

Multiword Expressions in the

MEDICAL DOMAIN:

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...

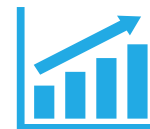


Electronic
Health Records



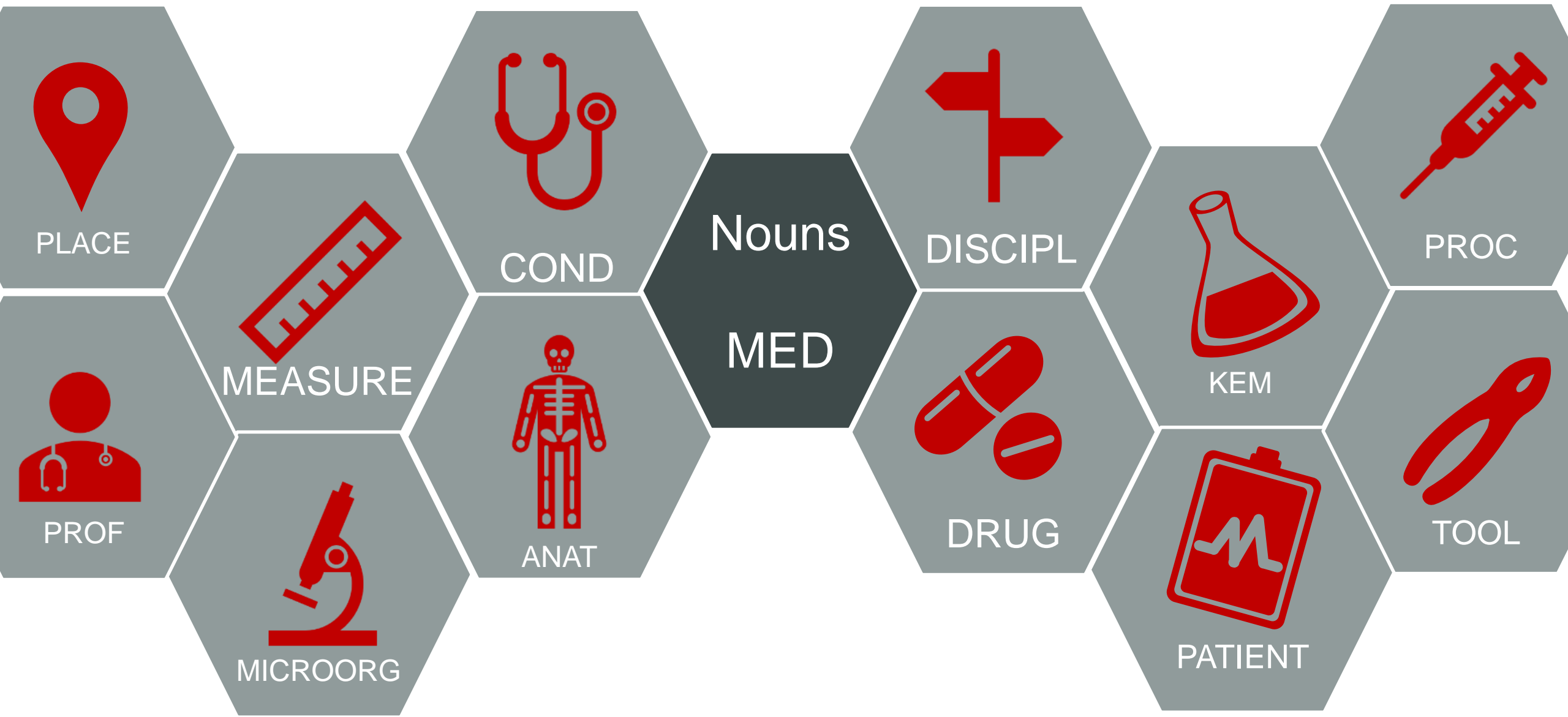
Benefits

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

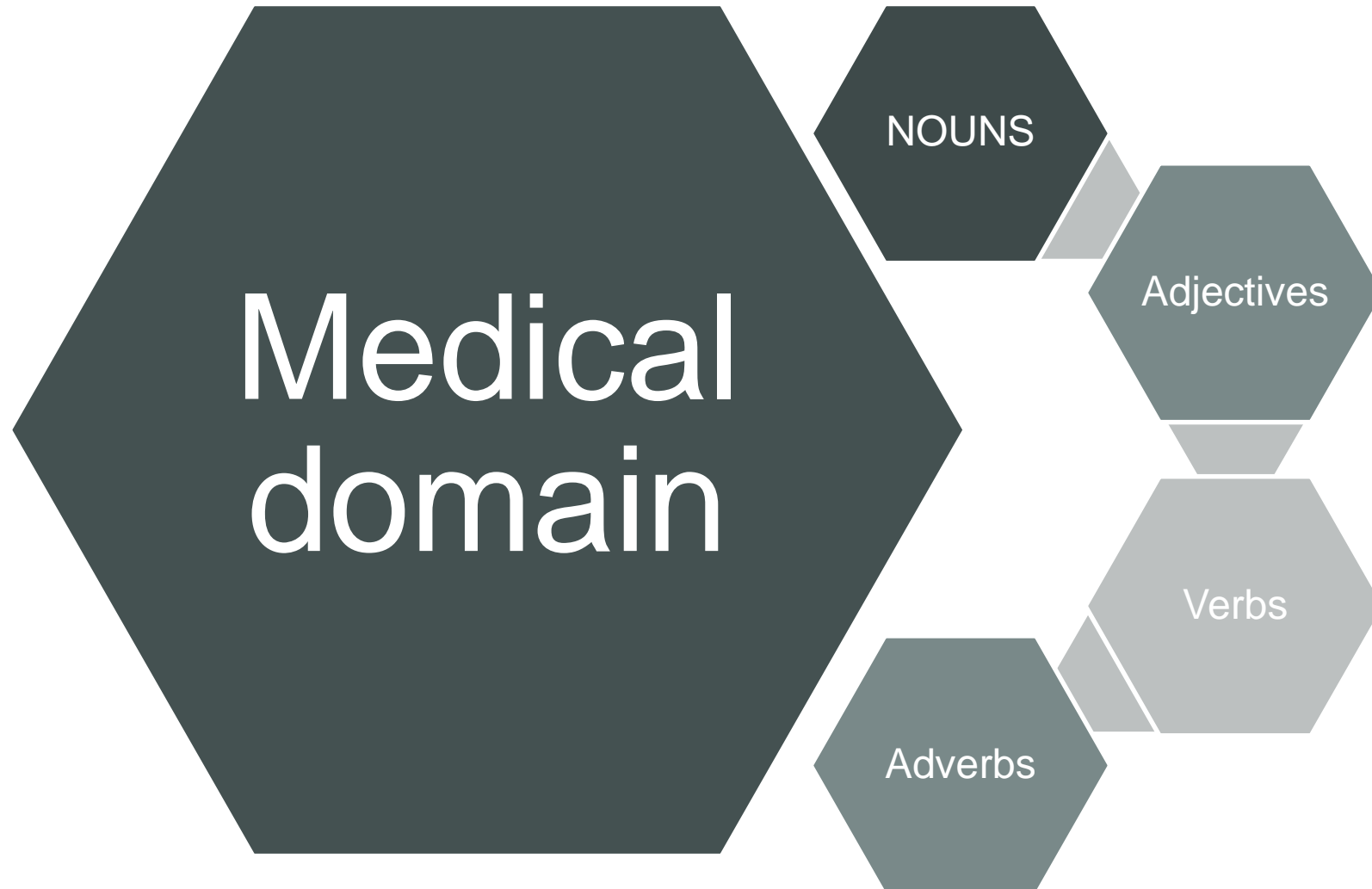


25 000 000 tokens
88 000 unknowns

Classification: **sub**-domains



Words, words everywhere!



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



The Quest

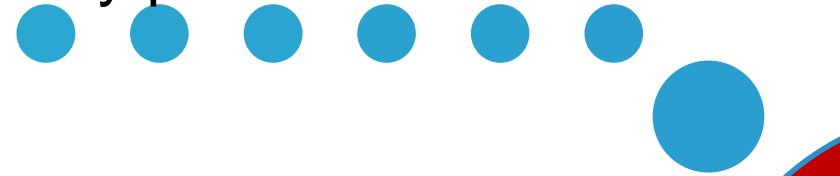
Asking the questions

And searching for answers

Which adjectives are A+MED?



