

An Initial Study of Indonesian Semantic Role Labeling and Its Application on Event Extraction



Ade Romadhony
Ayu Purwarianti
Lisa Madlberger

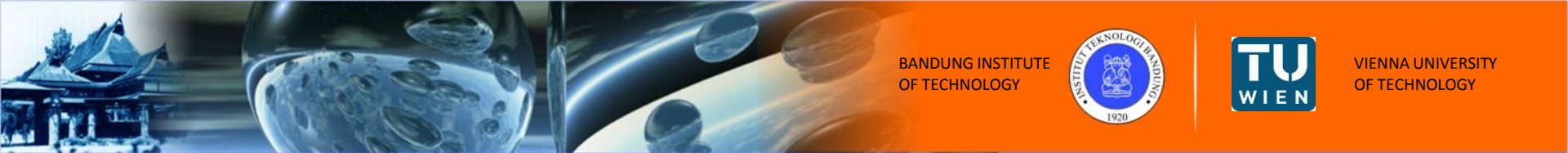


BANDUNG INSTITUTE OF TECHNOLOGY
INDONESIA



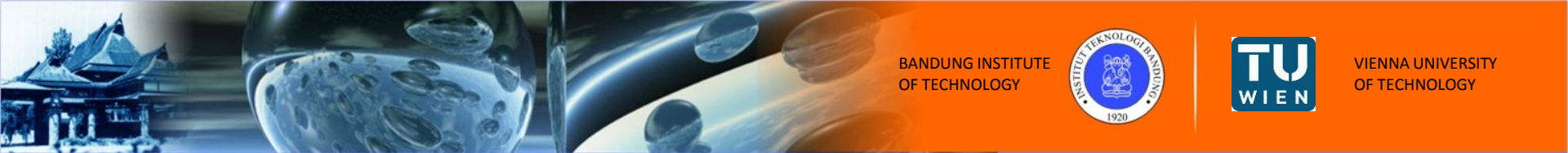
VIENNA UNIVERSITY OF TECHNOLOGY
AUSTRIA





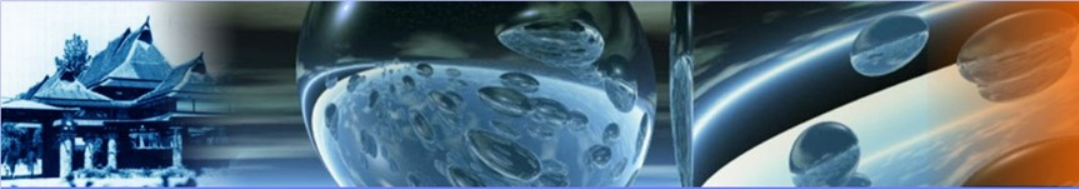
Background : Why Indonesian SRL?

- **Lack of research** on Indonesian Semantic Role Labeling (SRL)
 - Only one previous study about Indonesian SRL (FrameNet notation), built based on the translated sentence example in English FrameNet [Dewi, 2013].
- **Indonesian SRL** is required for several tasks. SRL has been widely used as the foundation, of several applications, as for example event extraction



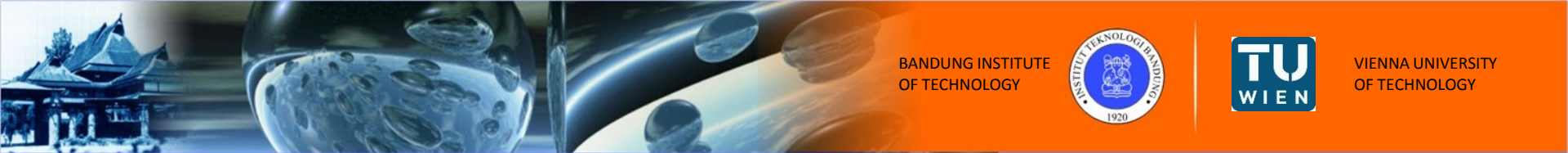
Our Work

*We design and implement an Indonesian SRL system using **Word-to-Word** and **Phrase-to-Phrase** features and employ the model for **extracting strike/demonstration** event attribute information.*



