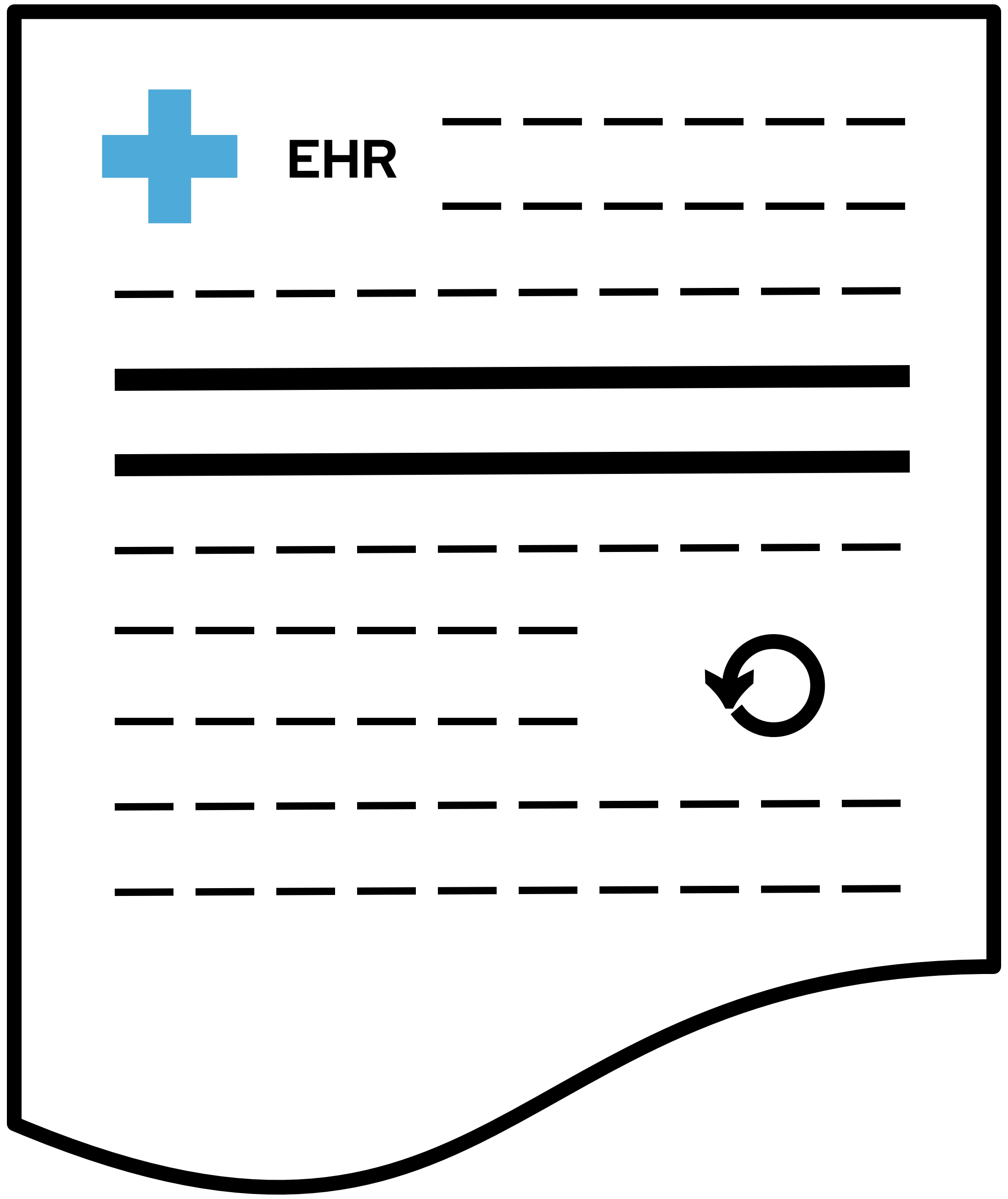
TEXT2STORY WORKSHOP @ ECIR

2ND APRIL 2023

END-TO-END TEMPORAL RELATION EXTRACTION IN THE CLINICAL DOMAIN

JOSÉ JAVIER SAIZ (Barcelona Supercomputing Center), BEGOÑA ALTUNA (HiTZ Center - IXA)

