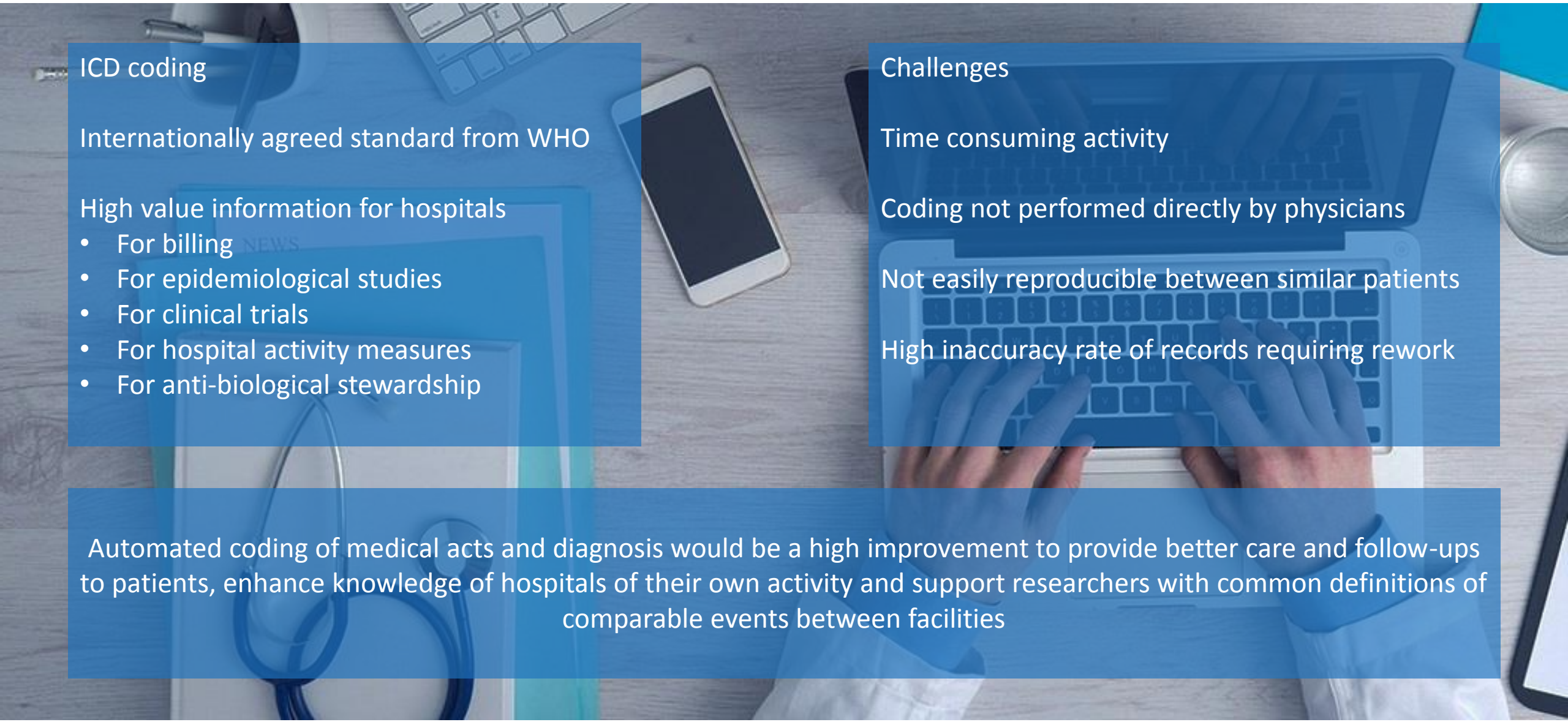
## Team 4


Remi Flicoteaux – Lead
Pierre Zweigenbaum
Vassilina Nikoulina
Yohan Taille
Sameh Megrhi
Claire Vinet
Cyril Grouin
Namik Taright
Thomas Séjourné
Michael Elbaz


## Dat-ICU 2018

# Automated text classification for ICD 9 diagnosis code assignment from MIMIC Database

**Hopital Rothschild - Campus Picpus**



## ICD coding

Internationally agreed standard from WHO

High value information for hospitals

- For billing
- For epidemiological studies
- For clinical trials
- For hospital activity measures
- For anti-biological stewardship

## Challenges

Time consuming activity

Coding not performed directly by physicians

Not easily reproducible between similar patients

High inaccuracy rate of records requiring rework

Automated coding of medical acts and diagnosis would be a high improvement to provide better care and follow-ups to patients, enhance knowledge of hospitals of their own activity and support researchers with common definitions of comparable events between facilities

