

Dialog Act Annotation for Twitter Conversations

Elina Zarisheva

Hasso-Plattner-Institut

elina.zarisheva@student.hpi.uni-potsdam.de

Tatjana Scheffler

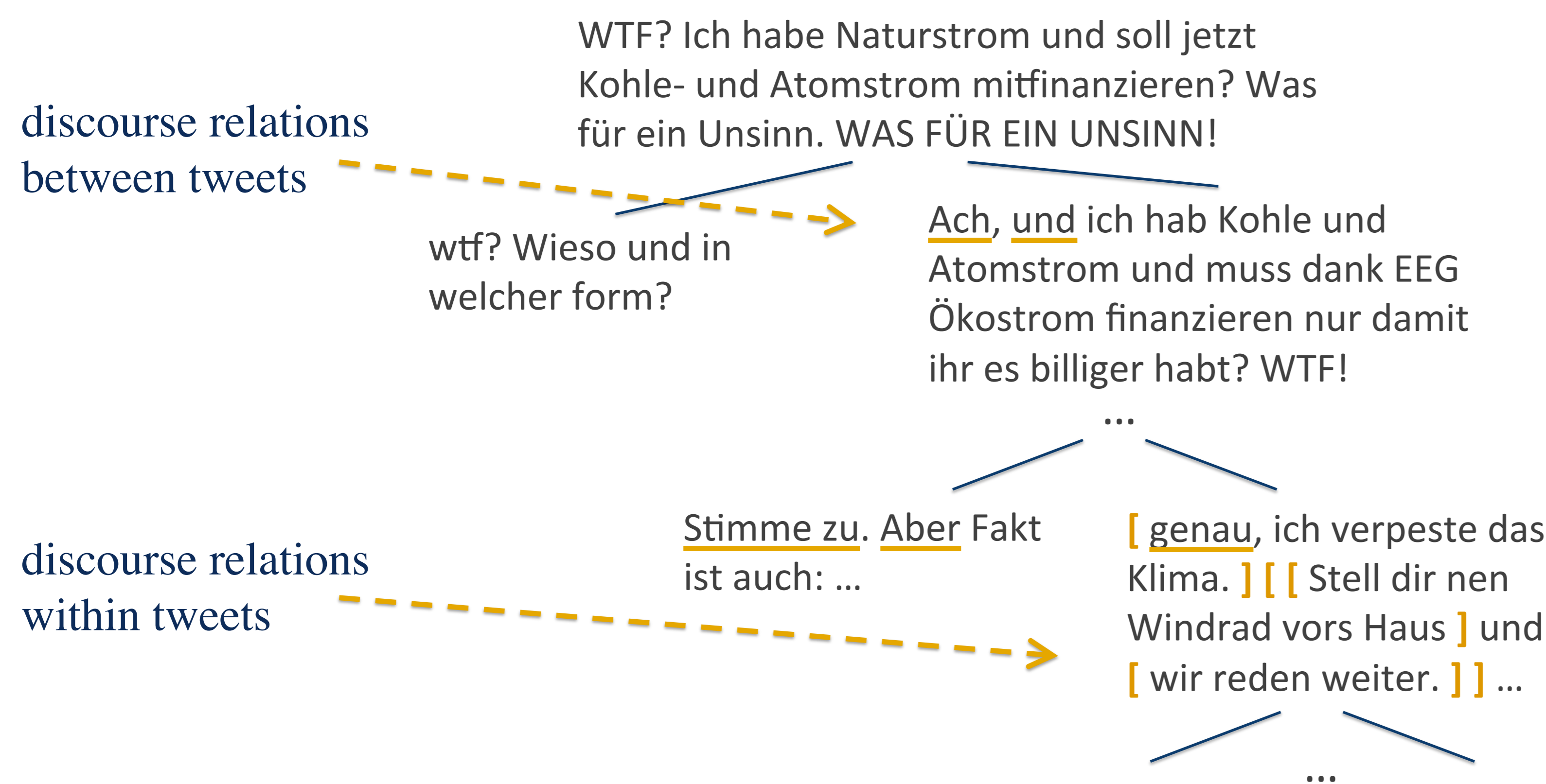
Universität Potsdam