Motivation: Problem background



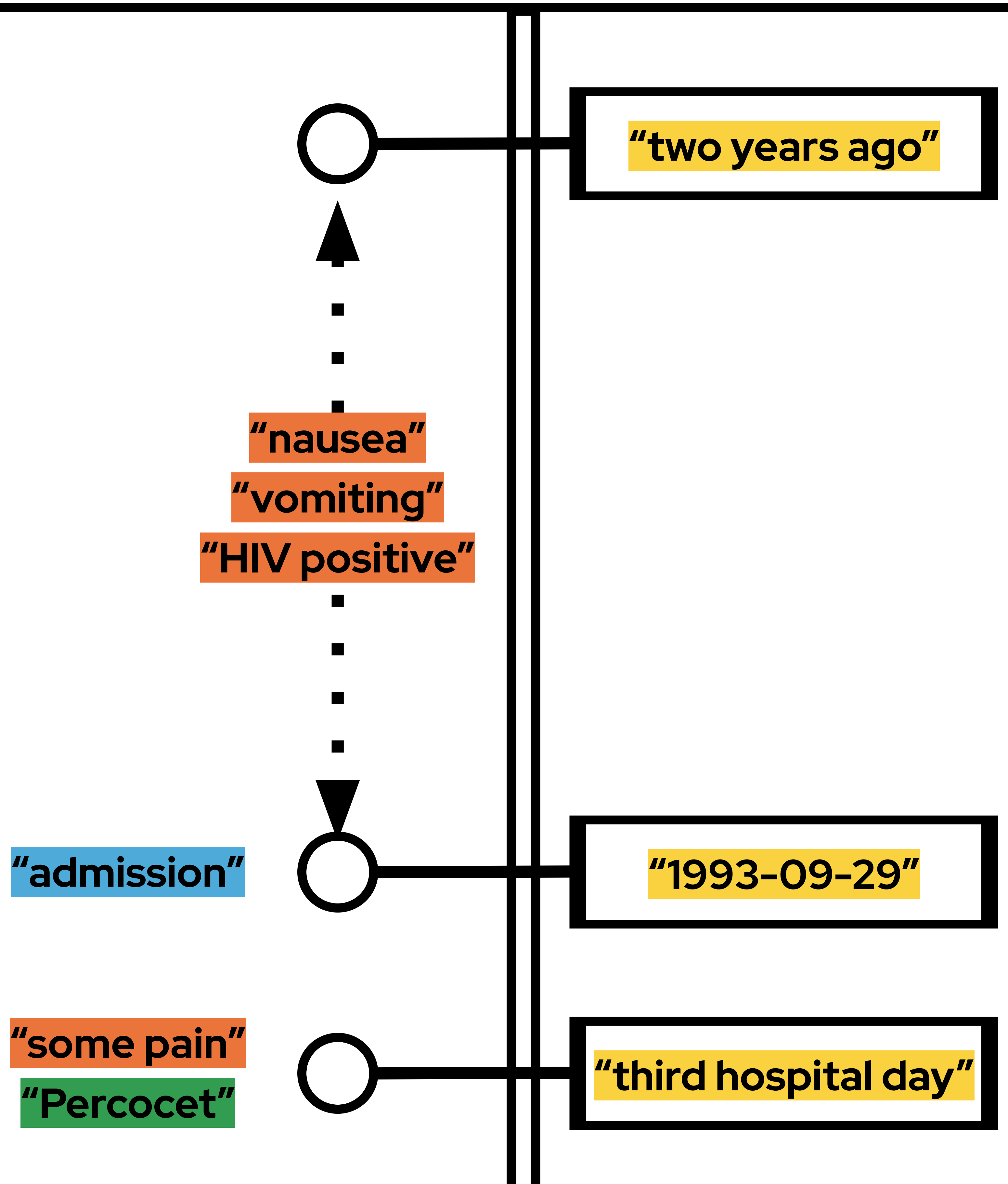
Information overload

Frequent redundancy

Poor accessibility to information ultimately leads to *worse clinical care*.