Any patterns?



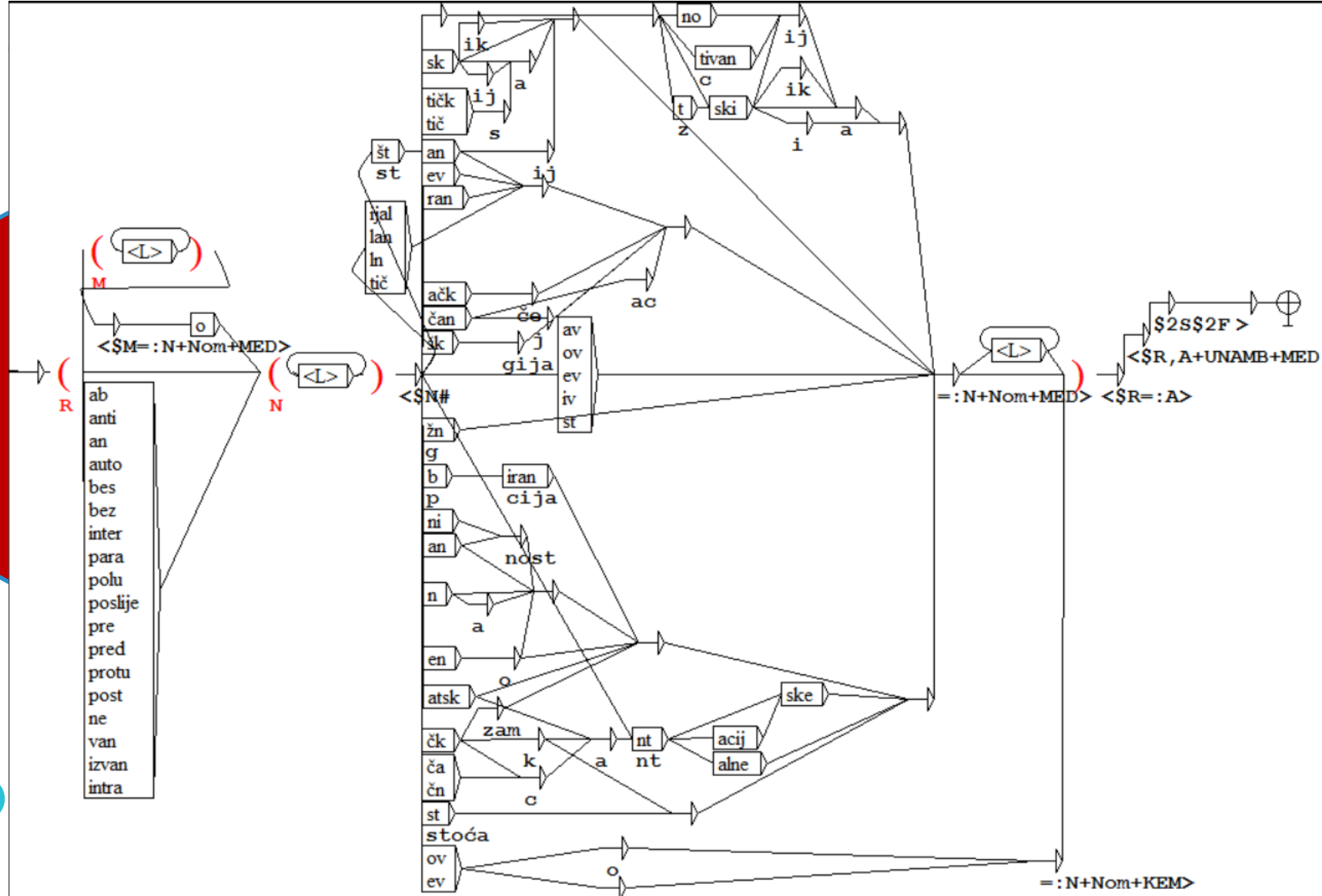
Recognized from N+MED?



Any special prefixes?



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



Which adjectives are A+MED?



infek
infek

tivan, A
cija, N

Any patterns?