tatjana.scheffler@uni-potsdam.de

@tschfflr

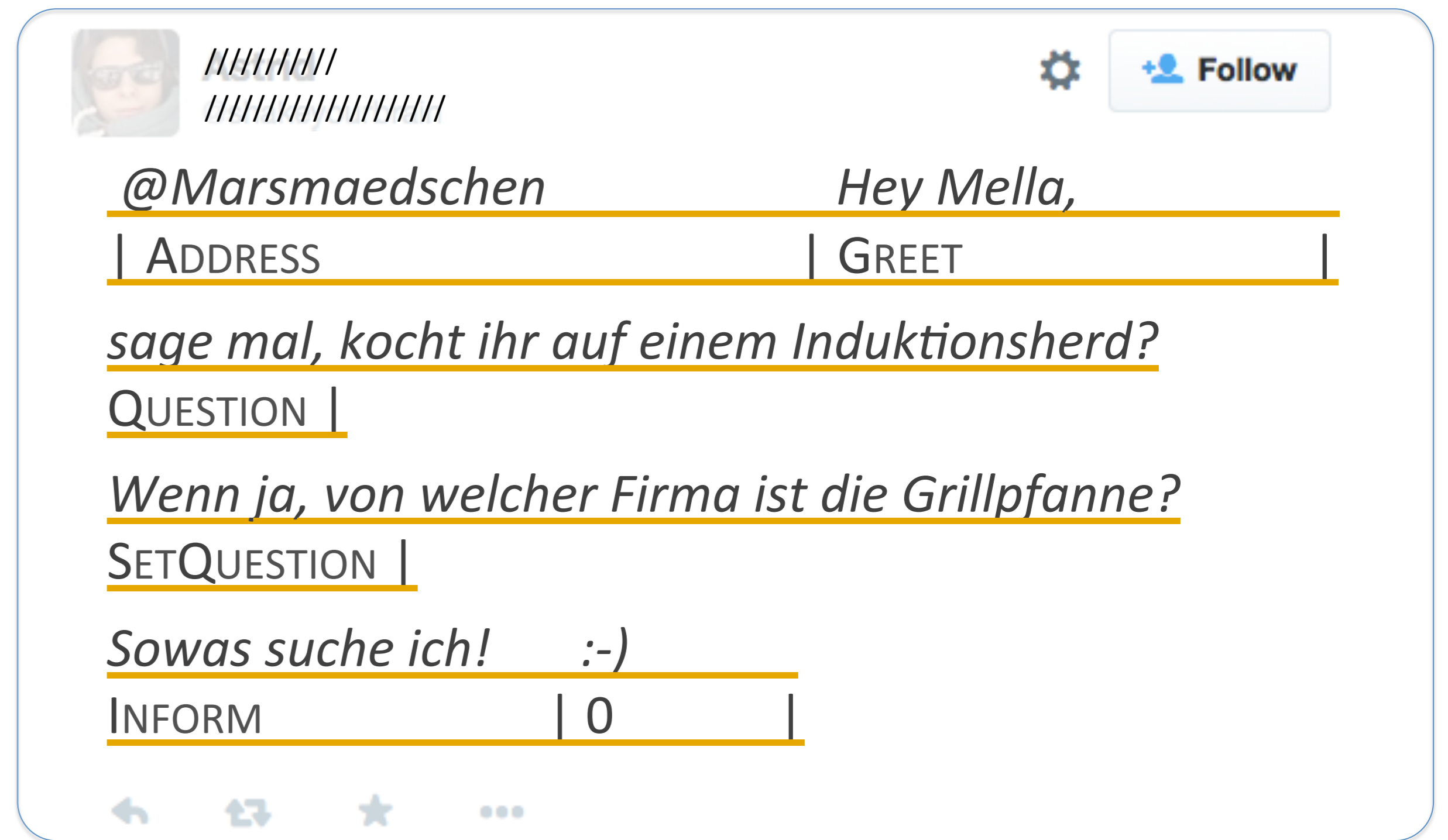
Twitter Conversations

- Up to 40% of German tweets are part of conversations.
- Answer-relation on Twitter creates discourse trees.