Clinical narratives are the *most reviewed* data in the clinical workflow because they contain *abundant* and *valuable* information.

Motivation: Problem background

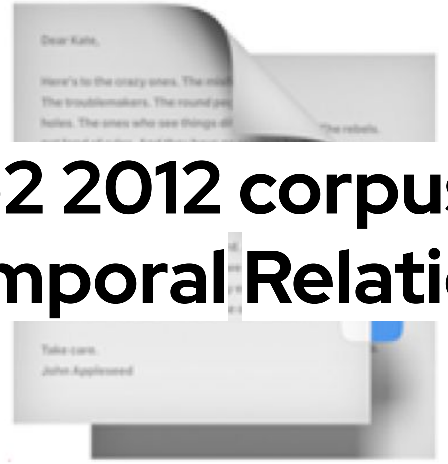


With **Temporal Relation Extraction (TRE)**, we can retrieve useful timelines that require *less effort* and *time* to assimilate the information.

Motivation: Previous approaches

	Other clinical TRE approaches	Our clinical TRE approach
Task	Focus on Relation extraction (Haq et al. 2021, Lin et al. 2020, Zhou et al. 2020)	End-to-End: Entity extraction + Relation extraction
Relation types	Explicit, intra-sentence temporal relations (Alfattni et al. 2021, Guan et al. 2020)	Explicit and implicit + Intra-sentence and cross-sentence temporal relations

Motivation: Our proposal



i2b2 2012 corpus: Temporal Relations

REBEL* pre-trained model (BART-based)

Fine-tuning

Fine-tuned model for clinical TRE

- Since we use shared model weights from the pre-training phase, fine-tuning allows to adapt a general-domain model with very *low resources and time*.

*Relation Extraction By End-to-end Language generation (Cabot et al. 2021)