## Aim of study

MIMIC contains both textual notes from discharge records completed by treating physicians and human-assigned ICD codes

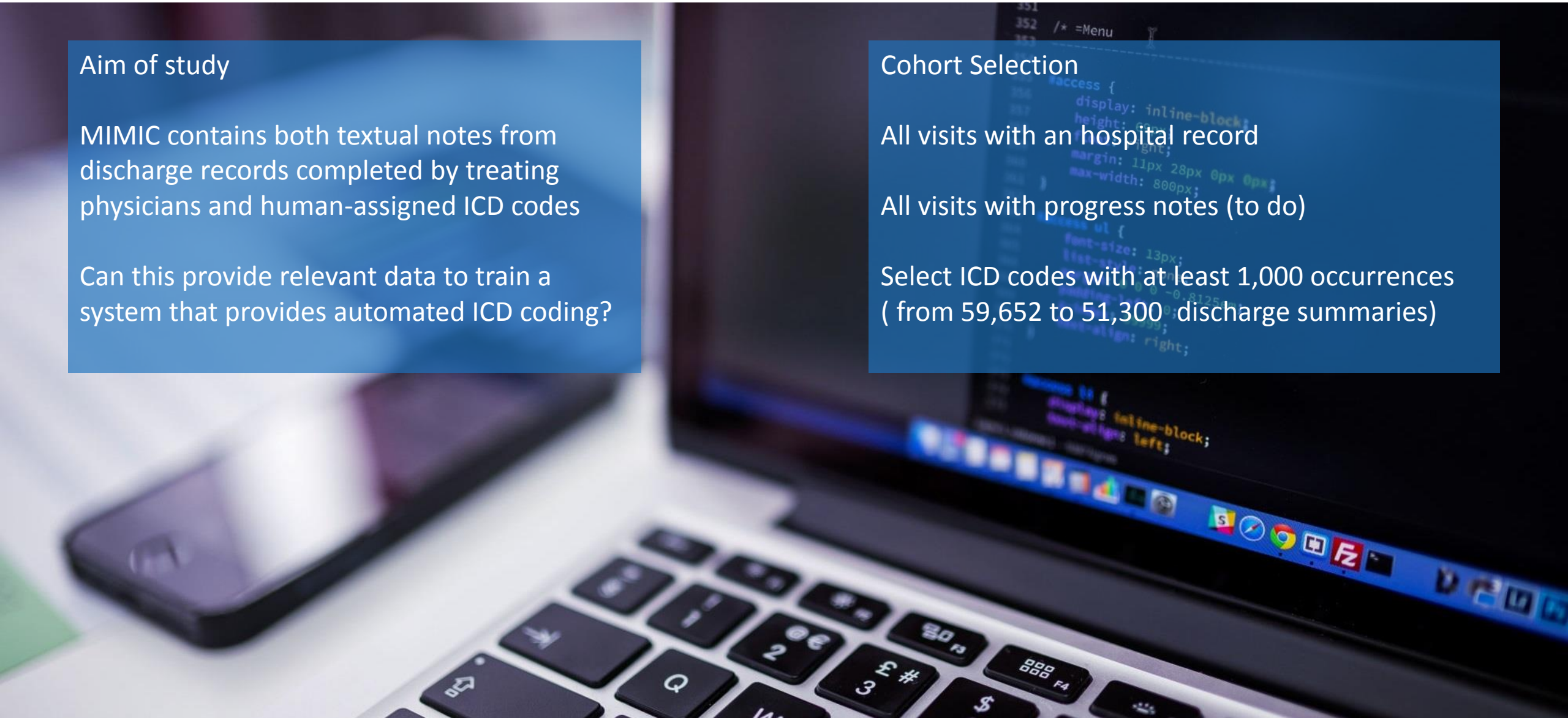
Can this provide relevant data to train a system that provides automated ICD coding?

