTEXT John Smith is an org. ) = █  
P( PERSON | NP John Smith, CONTEXT ) = ██████  
P( DATE | NP John Smith, CONTEXT ) = █  
P( ORG | NP John Smith, CONTEXT ) = █

- The textual entailment based inference is done in 3 steps: candidate generation, label verbalization (or hypothesis generation) and inference.
- The candidate generation process **identifies all possible valid spans or tuples** for a given task.
- The label verbalization step **generates hypotheses** by filling some predefined templates (verbalizations).
- Lastly, a textual entailment model **performs the inference** and returns the label with higher probability.

## Pipeline

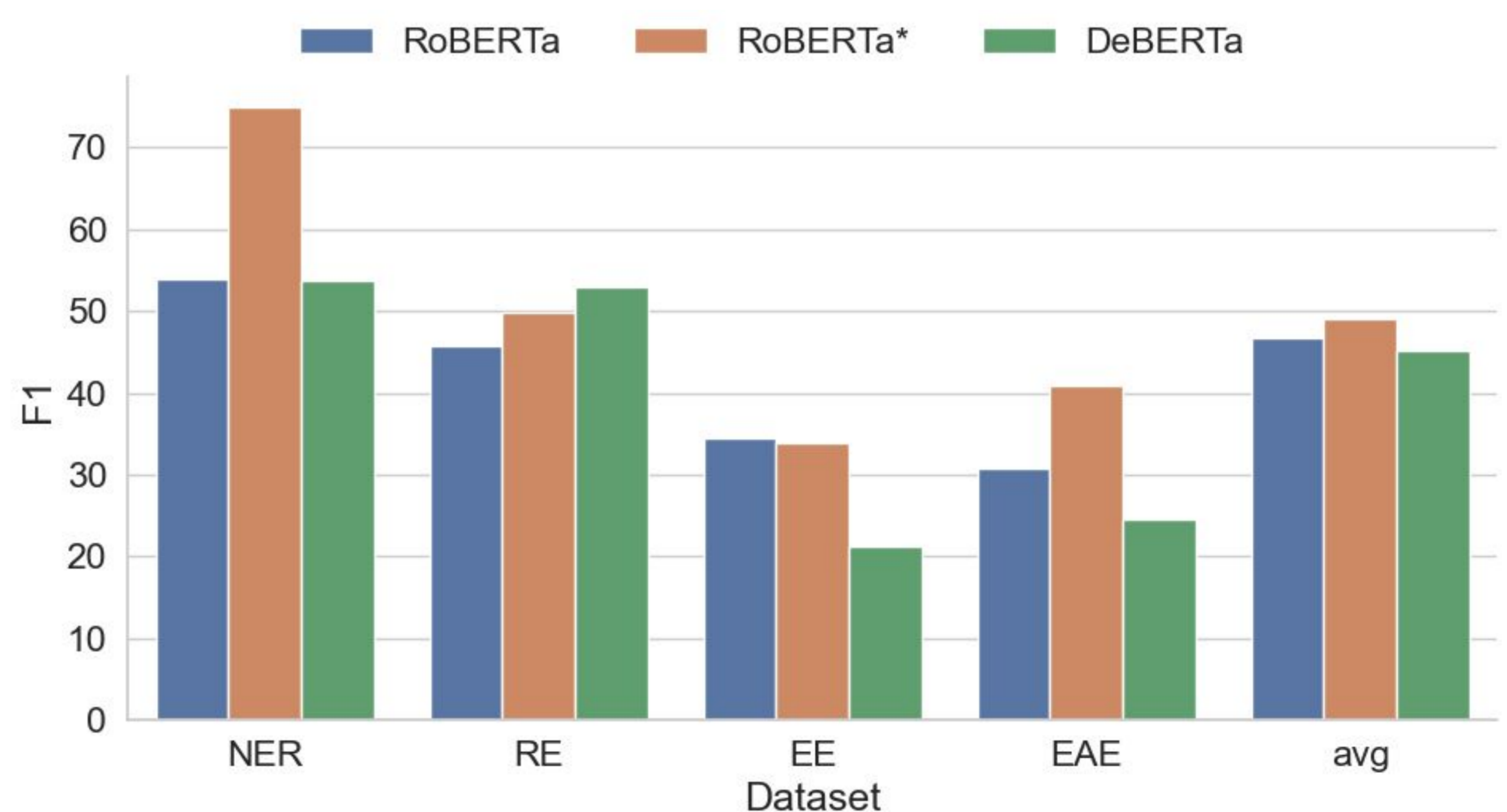
The system runs the following pipeline:



## Results

A zero-shot evaluation was carried on 4 different tasks and datasets:

- NER: CoNLL03 English (Tjong Kim Sang and De Meulder, 2003)
- RE: TACRED (Zhang et al., 2017)
- EE and EAE: ACE05 (Walker et al., 2006)



The above results suggests that verbalizing-while-defining workflow can have similar impact **as post-editing machine translated text**, where:

1. The users define their own schema.
2. Run the system on some unannotated data.
3. Post-edits (fixes the wrong predictions) **with much less effort**.