

# **PATIENT CASE SIMILARITY**

**TEAM : THE\_OUTLIERS**

NATIONAL INSTITUTE OF TECHNOLOGY, SILCHAR

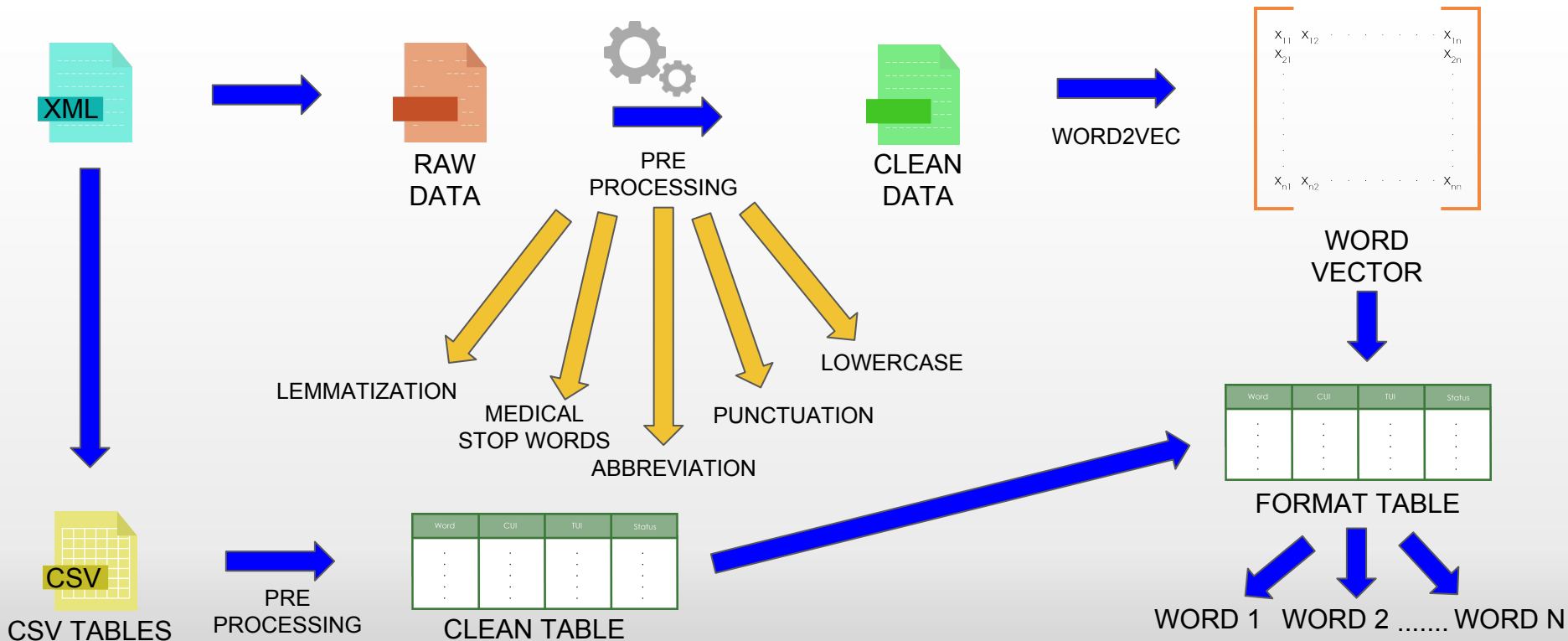
**ezDI – RG2 – 2**



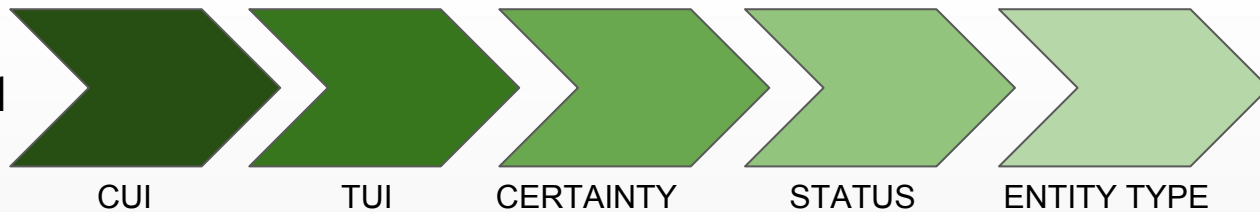
# CHALLENGES

- THE VAST SYNONYMY IN MEDICAL TEXT MAKES IT UNSUITABLE TO BE TREATED AS TRIVIAL NLP USE CASE.
- IT IS DIFFICULT TO INCORPORATE DOMAIN KNOWLEDGE.
- MEDICAL DATA IS NOT PUBLICLY AVAILABLE FOR ANALYTICAL PURPOSES.
- THE DATA IS HIGHLY VARIANT DEPENDING UPON THE DEMOGRAPHIC PARTICULARS OF THE PATIENT.