Methodology: Dataset

EVENT types

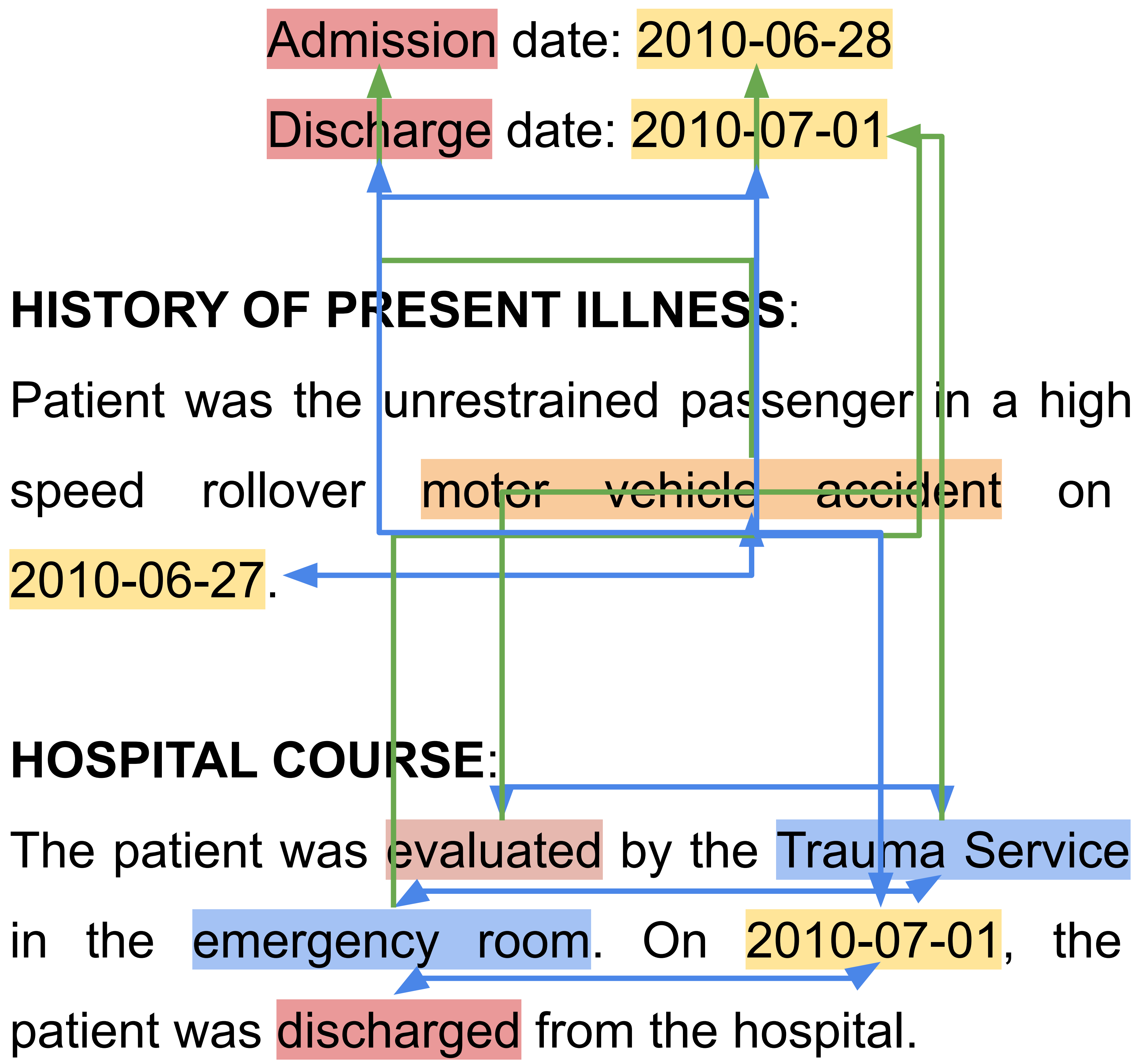
PROBLEM
TEST
TREATMENT
DEPARTMENT
OCCURRENCE
EVIDENTIAL

TIMEX3 types

DATE
TIME
DURATION
FRQUENCY

TLINK types

BEFORE
AFTER
OVERLAP



The i2b2 2012 corpus contains 310 discharge summaries annotated with *clinical events*, *temporal relations*, and *time expressions*.

Methodology: Problem formulation