Dialog Acts

- characterize general discourse function
- motivation:
 - understand discourse structure
 - enable further applications
- domain-independent, adapted from DIT++ schema



Annotation & Gold Standard

annotation:

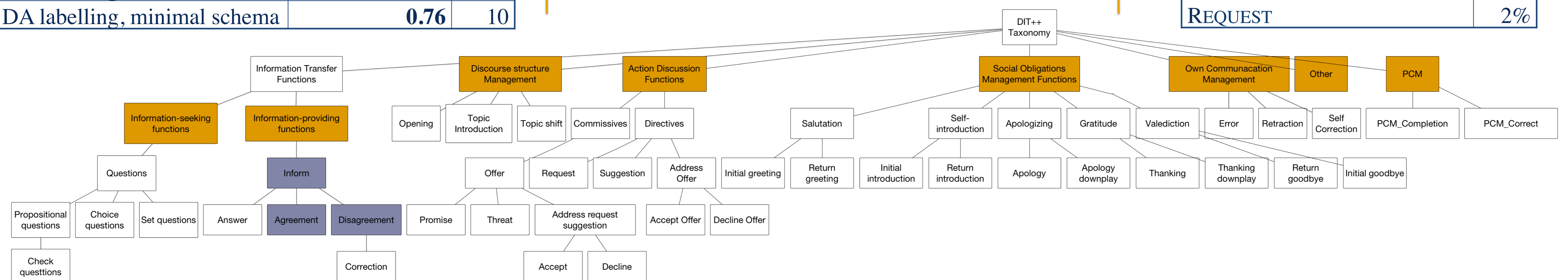
- segmentation and labelling
- 3 naïve annotators
- WebAnno: <https://webanno.github.io/>
- 172 conversations with 1566 tweets

semi-automatic creation of the gold standard:

- preprocessing: remove non-German tweets
- agreement
- majority vote
- back-off to more general DAs, using parent and sibling relationships
- manual adjudication if needed

Common Dialog Acts	%
INFORM	26%
AGREEMENT	11%
ANSWER	10%
SETQUESTION	7%
IT_IP	6%
DISAGREEMENT	5%
SOCIAL	4%
PROPQUESTION	4%
QUESTION	2%
TOPICSHIFT	2%
SUGGESTION	2%
REQUEST	2%

	Fleiss' multi- π	DAs
Segmentation	0.88	
DA labelling, full schema	0.56	57
DA labelling, reduced schema	0.65	14
DA labelling, minimal schema	0.76	10



Results & Conclusions

Many tweets contain more than one segment:

Number of segments per tweet	tweets
1 segment	89
2 segments	671
3 segments	320
4 segments	114
5 segments	17
6 segments	2
Total	1213

- It is not sufficient to assign acts to entire tweets
 - combinations of dialog moves (1)
 - sarcastic markers (4)

(1) True, unfortunately. | But what about the realization of high solar activity in the 70s and 80s?

The many faces of INFORM:

- dialogs often cannot be considered straightforward information exchanges
- many assertions carry strong secondary functions
 - metacommunication (3)
 - irony/sarcasm (4-5)
 - emotional evaluation (4)
- leads to annotation problems

(2) Coal 300 kWh, nuclear power 100 kWh

(3) The link still doesn't work.

(4) Wind/PV is enough? Lol.

(5) And the solar panels and wind power plants in the Middle Ages were great

Dialog transitions:

Long dialogs	Short dialogs
INFORM, INFORM	INFORM, INFORM
INFORM, AGREEMENT	<S>, OPEN
AGREEMENT, INFORM	SETQUEST., ANSWER
ANSWER, INFORM	ANSWER, INFORM
SETQUEST., ANSWER	<S>, TOPICINTRO