## Cohort Selection

All visits with an hospital record

All visits with progress notes (to do)

Select ICD codes with at least 1,000 occurrences ( from 59,652 to 51,300 discharge summaries)



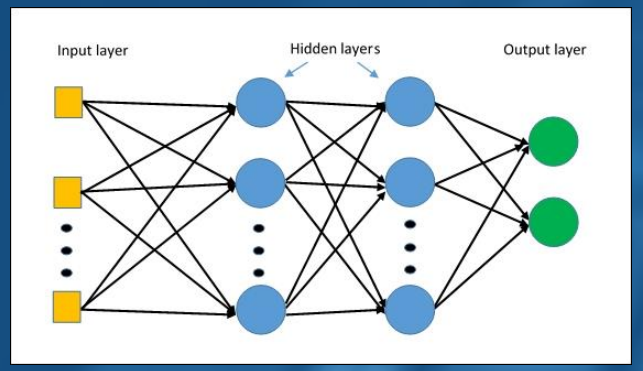
# Method

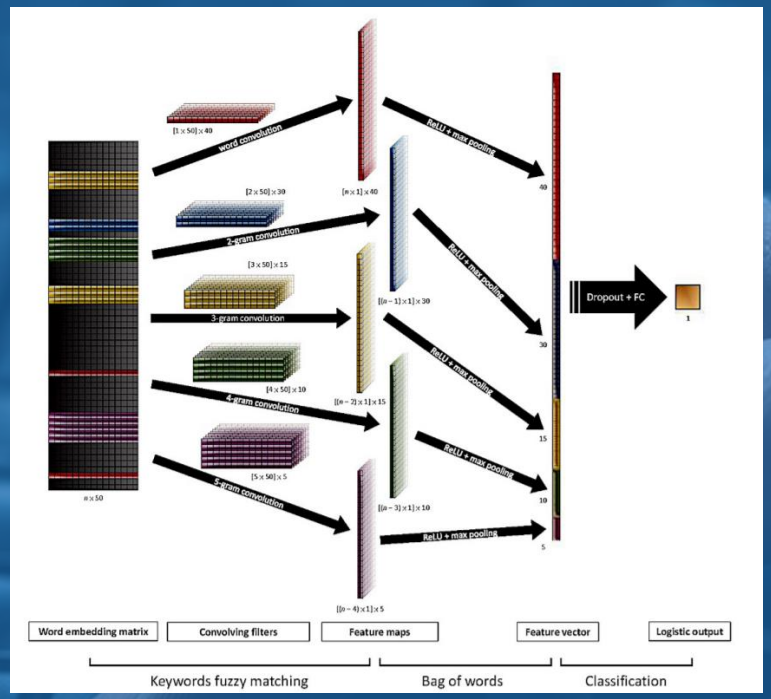
Data selection:
- SparkSQL
- Pandas

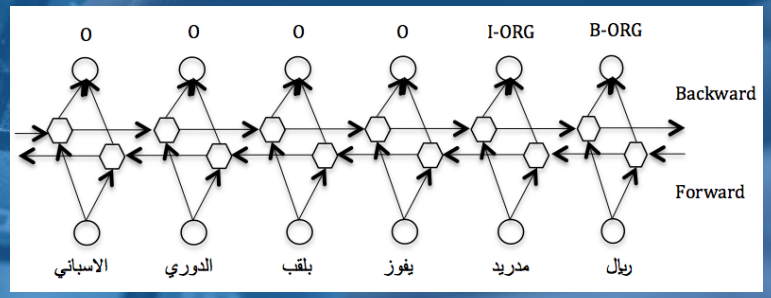
Feature extraction:
- bag of tokenized words: known ICD terms
- GloVe word embeddings trained on MIMIC texts

Classification:
- Convolutional Neural Network
- Bidirectional LSTM
- Multilayer perceptron