Given a clinical document x containing a set of n temporal relations, we want to maximize the probability of generating the desired sequence of relations y autoregressively.

$$\hat{y} = \arg\max_{y} p(y \mid x) = \arg\max_{y} \prod_{i=1}^{n} p(y_i \mid y_{<i}, x)$$

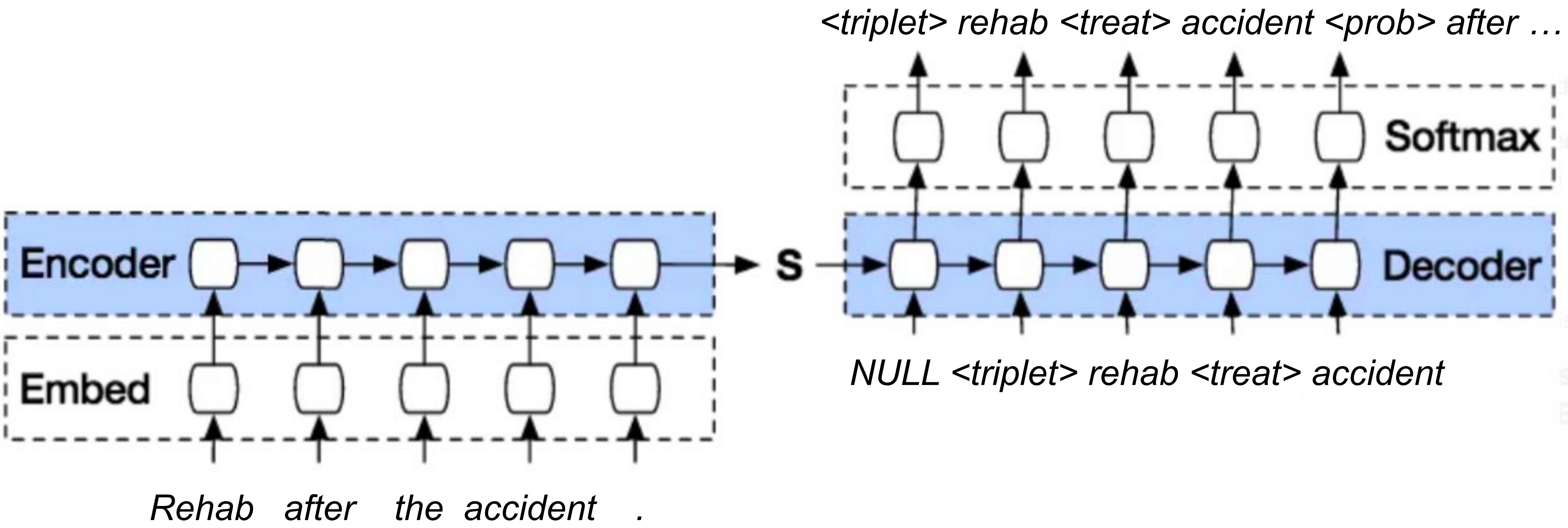
Any relation y_j is represented like a 6-tuple of tokens, such as:

$$y_j = \{ y_{j1}, y_{j2}, y_{j3}, y_{j4}, y_{j5}, y_{j6}, \ldots \}$$

<triplet> *accident* <prob> *rehab* <treat> before …

x = "The patient had an *accident* and went to *rehab*."

Methodology: Model overview

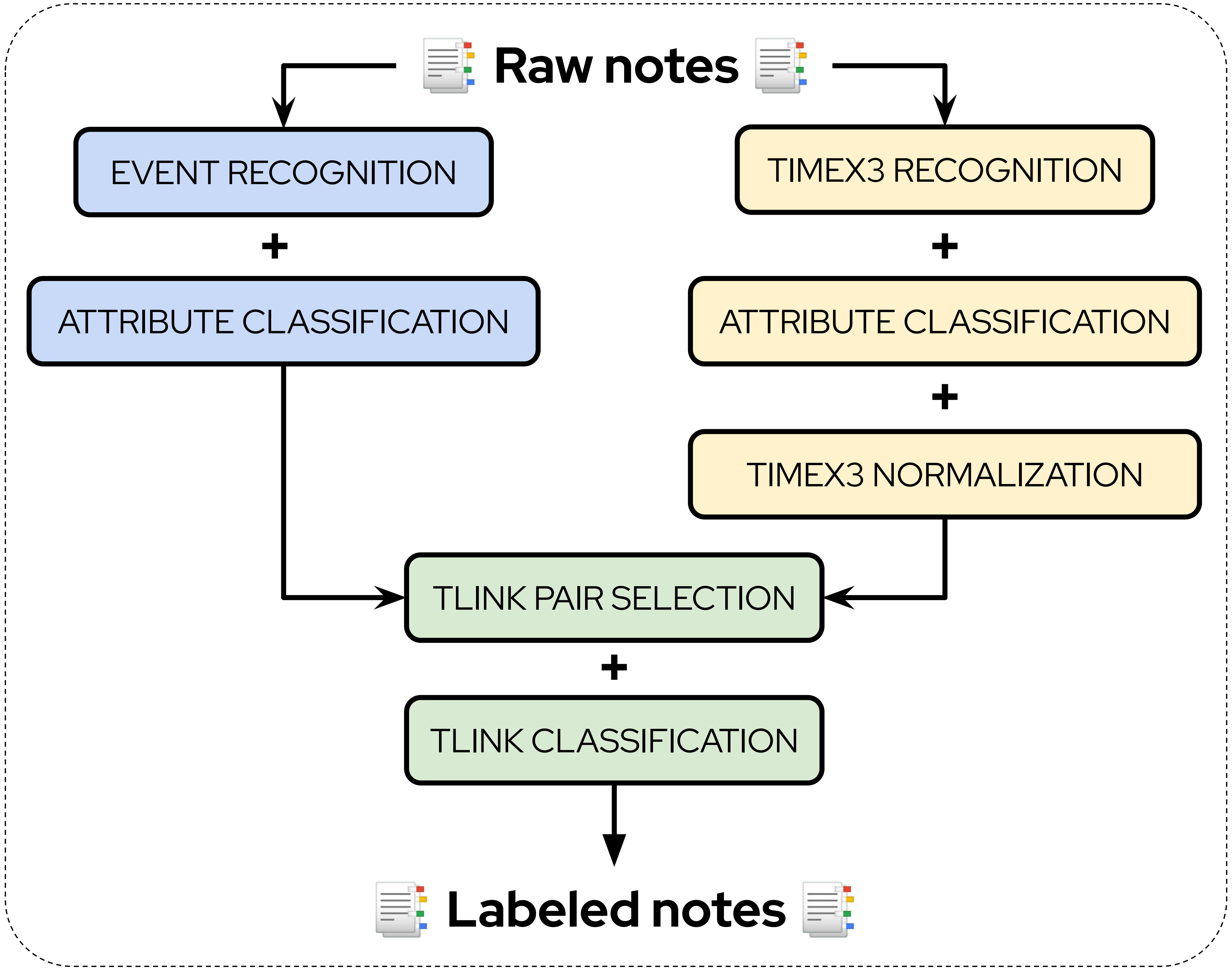


We are able to extract entities, time expressions and relations in one pass.

Experimentation: Evaluation

Baseline: Best systems in the i2b2 2012 challenge (End-To-End Track)

OVERVIEW OF A **TRE** PIPELINE APPROACH



Evaluation method: TempEval3

F-score, **Precision**, **Recall** for *time relations* are the main metrics and are calculated based on *temporal awareness*.

$$Precision = \frac{|Sys^{-}_{relation} \cap Ref^{+}_{relation}|}{|Sys^{-}_{relation}|}$$

