Multiword Expressions in the MEDIA\CAL\ DOMA\N: 
Who Carries the Domain Specific Meaning

Kristina Kocijan
Krešimir Šojat
Silvia Kurolt
The Context

How did we arrive at today…
The Context of the project

Some time ago, in the galaxy not so far away…

Benefits

- digital format of records
- faster available data
- easier shared data

25 000 000 tokens
88 000 unknowns
Classification: sub-domains

- PLACE
- MED
- COND
- Nouns
- DISCIPL
- KEM
- PROF
- ANAT
- DRUG
- TOOLS
- MICROORG
- PATIENT
- PLACE
- MED
- COND
- Nouns
Words, words everywhere!

- **temperaturna lista** (temperature list)
- **zdravstveni radnik** (medical worker)
- **limfni sustav** (lymphatic system)
- **srčani udar** (heart attack)
- **bolesnički krevet** (sick bad)

Medical domain
The Quest

Asking the questions
And searching for answers
Which adjectives are A-MED?

Any patterns?

Recognized from N-MED?

Any special prefixes?

A-MED and N-MED = MWE+MED?
Which adjectives are A+MED?

Any patterns?

Recognized from N+MED?

Any special prefixes?

A-MED and N-MED = MWE+MED?
Which adjectives are A+MED?

Any patterns?

Recognized from N+MED?

Any special prefixes?

A-MED and N-MED = MWE+MED?
Which adjectives are A+MED?

Any patterns?

Recognized from N+MED?

Any special prefixes?

A-MED and N-MED = MWE+MED?
The Discovery

Augmenting the dictionary
% of MED domain words at the beginning

**NOUNS**
- MED domain: 25%
- GEN domain: 75%
- Total: 27,813

**ADJECTIVES**
- MED domain: 0%
- GEN domain: 100%
- Total: 14,635
% of MED domain words at the end

NOUNS

27,813

25%

ADJECTIVES

14,635

95%

75%

5%
The Challenges
A-MED N-MED = MWE+MED

žarišna točka
{hot spot}

vodene kozice
{chicken pox}

N+MED->A-MED

prevratnički
[pre+neck+sufix]
{revolutionary}

prijelomnan
[fracture+sufix]
{turning point}

zatvorjski

? -> A+MED

kardiovaskularni
{cardiovascular}

lumbalni
{lumbal}

oralni
{oral}

Manual post-processing

{of the prison}
To be continued…

Multiword Expressions in the **MEDICAL DOMAIN**: Who Carries the Domain Specific Meaning

Kristina Kocijan, Krešimir Šojat, Silvia Kurolt