# MR2VEC



WORD 1



$x_{11}$	$x_{12}$	...	$x_{1n}$
$x_{21}$			$x_{2n}$
...			...
$x_{n1}$	$x_{n2}$	...	$x_{nn}$

FINAL WORD  
VECTORS

MAPPING



RAW  
DATA

SIMILARITY SCORE

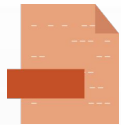
$x_{11}$	$x_{12}$	...	$x_{1n}$
$x_{21}$			$x_{2n}$
...			...
$x_{n1}$	$x_{n2}$	...	$x_{nn}$

DOCUMENT  
VECTORS

WILL FIND  
SUM, AVG.

$x_1$	$x_1$	$x_1$
$x_2$	$x_2$	$x_2$
...	...	...
$x_n$	$x_n$	$x_n$

VECTORS



RAW  
DATA



WMD



SIMILARITY SCORE



SIMILARITY SCORE



MR2VEC  
SIMILARITY SCORE

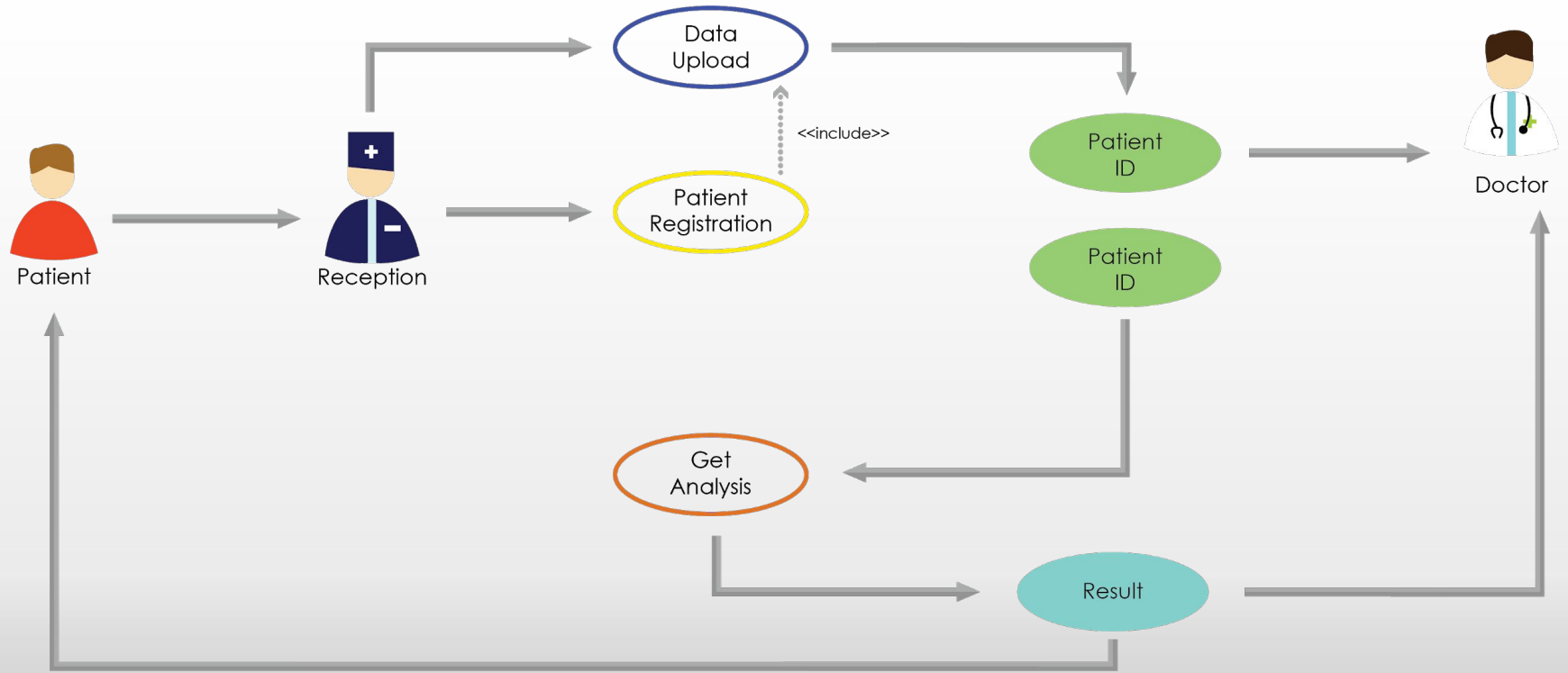
# FEATURES USED

- CUI
- TUI
- CERTAINTY