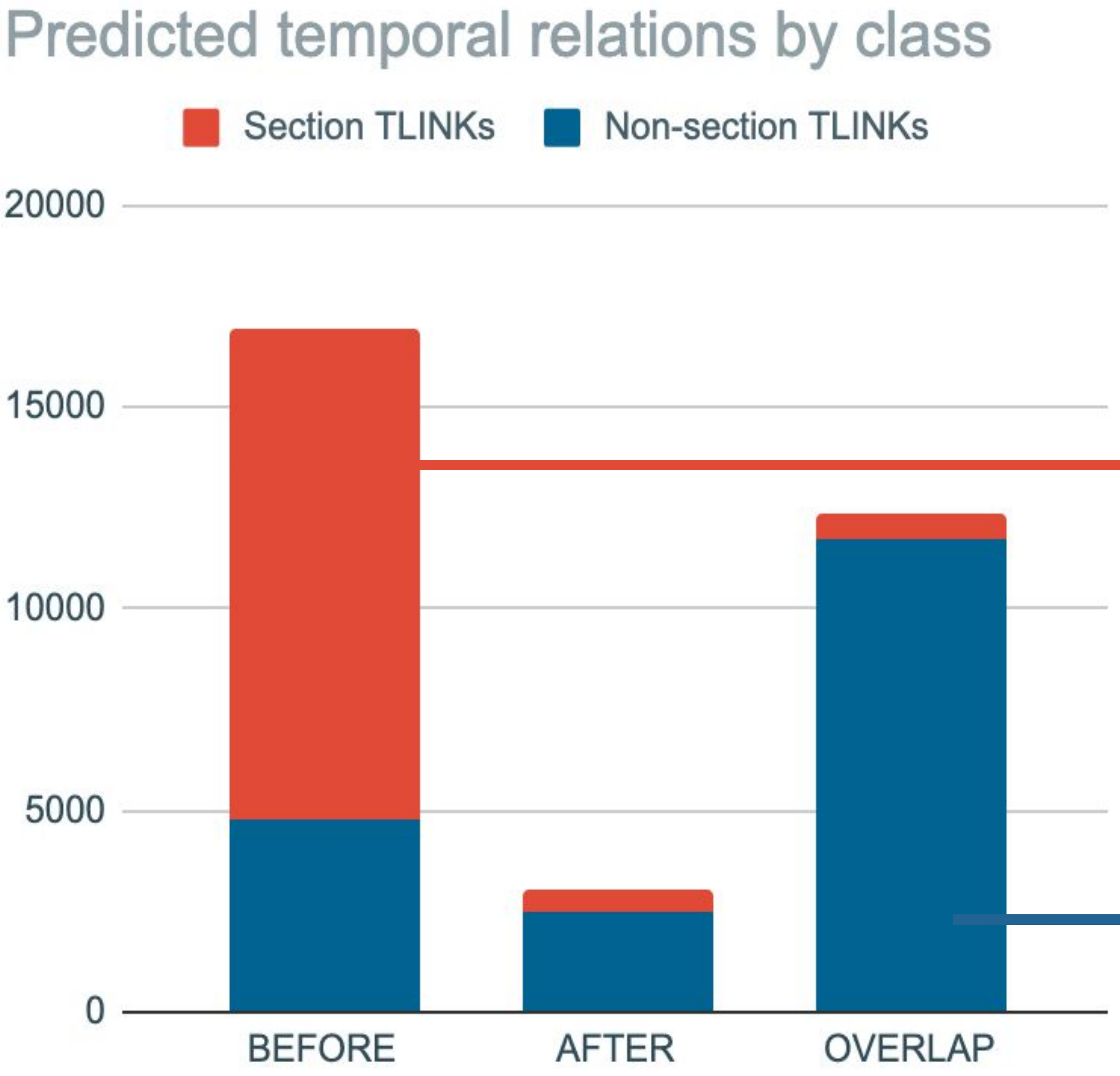
$$Recall = \frac{|Ref^{-}_{relation} \cap Sys^{+}_{relation}|}{|Ref^{-}_{relation}|}$$

Experimentation: Results

System	TLINK			EVENT		TIMEX3	
	F1	P	R	Span F1	Type P	Span F1	Type P
Tang et al., 2013	**0.63**	**0.7**	0.57	0.9	0.84	0.87	0.85
Xu et al., 2013	0.59	0.59	**0.59**	**0.92**	**0.86**	**0.91**	**0.88**
Our proposal*	0.58	0.65	0.52	0.78	0.72	0.77	0.65
Roberts et al., 2013	0.53	0.48	0.57	0.89	0.8	0.89	0.78

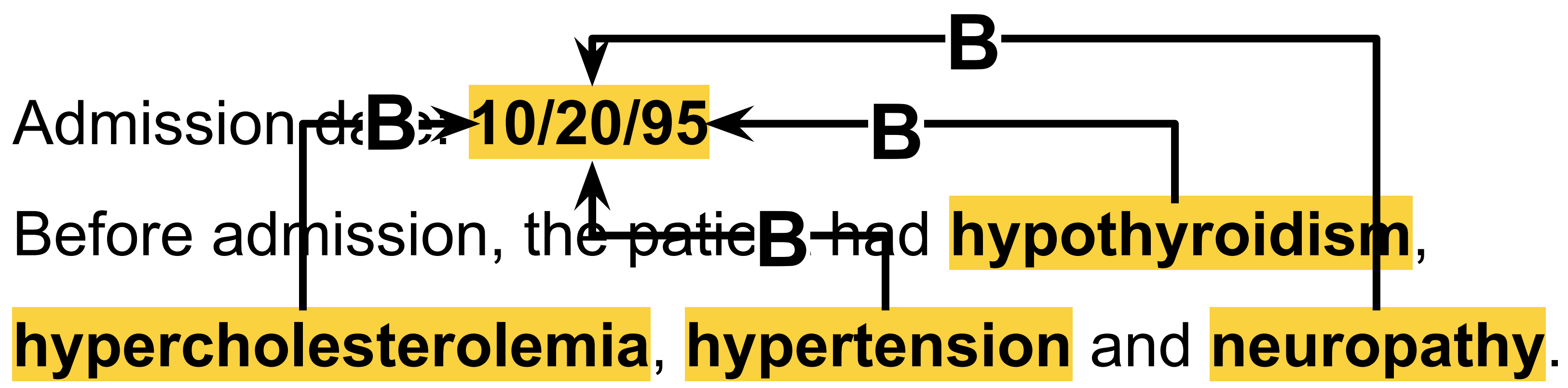
*Results after 10 epochs of fine-tuning on a single Tesla P100 GPU.

