

# COMMIT at SemEval-2017 Task 5: Ontology-based Method for Sentiment Analysis of Financial Headlines

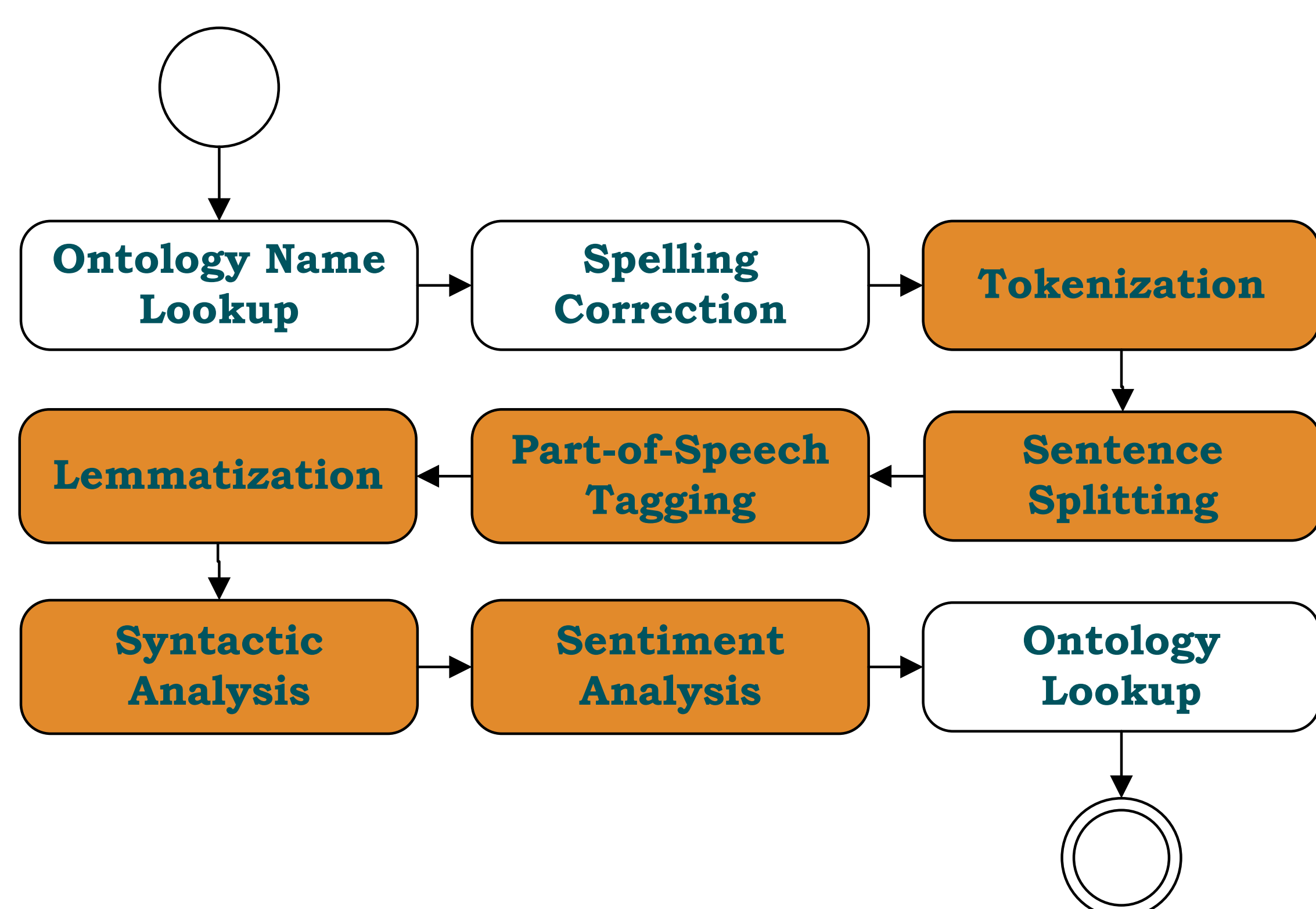
Kim Schouten, Flavius Frasinca, and Franciska de Jong

schouten@ese.eur.nl, frasinca@ese.eur.nl, f.m.g.dejong@eshcc.eur.nl

## Abstract

- > Task 5: Fine-Grained Sentiment Analysis on Financial Microblogs and News
- > This work only addresses track 2: News Headlines
- > Base approach: SVM for regression with unigrams and sentiment features
- > Added feature: ontology features
- > Domain ontology is created
- > Sentiment of actions is dependent on the entity they are affecting

