

# **Towards Partial Word Sense Disambiguation Tools for Czech**

**Tomáš Čapek, Pavel Šmerk**

**RASLAN 2010**

# WSD

cornerstone of NLP

black box vs toolbox solutions

Making Sense about Sense

desamb, mrec

# Desamb

hybrid tagger (rule-based and statistical)

ajka + guesser + chunk parser + HMM tagger

DIS performance issue, more benchmark ahead

# mrec

multi-word expression recogniser

domain-chunked MWE DB

$C_1, C_2, \dots, C_n \# l_1, \dots, l_k \backslash l_{k+1}, \dots, l_m$

# Semantic Annotation

SAFT as a use case

Merging MWE DB and CZWN

Optimizations for both *Desamb* and *mrec*

Thank you for your attention

Questions?