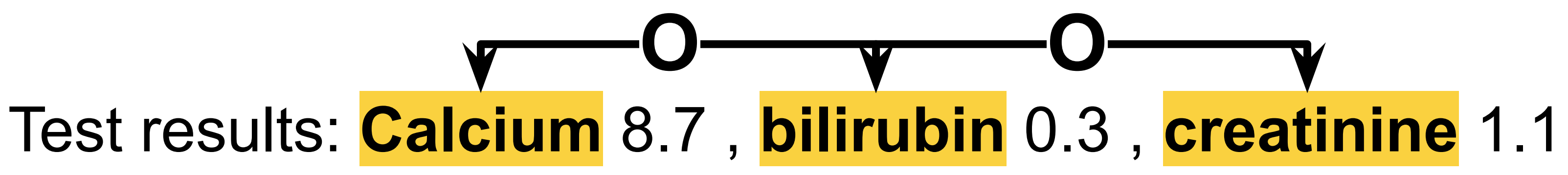
Analysis



Predictions are skewed by the *class imbalance* in the training corpus, where most of the:

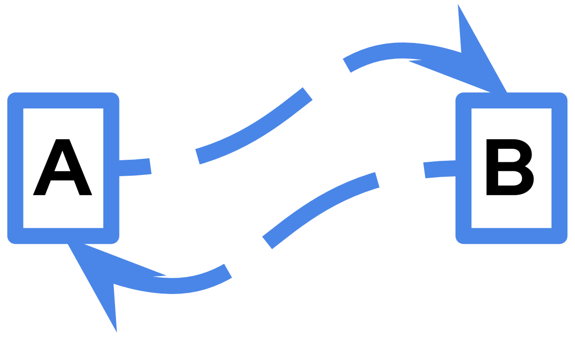
- Section TLINKs are of **B**EFORE type
- Non-section TLINKs are of **O**VERLAP type

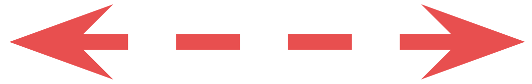
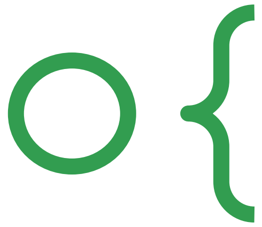


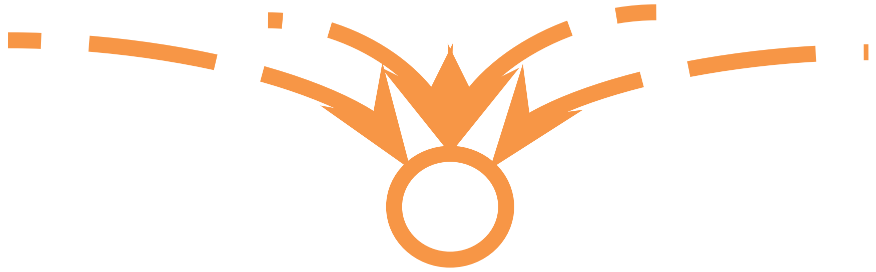


Conclusion & Future work

- Demonstrated that clinical TRE can be performed as a *single task* in a *resource- and time-limited* environment, but ...
- Calculating the **transitive closure** would allow to:



  - mitigate the class imbalance.
  - take into account a higher number of implicit relations. 
  - expand the training corpus. 



- Use a training corpus with **narrative containers**.