## Preprocessing



## Ontology Axioms

### General

- Increase  $\cap$  PosEntity  $\rightarrow$  Positive
- Increase  $\cap$  NegEntity  $\rightarrow$  Negative
- Decrease  $\cap$  PosEntity  $\rightarrow$  Negative
- Decrease  $\cap$  NegEntity  $\rightarrow$  Positive

### Custom

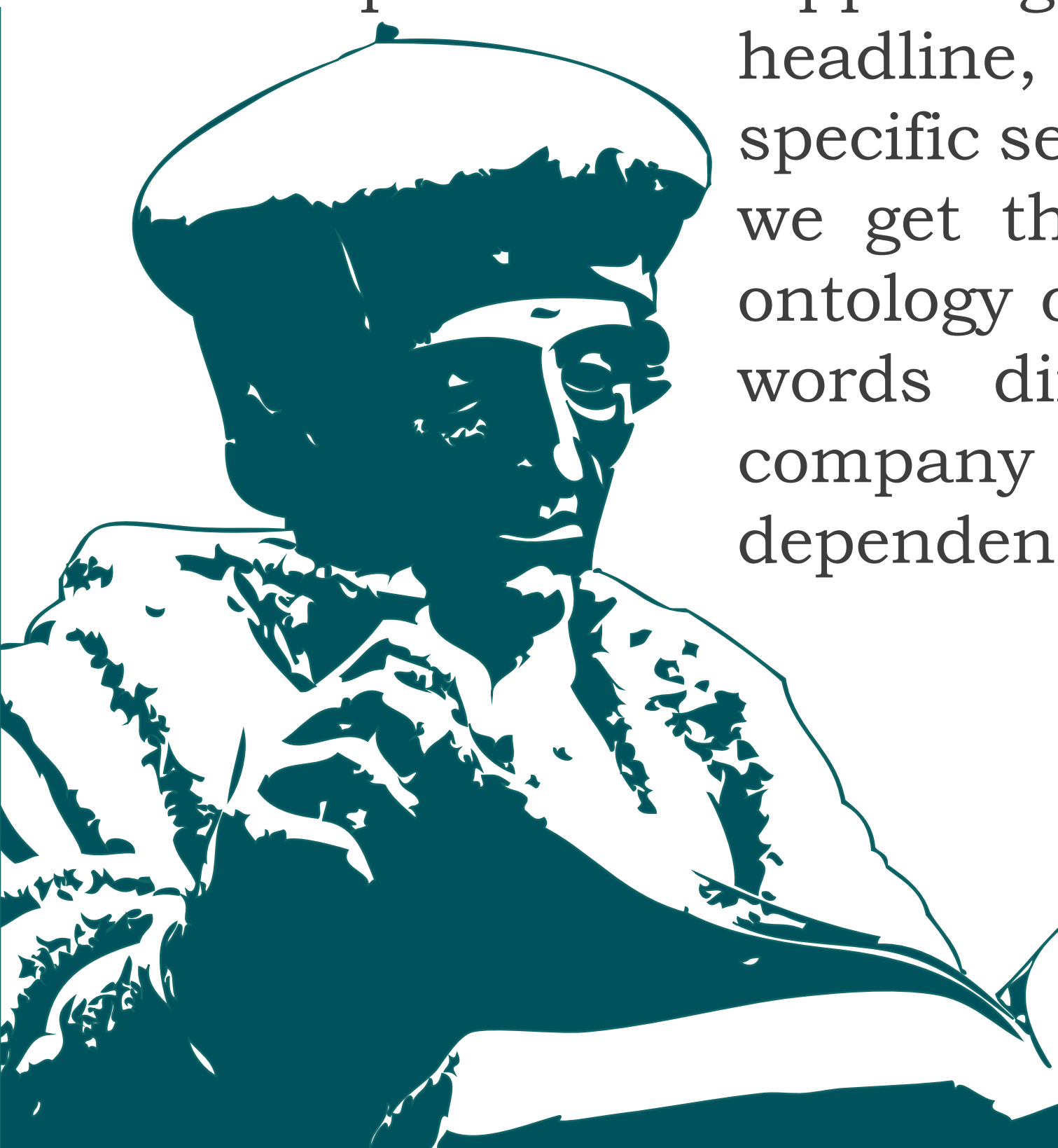
- Close  $\cap$  Deal  $\rightarrow$  Positive
- Close  $\cap$  CompanyPart  $\rightarrow$  Negative
- Open  $\cap$  CompanyPart  $\rightarrow$  Positive

