

**Евалуација семантичких ознака  
у ручно аотираним корпусима**

Миљана Младеновић

Београд, децембар 2016.

- **Појам аотације корпуса**
- **Примери аотације**
- **Циљеви аотације корпуса**
- **Начини спровођења ручне аотације**
- **Скале мерења**
- **Методе статистичке оцене ручно аотираног корпуса**
- **Приказ Крипендорфове алфа оцене на ручно аотираном корпусу ироничних ТВИТОВА**

# Појам анотације корпуса

Анотација корпуса је поступак којим се деловима корпуса придружују додатне информације. Развој и ручна анотација малих делова електронских корпуса значајни су због примене у методама којима се генеришу модели аутоматске анотације великих електронских корпуса као и због примене у задацима обраде природног језика који користе поступке анотирања као што су: кореновање (енг. stemming), лематизација (енг. lemmatisation), означавање врстама речи (енг. PoS tagging), семантичка анотација и др. Како квалитет ручне анотације утиче на успешност ових модела и метода, то је врло важно развити и применити поступке за оцену квалитета ручне анотације.

# Примери анотације

- **XML етикете**

<time.duration.abs>hiljadu i 200 godina</time.duration.abs>

- **морфолошке етикете у LADL формату:**

генерацију, генерација.N600:fs4 – именица (N) у акузативу (4) сингулара (s) женског рода (f).

- **етикете као саставни део текста**

On je dodao da je što se tiče obnove voznog parka gradskog saobraćajnog preduzeća grad Beograd S\_LOCATION potpisao ugovor sa sa Ikarbusom S\_ORGANIZATION za novih 30 autobusa i da se očekuje još jedno potpisivanje ugovora za još 20 autobusa posle rebalansa budžeta.

од {{8. марта,.NE+time+date+rel}} до {{7. априла,.NE+time+date+rel}},  
.NE+time+date+period}}

## Циљеви аотације корпуса

- Аутоматска анализа корпуса
- Олакшано истраживање и екстракција информација
- Могућност вишеструког коришћења корпуса
- Генерисање скупова за учење који се користе у задацима машинског учења

## Начини спровођења ручне аотације

- Анкете
- Електронски формулари (нпр. Google forms)
- Системи за групну расподелу рада (енг. Crowdsourcing) као што су *Amazon Mechanical Turk*, *CrowdFlower*, *GATE Crowd* и др.

# Скале мерења

Номинална скала (квантитативна или квалитативна)

Пол особе:

1. женски
2. мушки

Уређена скала

Како сте задовољни пруженом услугом:

1. Веома задовољни
2. Делимично задовољни
3. Неутрални
4. Делимично незадовољни
5. Веома незадовољни

Интервална скала

Одаберите термин пријављивања у опсегу:

1. Од 17 до 18 сати
  2. Од 18 до 19 сати
  3. Од 19 до 20 сати
- 

Размерна скала

Структура високообразованих у фирми је:

1. 30 на 100 запошљених
2. 50 на 100 запошљених
3. 70 на 100 запошљених

# Методе статистичке оцене ручно аотираног корпуса

- Сагласност аотације два или више аотатора је оцена у којој мери два или више аотатора дају исту етикету за одређену категорију.
- Најчешће коришћене методе за оцену квалитета ручне аотације два или више аотатора из класе статистичких оцена сагласности су Коенов капа (Cohen's kappa), Флајсоф капа (Fleiss' kappa) и Крипендорфов алфа (Krippendorff's alpha) коефицијент.
- Коенов капа коефицијент је статистика која мери сагласност у аотацији између два оцењивача номиналном скалом мерења. Посматрајмо најпре тзв. матрицу сагласности у случају два аотатора и три категорије.

Табела 1. Матрица сагласности оцена 2 анотатора и 3 категорије

		Анотатор А			
Анотатор В	Категорија	1	2	3	$p_{i_B}$
	1	0.25 (0.20)	0.13 (0.15)	0.12 (0.15)	0.50
	2	0.12 (0.12)	0.02 (0.09)	0.16 (0.09)	0.30
	3	0.03 (0.08)	0.15 (0.06)	0.02 (0.06)	0.20
	$p_{i_A}$	0.40	0.30	0.30	$\sum p_i = 1.00$

Анотатор А је 40% својих оцена доделио класи А, а анотатор В је истој класи доделио 50% својих оцена. То значи да би очекивано слагање оцена анотатора у овој класи било  $0.50 \cdot 0.40 = 0.20$

На исти начин оцењујемо очекивано слагање анотатора и по осталим класама и те вредности се налазе у заградама у матрици сагласности.