Experimental Setup

- **Feature Design**
- **Dataset Preparation**



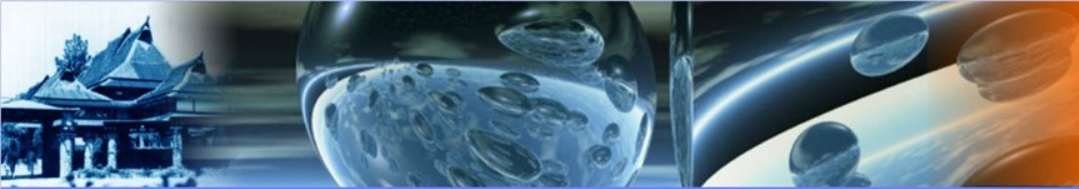
Experimental Setup

- **Feature Design**
 - **Word-to-Word vs Phrase-to-Phrase [Hacioglu et al, 2004]**

Word	POSTag	Phrase	Predicate
<i>Mereka / they</i>	PRP	B-NP	-
<i>menjaga / keep</i>	VBT	B-VP	menjaga
<i>aksi / action</i>	NN	B-NP	-
<i>ini / this</i>	DT	I-NP	-
<i>agar / to</i>	SC	B-SC	-
<i>berjalan / run</i>	VBI	B-VP	-
<i>tertib / orderly</i>	JJ	B-ADJP	-

Word-to-Word





Experimental Setup

- **Feature Design**
 - **Word-to-Word vs Phrase-to-Phrase [Hacioglu et al, 2004]**

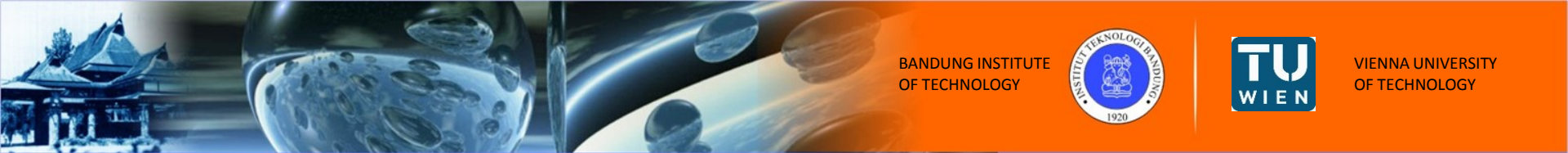
Word	POSTag	Phrase	Predicate
<i>Mereka / they</i>	PRP	B-NP	-
<i>menjaga / keep</i>	VBT	B-VP	menjaga
<i>aksi / action</i>	NN	B-NP	-
<i>ini / this</i>	DT	I-NP	-
<i>agar / to</i>	SC	B-SC	-
<i>berjalan / run</i>	VBI	B-VP	-
<i>tertib / orderly</i>	JJ	B-ADJP	-

Word-to-Word

Word	POSTag	Phrase	Predicate
<i>Mereka / they</i>	PRP	B-NP	-
<i>menjaga / keep</i>	VBT	B-VP	menjaga
<i>aksi / action</i>	NN	B-NP	-
<i>berjalan / run</i>	VBI	B-VP	-
<i>tertib / orderly</i>	JJ	B-ADJP	-

Phrase-to-Phrase





Experimental Setup

- Feature Design
 - Word-to-Word vs Phrase-to-Phrase [Hacioglu et al, 2004]

Word	POSTag	Phrase	Predicate
<i>Mereka / they</i>	PRP	B-NP	-
<i>menjaga / keep</i>	VBT	B-VP	menjaga
<i>aksi / action</i>	NN	B-NP	-
<i>ini / this</i>	DT	I-NP	-
<i>agar / to</i>	SC	B-SC	-
<i>berjalan / run</i>	VBI	B-VP	-
<i>tertib / orderly</i>	JJ	B-ADJP	-