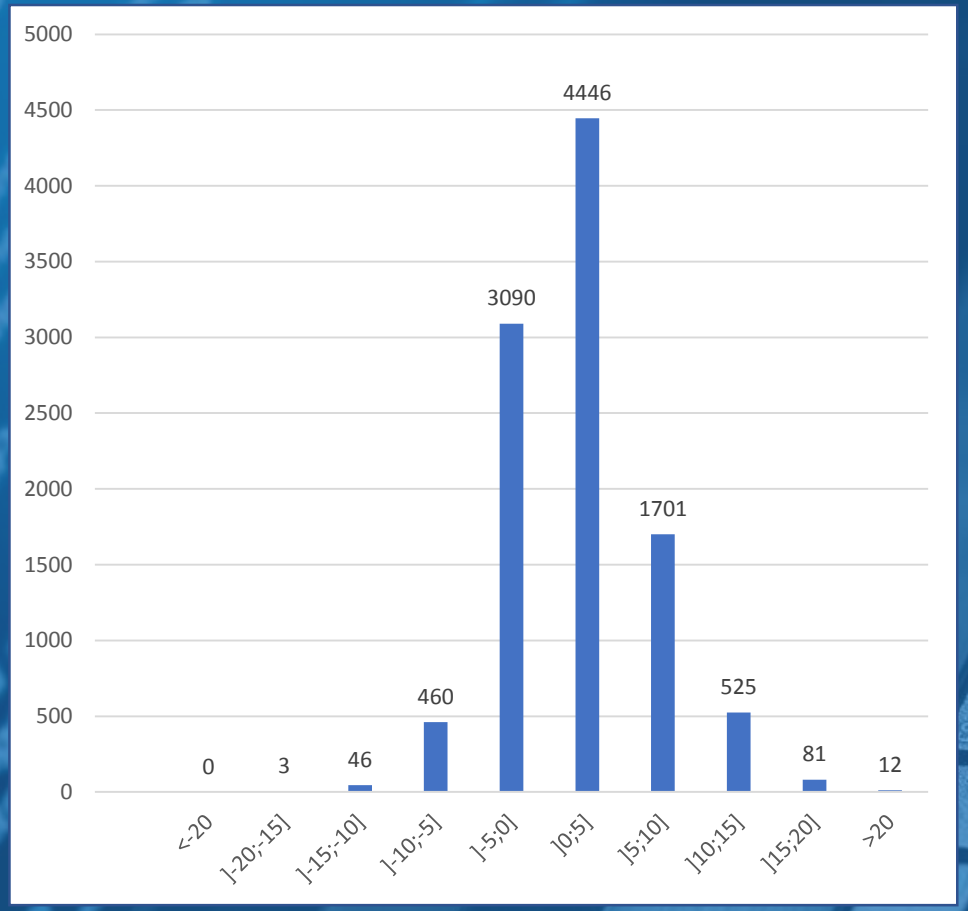
Evaluation:
- F-Score



## Bidirectional LSTM evaluation

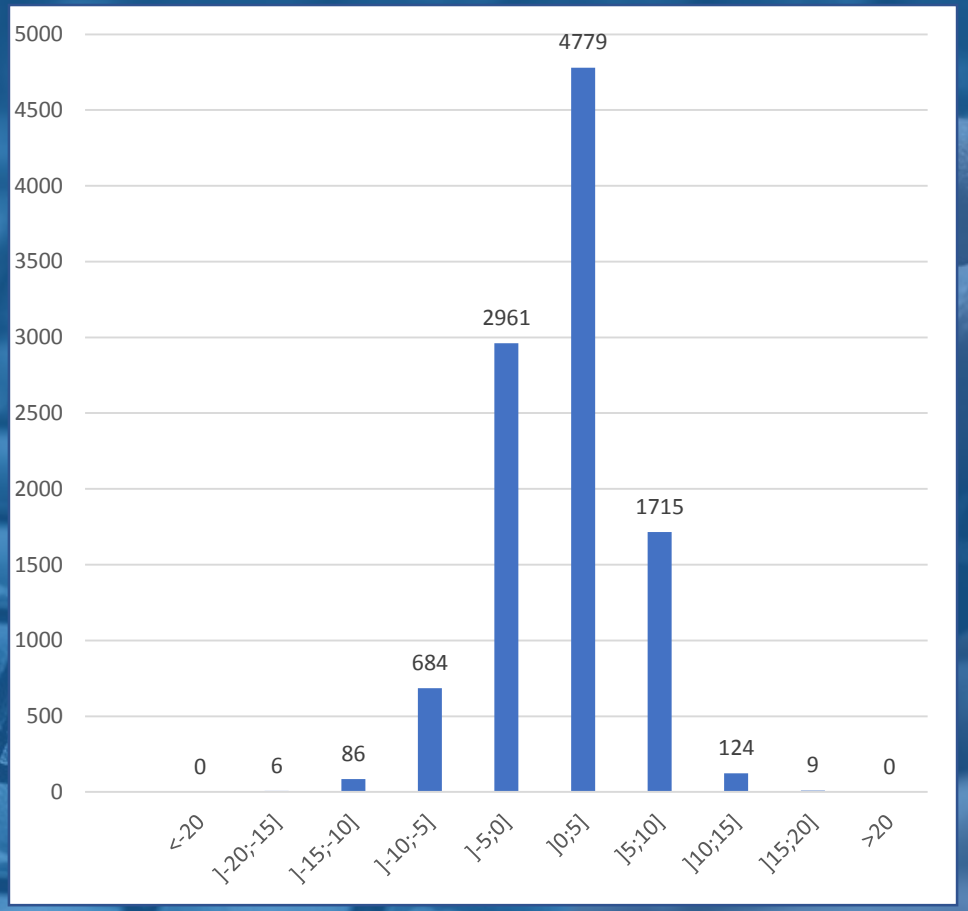



Threshold 0,25

On 10,354 expected

Max F-Score with a 0,3 thershold

## Convolutional Neural Network Evaluation



## Further developments & Discussion