Уведимо ознаке  $P_o$  - однос броја анотација у којима постоји слагање и  $P_e$  - очекивани однос броја анотација у којима постоји слагање

$$P_o = 0.25 + 0.02 + 0.02 = 0.29$$

$$P_e = 0.20 + 0.09 + 0.06 = 0.35$$

Уочимо да се ради о независним случајним догађајима које посматрамо у паровима (догађај да је анотатор  $X_1$  доделио етикету  $Y_1$  категорији Z и догађај да је анотатор X доделио етикету  $Y_2$  категорији Z)



Коенов капа степен сагласности ( $A_m$  – agreement measure) изражава се као однос

$$A_m = \frac{P_o - P_e}{1 - P_e}$$

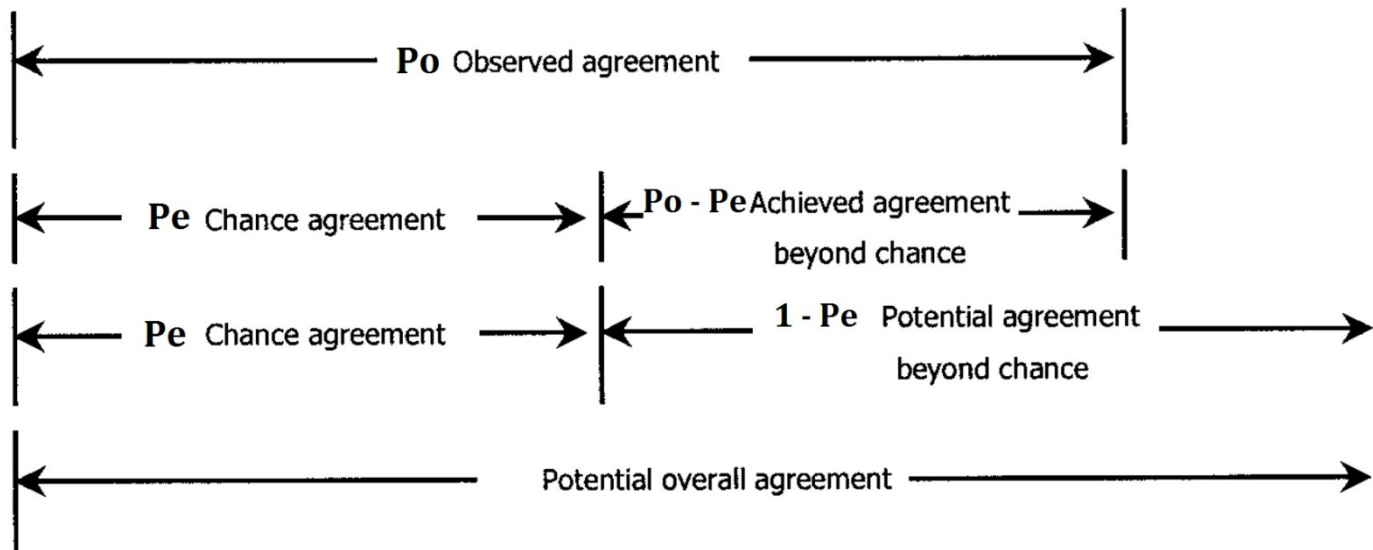
У претходном примеру

$$A_m = \frac{0.29 - 0.35}{1 - 0.35} = -0.092$$

$P_o - P_e$  представља оне случајеве који су се појавили као слагање изван очекиваног слагања

$1 - P_e$  представља вероватноћу очекиваног неслагања

$A_m$  се интерпретира као удео слагања изван очекиваног у делу у коме се очекује неслагање



Интерпретација резултата степена сагласности на основу Коеновог капа коефицијента дата је у следећој табели:

Коенов капа кефицијент	Ниво сагласности
< 0.00	слаб
0.00 – 0.20	мали
0.21 – 0.40	коректан
0.41 – 0.60	умерен
0.61 – 0.80	висок
0.81 – 1.00	веома висок

Даља унапређења статистичке оцене из класе капа коефицијената су **Флајсов капа** (Fleiss' kappa) и **Крипендорфов алфа** (Krippendorff's alpha) коефицијент.

Флајсов капа узима у разматрање већи број анотатора, а не само два, а Крипендорфов алфа омогућава оцену сагласности анотације када је:

- број анотатора већи од 2
- број категорија (ознака) већи од 2
- скала мерења може бити произвољна (номинална, уређена, интервална, размерна)
- анотатори не морају свакој јединици за анотацију доделити етикету (енг. missing data)
- скупови за анотацију немају захтевани минимум података.

На слици 1 дат је приказ примера анотације коју Крипендорфов алфа алгоритам може обрадити.

<b>Jedinica : za anotaciju</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>
<b>Anotator <i>A</i>:</b>	1	2	3	3	2	1	4	1	2	.	.	.
<b>Anotator <i>B</i>:</b>	1	2	3	3	2	2	4	1	2	5	.	3
<b>Anotator <i>C</i>:</b>	.	3	3	3	2	3	4	2	2	5	1	.
<b>Anotator <i>D</i>:</b>	1	2	3	3	2	4	4	1	2	5	1	.
<b>Kategorije :</b>	1	2	3	4	5							

Слика 1. Резултати анотације употребом номиналне скале када је број анотатора већи од 2, број категорија већи од 2, а неке од јединица за анотацију немају етикету.

<http://web.asc.upenn.edu/usr/krippendorff/>

Крипендорфов алфа алгоритам може се користити као један од онлајн алата доступних на сајту Друштва за језичке ресурсе и технологије на адреси

<http://sm.jerteh.rs/Krippendorff.aspx>

Формат обрасца који се процесира је Excel (xlsx), у коме један ред представља све оцене једног посматрача. Недостајајуће оцене се означавају тачком (.)