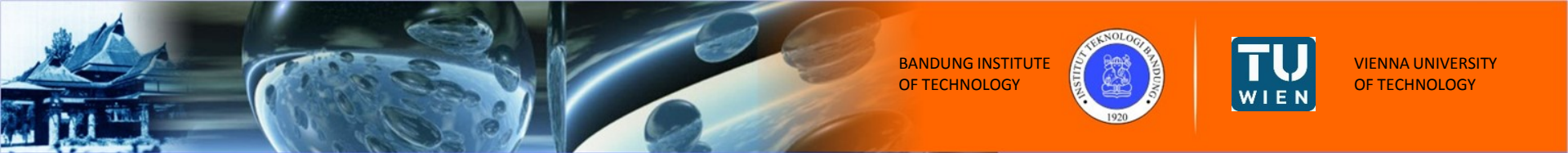
Word-to-Word

Word	POSTag	Phrase	Predicate
<i>Mereka / they</i>	PRP	B-NP	-
<i>menjaga / keep</i>	VBT	B-VP	menjaga
<i>aksi / action</i>	NN	B-NP	-
<i>berjalan / run</i>	VBI	B-VP	-
<i>tertib / orderly</i>	JJ	B-ADJP	-

Headword
[Dewi, 2013]

Phrase-to-Phrase





Experimental Setup

- **Feature Design**

- **Word-to-Word vs Phrase-to-Phrase**

Aside from the base features, we also extracted the following features:

- **Word-to-Word**

- words window, token position, distance

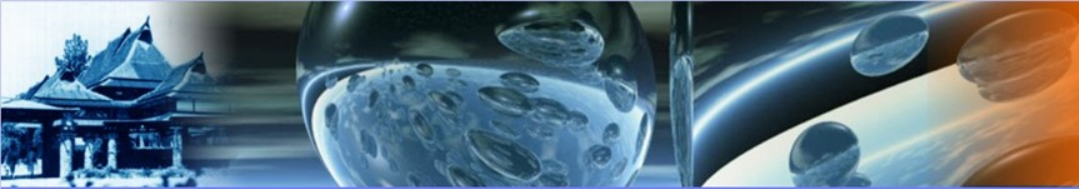
- **Phrase-to-Phrase**

- headword, preposition, voice, position



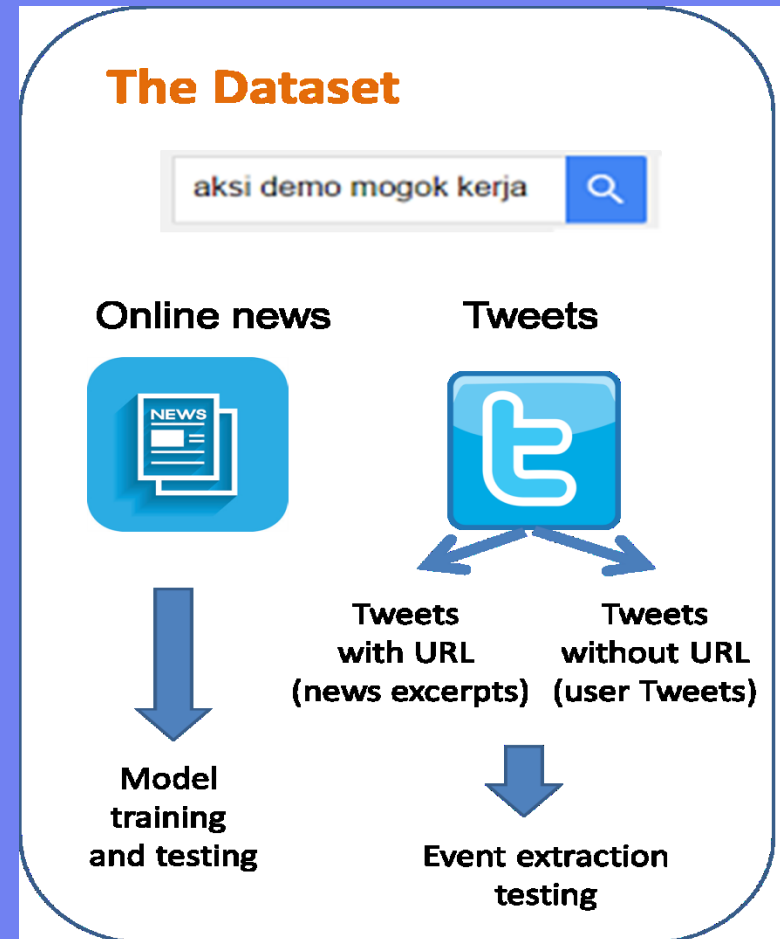
Experimental Setup

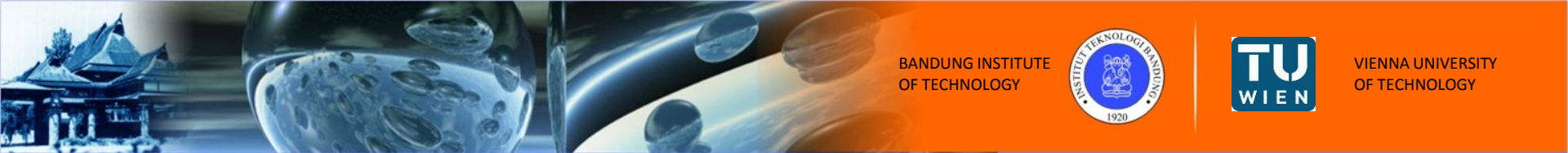
- **Feature Design**
 - **Word-to-Word vs Phrase-to-Phrase [Hacioglu et al, 2004]**
 - **Base features obtained using Indonesian language processing tools:**
 - **POSTag**
 - **Phrase Tag**
 - **Named Entity Tags**



Experimental Setup

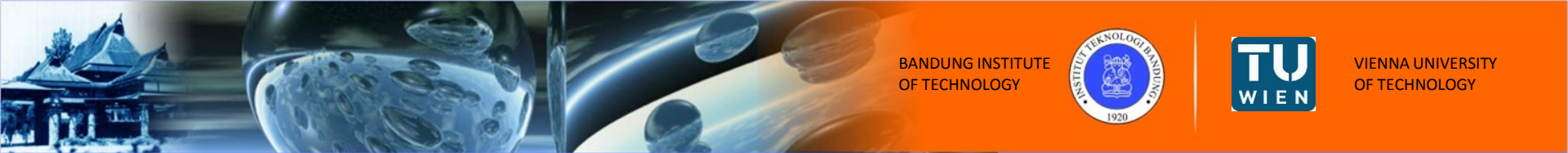
- Feature Design
- Dataset Preparation
 - News dataset: for training
 - Tweet dataset: for testing on event extraction





Experimental Setup

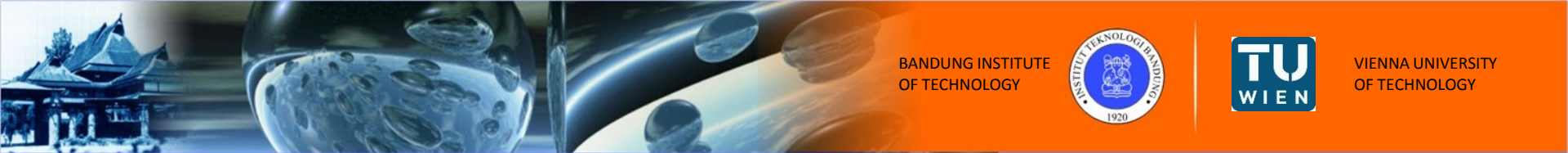
- **Feature Design**
- **Dataset Preparation**
 - **News dataset: for training**
 - News articles with strike/demonstration topic from various website in May-June 2015 period.
 - Select sentences that contains words related with the topic, e.g: *unjuk rasa / protest, demo / demonstration, mogok kerja / strike.*
 - Process the selected sentences with the Indonesian language processing tools.
 - Manually annotated the Semantic Role Labels according to the PropBank-styled annotation guideline [Carreras and Marquez, 2005]