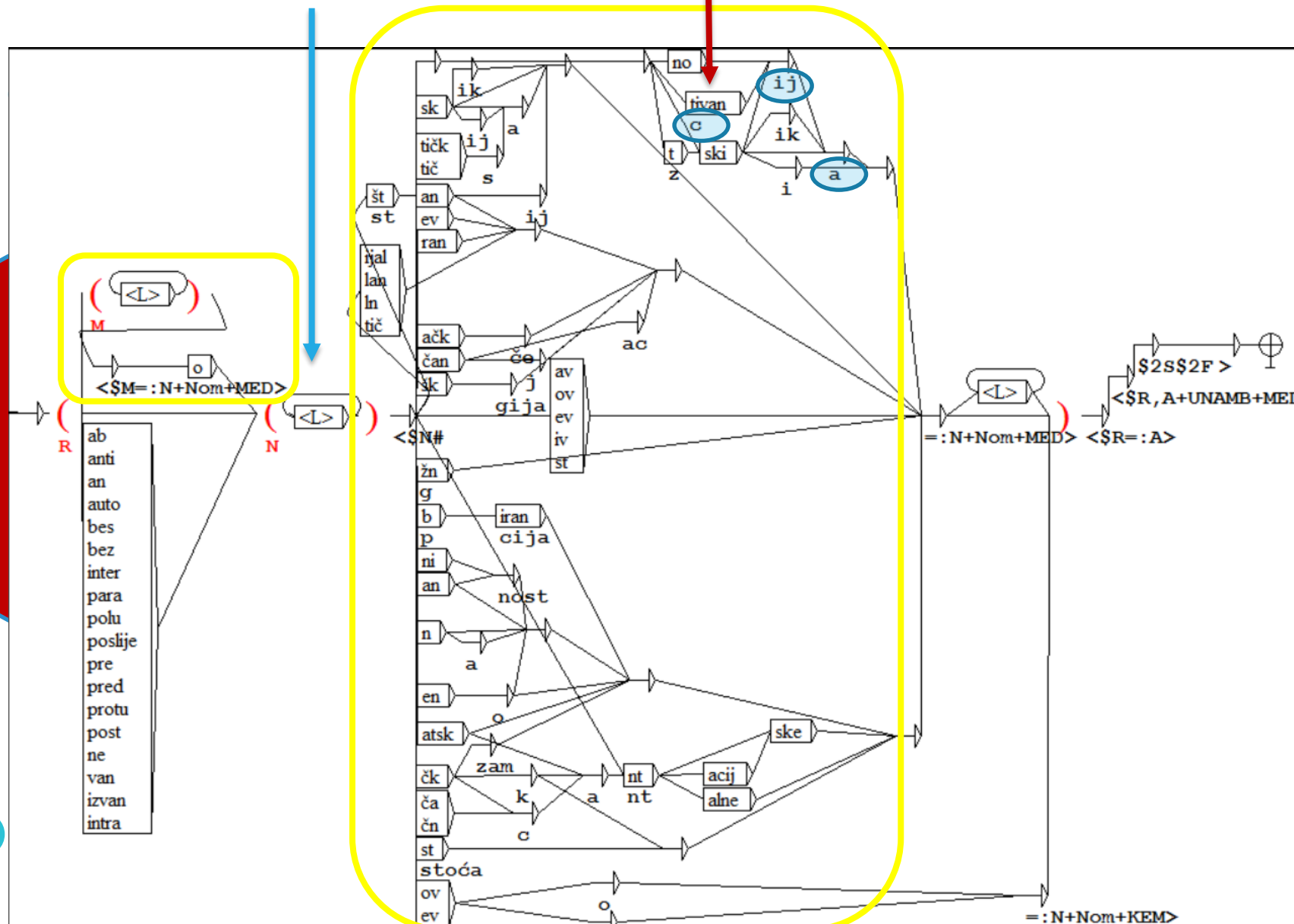
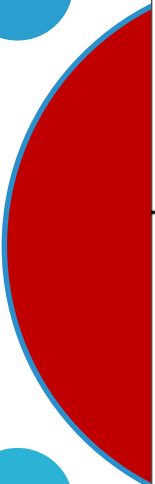
Recognized from N+MED?