## Ontology Features

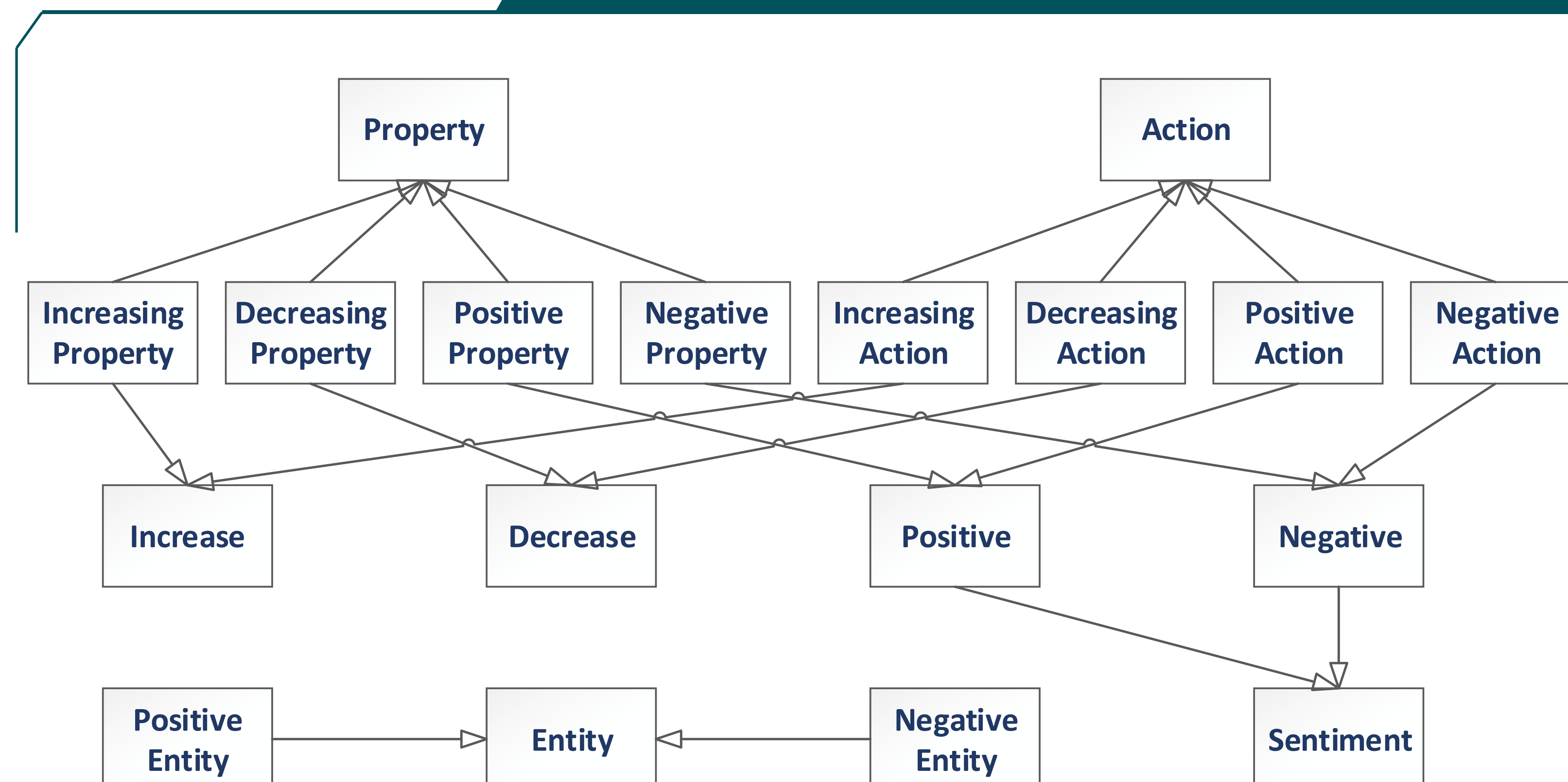
Lemmas in the text are looked up in a manually created domain ontology, and discovered concepts, as well as all superclasses, are added to the feature set (bag model).

When necessary, the reasoner infers the sentiment of concept combinations.