Experimental Setup

- Feature Design
- Dataset Preparation
 - News dataset: for training
 - Tweet dataset: for testing on event extraction
 - Retrieved Tweet matching with demo/strike topic using similar keywords used in News dataset, in similar period (May-June 2015)
 - Divided based on the existence of url: News Excerpt and User Tweets
 - Performed preprocessing to obtain cleaner text



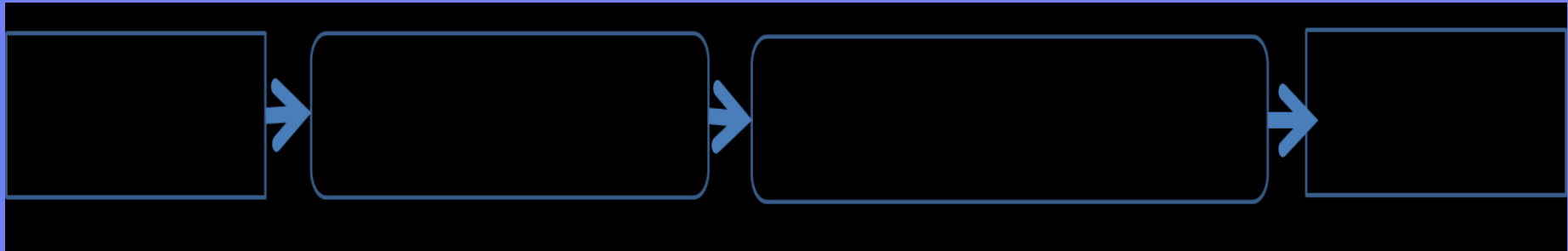


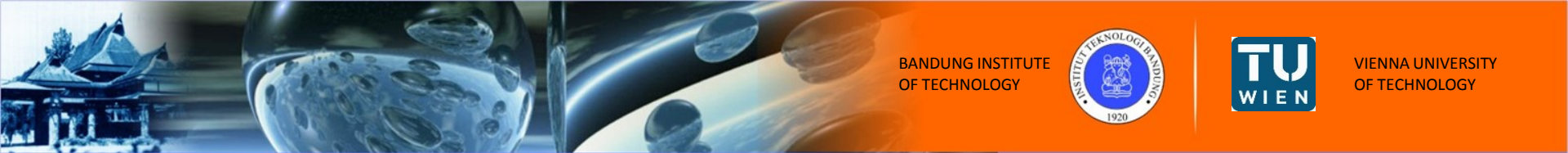