Any special prefixes?



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



Which adjectives are A+MED?

infek
infek

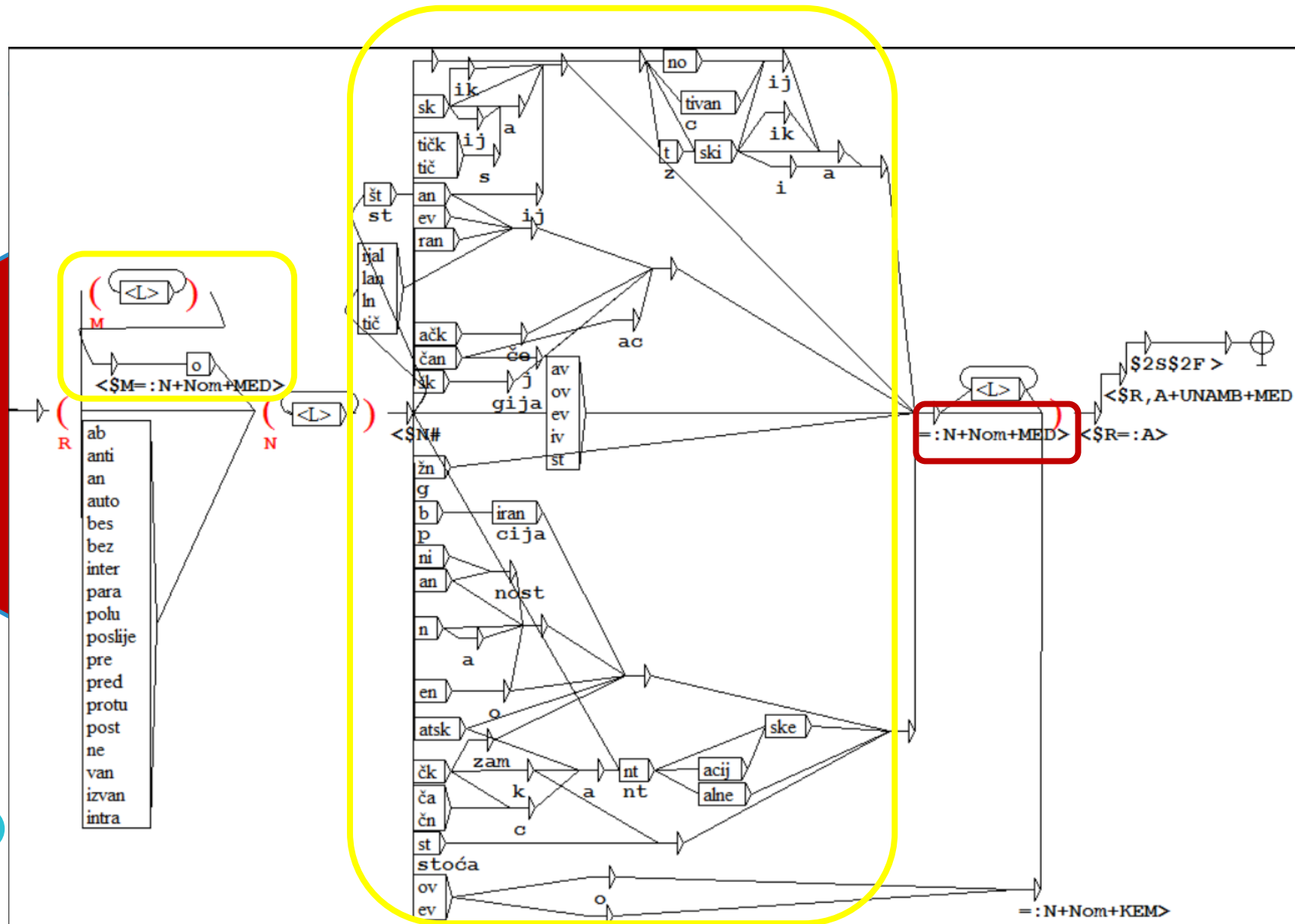
tivan, A
cija, N

Any patterns?

Recognized from N+MED?

Any special prefixes?