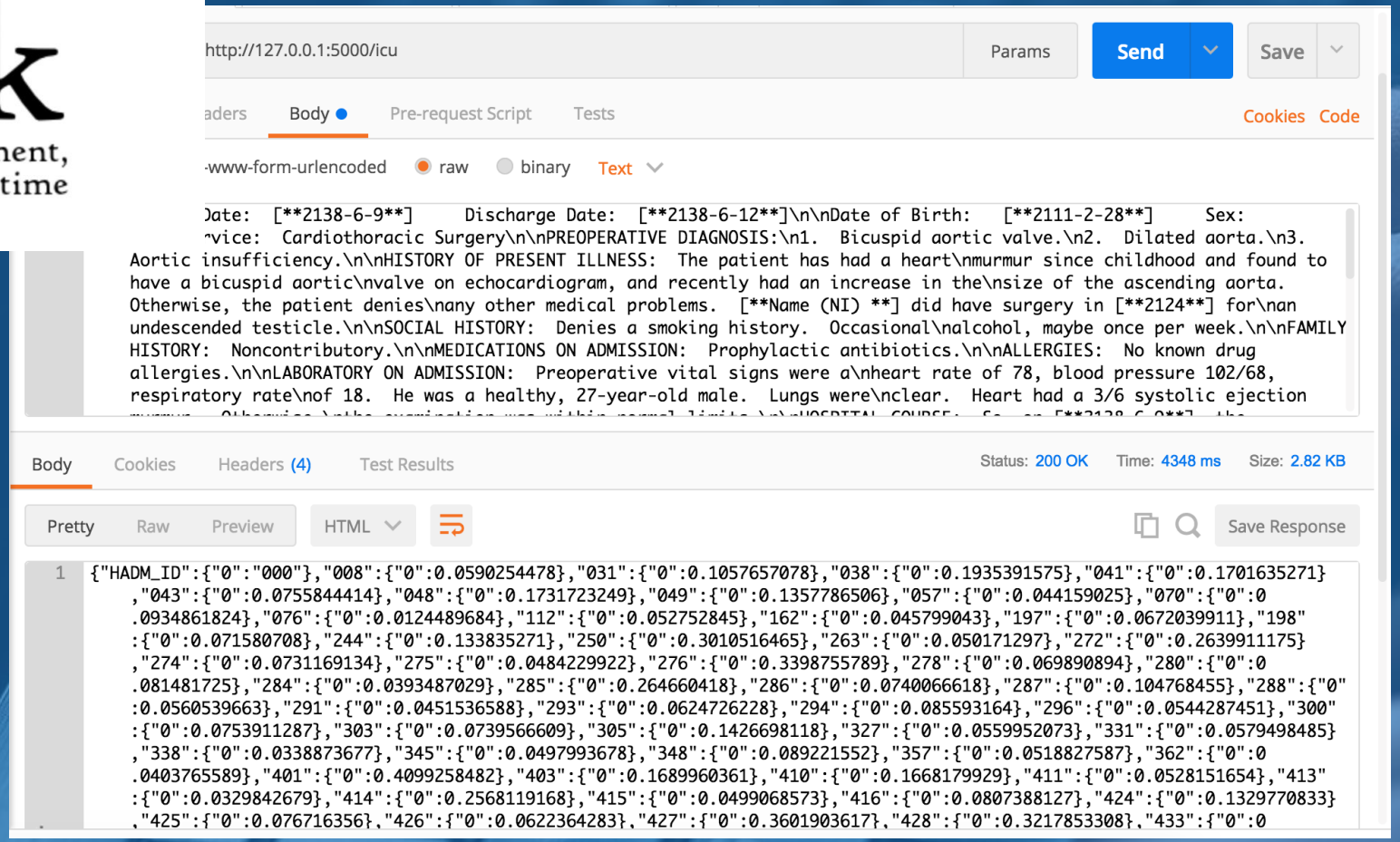
## Interaction with other teams

(active learning procedure for production…)

## Further improvements & testing

## Web service deployment

Flask
web development, one drop at a time




IQVIA™ cnrs LABS NAVER LABS EUROPE

# Thank You to our hosts