- STATUS
- SECTION NAME
- DOCUMENT TYPE
- RANK

# NOVELTY

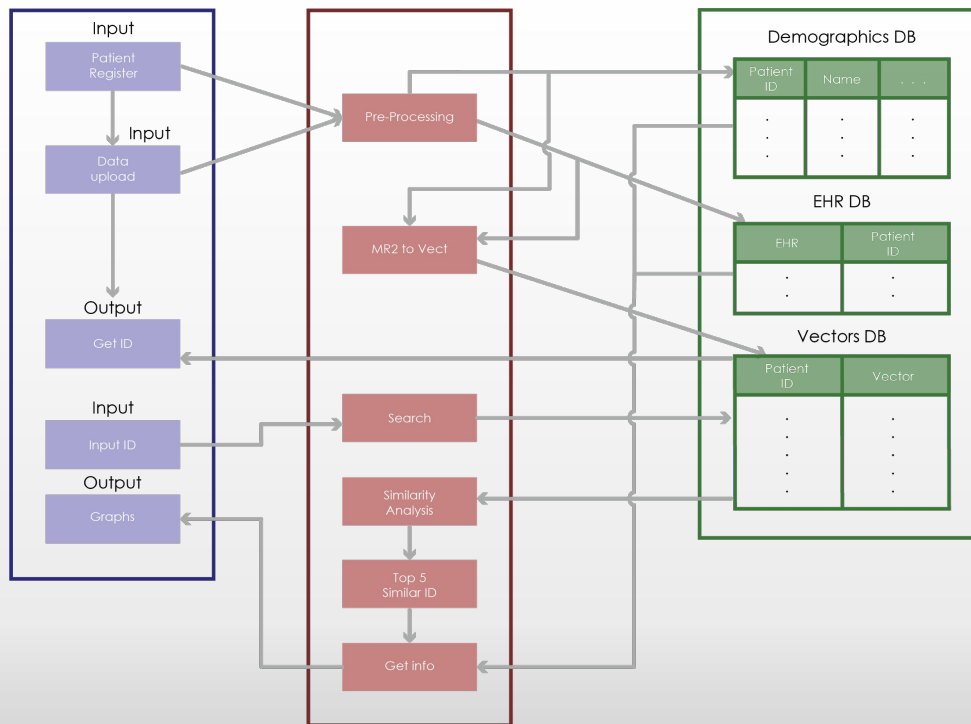
- WORD MOVERS DISTANCE
- CONCATENATION OF WORD VECTORS AND MANUAL FEATURE VECTORS
- CREATING PATIENT SPECIFIC WORD VECTORS

# FUTURE PROSPECTS





# FUTURE PROSPECTS(CONTD.)



# MEET THE TEAM

## THE\_OUTLIERS

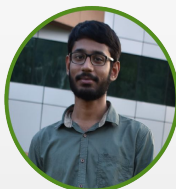
NATIONAL INSTITUTE OF TECHNOLOGY, SILCHAR



**APOORVA  
VIKRAM SINGH**



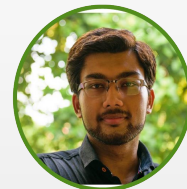
**DIVYANSHA**



**KHILJI ABDULLAH  
FAIZ UR RAHMAN**



**SACHIN RAI**



**ANUBHAV  
SACHAN**



**TUSHAR  
AGARWAL**

**THANK YOU**