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

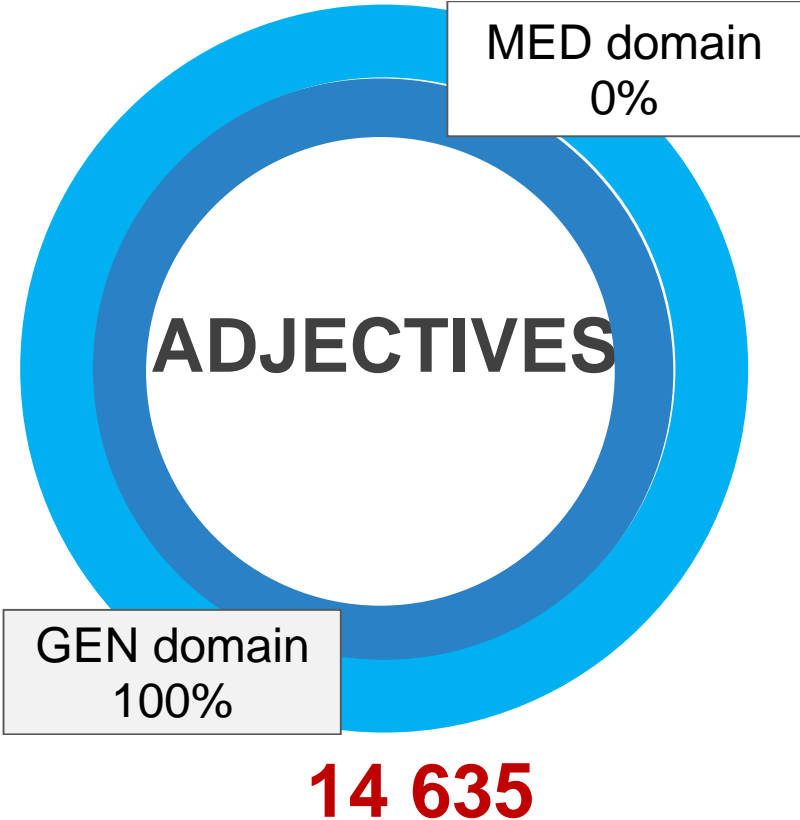
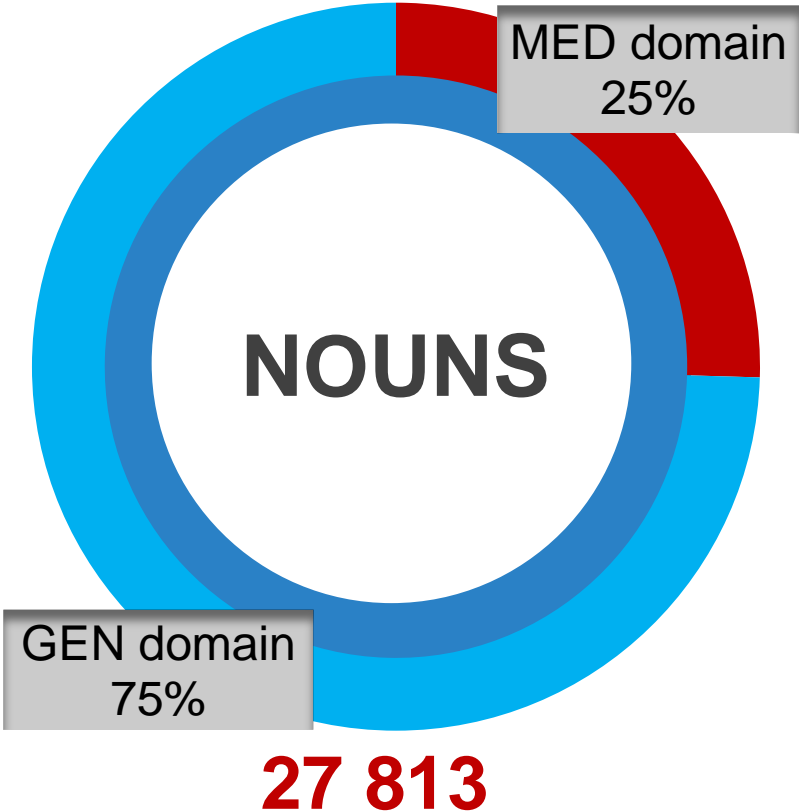