Experiment



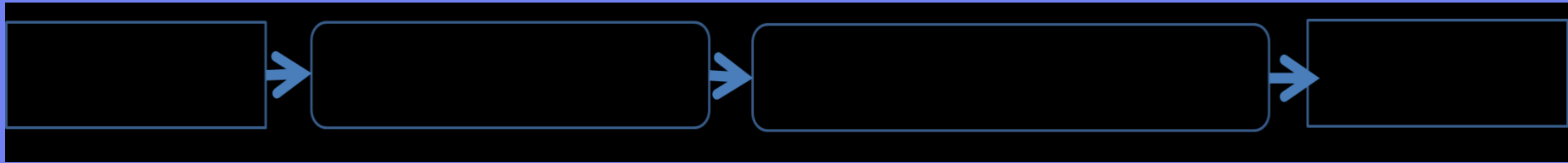


Event Extraction



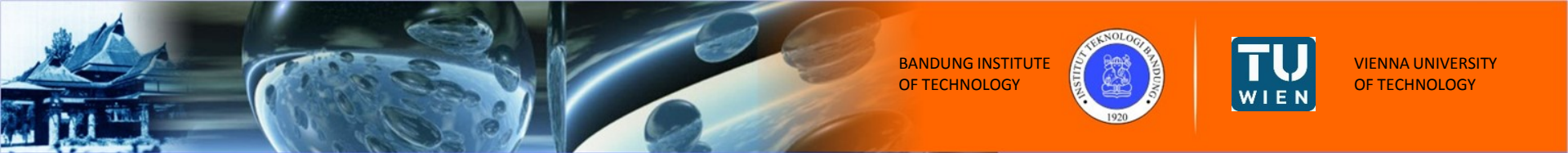


Event Extraction

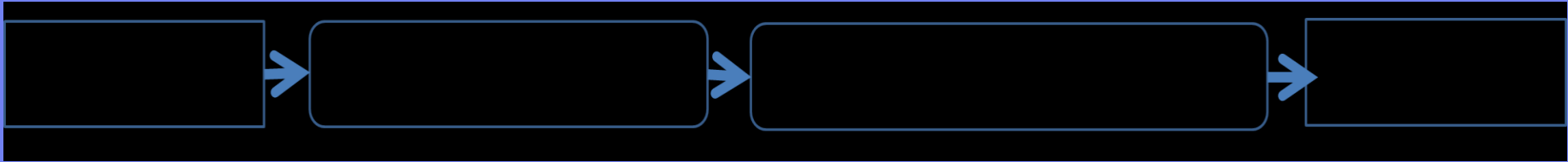


Example Input:

Banter melakukan aksi demo di Bandung / Banter is doing demonstration in Bandung

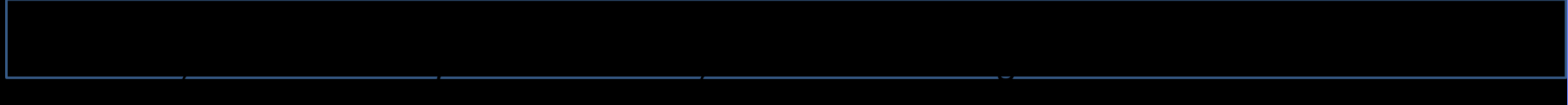


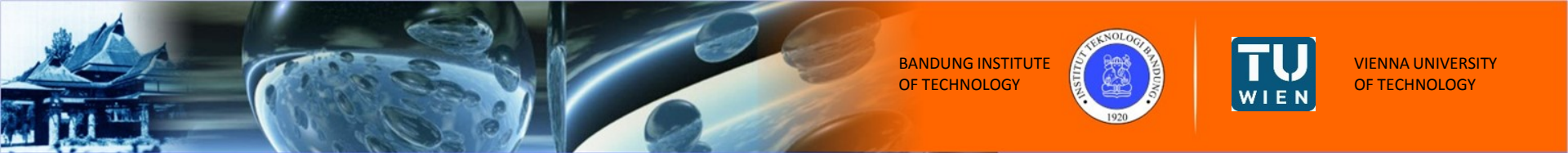
Event Extraction



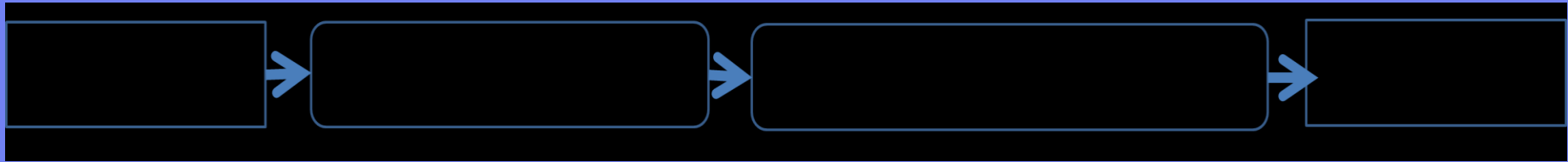
Example Input:

Banter melakukan aksi demo di Bandung / Banter is doing demonstration in Bandung



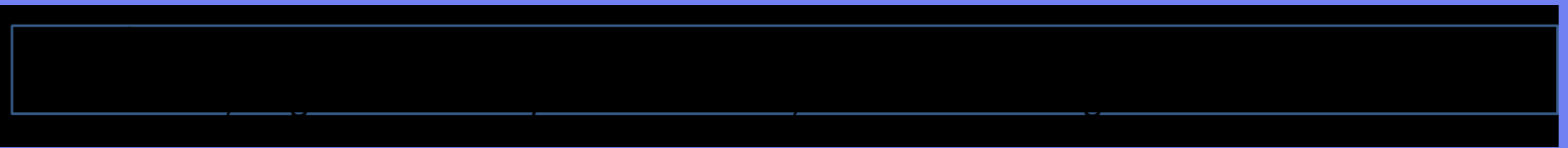
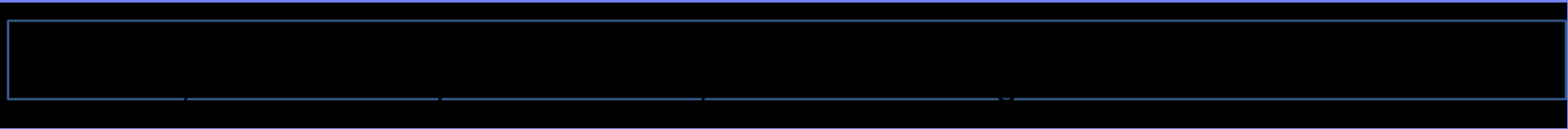


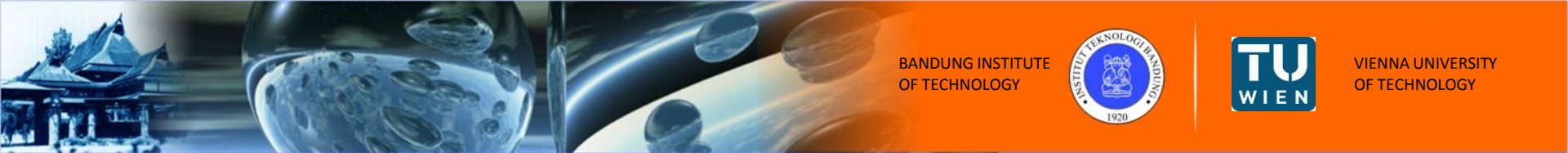
Event Extraction



Example Input:

Banter melakukan aksi demo di Bandung / Banter is doing demonstration in Bandung





Experimental Result: SRL Model Evaluation

Finding: in several cases, the NE feature could not be determined by the NE Tagger.

Word-to-Word vs Phrase-to-Phrase:

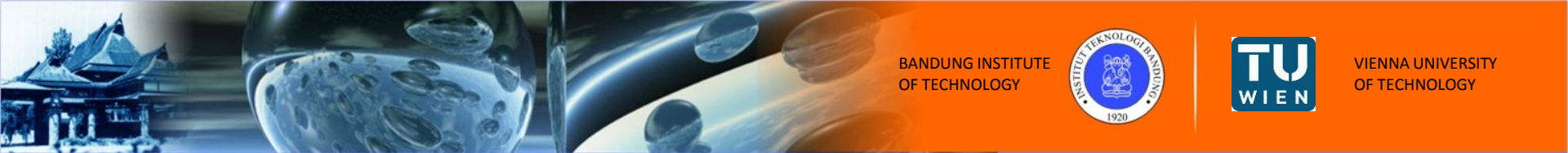
a. With Named Entity feature

	P-W	P-P	R-W	R-P	F1-W	F1-P
A0	0.82	0.74	0.75	0.62	0.78	0.67
A1	0.72	0.64	0.81	0.76	0.76	0.69
AM-LOC	0.75	0.74	0.72	0.42	0.73	0.54
AM-TMP	0.99	1.00	0.44	0.13	0.61	0.23

b. Without Named Entity feature

	P-W	P-P	R-W	R-P	F1-W	F1-P
A0	0.81	0.76	0.71	0.64	0.76	0.69
A1	0.71	0.61	0.82	0.78	0.76	0.68
AM-LOC	0.76	0.80	0.68	0.38	0.72	0.52
AM-TMP	1.00	1.00	0.38	0.03	0.55	0.06





