Natural Language Inference for Dutch

Natural language inference (NLI):

- Identify semantic relations (entails, contradicts, is neutral)

A group of scouts are hiking through the grass
Some people are walking

A hiker is on top of the mountain and is dancing
There is no hiker dancing on top of the mountain

The black woman is wearing glasses over the headdress
A woman is wearing an Egyptian headdress

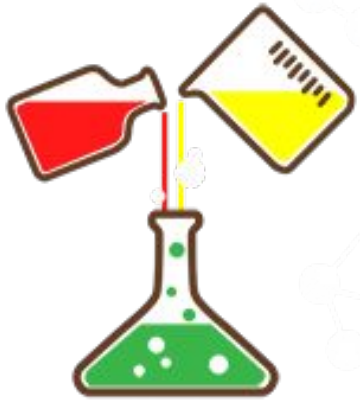
- Central manifestation of semantic competence
- Popular NLP task (FraCas, RTE-N, SICK, SNLI, ...)
- Under-researched for Dutch

To do (1): create NLI dataset in Dutch

How?	By translating an existing one
Which?	SICK (Sentences Involving Compositional Knowledge)
Why?	Offers a good starting point
Time & efforts?	Not much, 10K sen (6K diff) (e.g., SICK-BR)

A group of scouts are camping in the grass.	Een groep verkenners kampeert in het gras.
A group of scouts are hiking through the grass.	Een groep verkenners wandelt door het gras.
Some people are walking	Sommige mensen lopen
People are walking	Mensen lopen
Some people are walking	Sommige mensen lopen
A group of scouts are camping in the grass	Een groep verkenners kampeert in het gras
A group of scouts are hiking through the grass	Een groep verkenners wandelt door het gras
Some people are walking	Sommige mensen lopen

To do (2): work with the data



Several possibilities:

- Create strong baselines
- Reproduce results of related work (SemEval-2014: task 1)
- Design your own reasoning monster

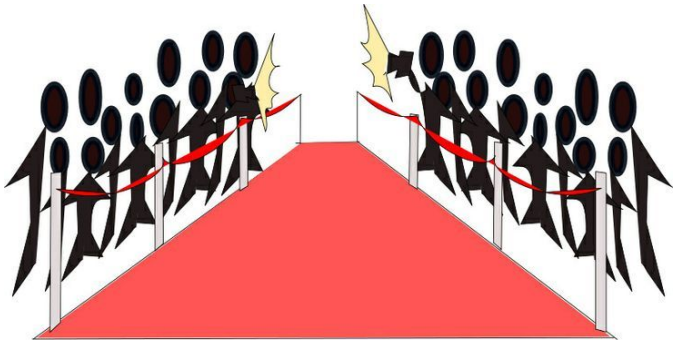
Why should I do it?

- You contribute to Dutch language technology
- You work with a generic task which has many applications
- You create a dataset that other will use

contact: Lasha Abzianidze (1311.436)

Why should I do it?

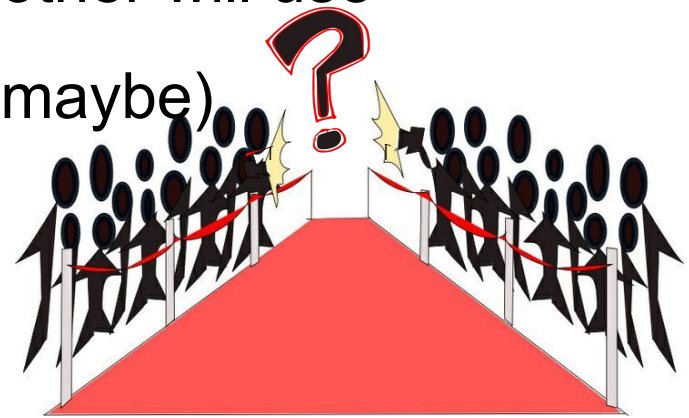
- You contribute to Dutch language technology
- You work with a generic task which has many applications
- You create a dataset that other will use
- You will become famous



contact: Lasha Abzianidze (1311.436)

Why should I do it?

- You contribute to Dutch language technology
- You work with a generic task which has many applications
- You create a dataset that other will use
- You will become famous (maybe)



contact: Lasha Abzianidze (1311.436)