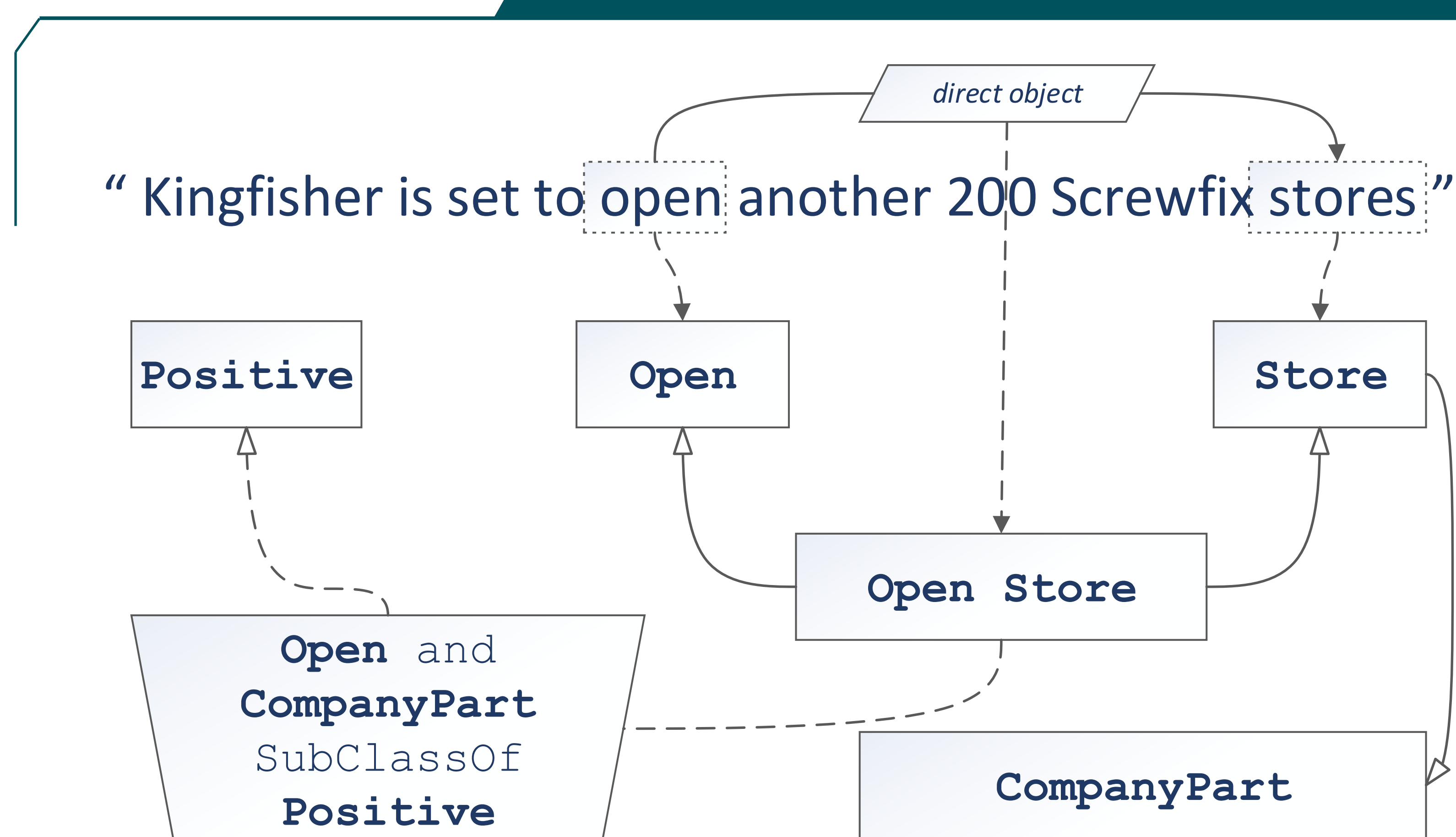
To be able to differentiate between two companies with opposing sentiment in the same headline, we add a company-specific sentiment feature, where we get the sentiment from the ontology of concepts denoted by words directly related to the company mention in the dependency graph.



## Data Snippet



## Example



## Results

Official test data results: cosine distance of **0.6810** (12th position)  
The results below are obtained using 10-fold cross-validation on the training data, to test the effect of the different types of features:

	avg. cosine	
	distance	st.dev.
base (B)	0.6311	0.0482
B + entities (E)	0.6361	0.0455
B + properties (P)	0.6300	0.0478
B + actions (A)	0.6815	0.0498
B + E + P + A	0.6883	0.0502
B + E + P + A + class axioms	0.7041	0.0450
B + E + P + A + class axioms + company-specific sentiment	0.7050	0.0441