Experimental Result: Tweet Event Extraction Evaluation

Metric: Accuracy =
$$\frac{\# (\text{overlap} \mid \text{exact match}) \text{ SRL Labels}}{\# \text{ Gold Labels}}$$

Number of Tweets Having Gold Labels

Tweet Types	Number of Gold Labels			
	Actor	Target	Location	Time
News Excerpt	82	28	19	10
User Tweet	25	3	13	8

Example of Overlap and Exact Match Condition

TweetId	Gold Actor	SRL-A0	#Overlap	Exact Match
1	<i>Dua orang</i> / two persons	<i>Dua orang</i> / two persons	2	yes
2	Sekuriti UIN / UIN security	<i>Sekuriti UIN ancam</i> / UIN security threatened	2	no



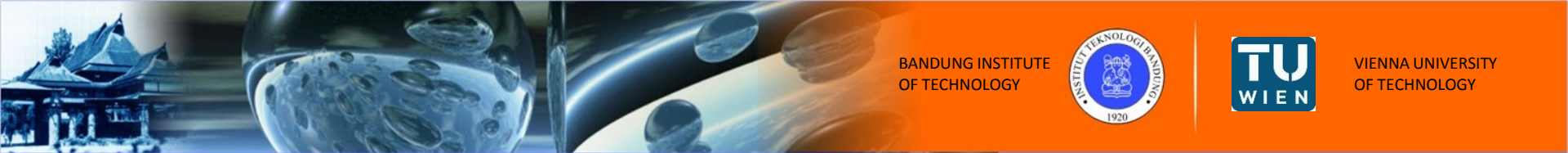
Experimental Result: Tweet Event Extraction Evaluation

Evaluation on News Excerpt Dataset:

Feature	Actor		Target		Location		Time	
	OV	EM	OV	EM	OV	EM	OV	EM
W2W-NE	0.54	0.15	0.50	0.00	0.32	0.11	0.40	0.00
W2W-WONE	0.94	0.17	0.89	0.00	0.79	0.32	0.47	0.00
P2P-NE	0.32	0.04	0.36	0.00	0.26	0.00	0.30	0.00
P2P-WONE	0.63	0.06	0.68	0.00	0.74	0.00	0.60	0.00

Evaluation on User Tweet Dataset:

Feature	Actor		Target		Location		Time	
	OV	EM	OV	EM	OV	EM	OV	EM
W2W-NE	0.32	0.04	0.67	0.00	0.15	0.08	0.50	0.00
W2W-WONE	0.68	0.04	0.67	0.00	0.62	0.15	0.88	0.00
P2P-NE	0.20	0.00	0.67	0.00	0.08	0.00	0.25	0.00
P2P-WONE	0.48	0.00	0.67	0.00	0.38	0.00	0.63	0.00



Conclusion and Future Work

- W2W approach outperforms the P2P approach
- The performance of language processing tools (ie., NER) needs to be improved
- Challenge in Indonesian SRL: multi-word predicate, light verb construction
- The experiment could be expanded using a general domain and larger dataset



BANDUNG INSTITUTE OF TECHNOLOGY
INDONESIA

Thank You



Research reported in this publication was jointly supported by the ASEAN-European Academic University Network (ASEA-UNINET), the Austrian Federal Ministry of Science, Research and Economy and the Austrian Agency for International Cooperation in Education and Research (OeAD-GmbH). We also thank Mochamad Ibrahim and the University of Indonesia for their cooperation during this project.



Imam Santosa © LPPM ITB 2011



VIENNA UNIVERSITY OF TECHNOLOGY
AUSTRIA