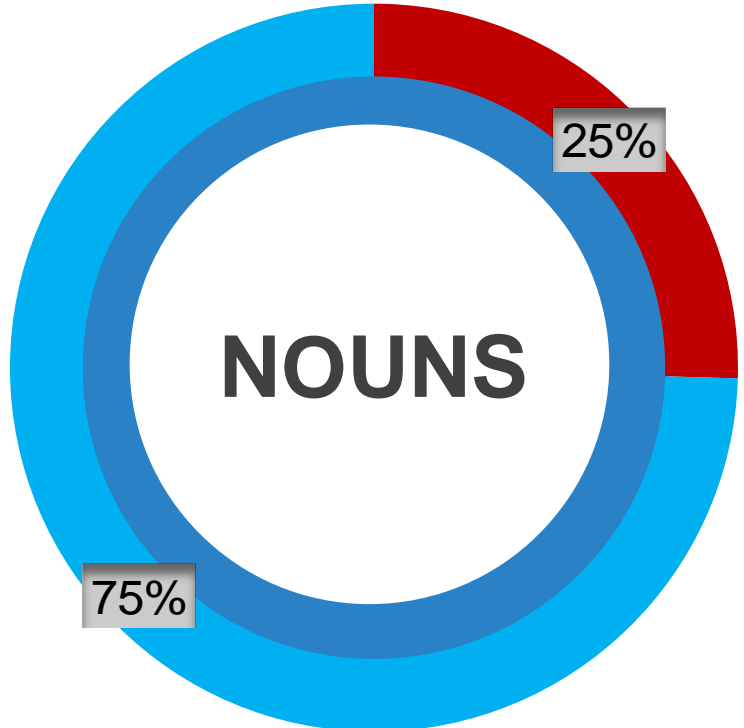
The Discovery

Augmenting the dictionary

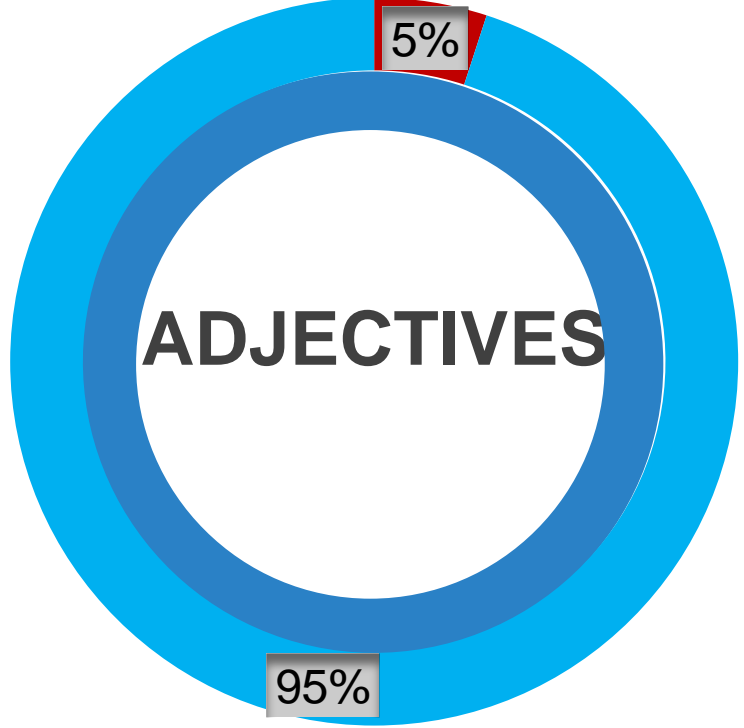
% of MED domain words at the beginning



% of MED domain words **at the end**



27 813



14 635



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}

prijeloman

[**fracture**+sufix]

{turning point}

zatvorski

{of the prison}

? -> A+MED

kardio**vaskular**ni

{cardiovaskular}

lumbalni

{lumbal}

oralni

{oral}

Manual post-processing



To be continued...

Multiword Expressions in the

MEDICAL DOMAIN:

Who Carries the Domain Specific Meaning

Kristina Kocijan, Krešimir Šojat, Silvia Kurolt