



JERTEH seminar, 26.11.2019.



Distant Reading WG2

“Methods and tools”

aktivnosti

Ranka Stanković, Branislava Šandrih, Cvetana Krstev



COST-DR WG2

“Methods and tools”



- Cilj: deljenje, evaluacija i poboljšanje metoda i alata za istraživanje u oblasti DR-a
- podstaknuti razmišljanje o različitim razvojnim okruženjima, alatima, paketima i bibliotekama
- objediniti postojeće tehnologije obrade prirodnog jezika u odblast digitalne humanistike (npr. NER) i spojiti ih što je više moguće u jednu platformu
- Analizirati mogućnosti i nedostatke postojećih pristupa u ovoj oblasti

<https://www.distant-reading.net/wg-2/>



Anotacije korpusa

„Annotation pipeline“ za romane ELTeC korpusa za takozvani nivo 2, uključuje 5 nivoa:

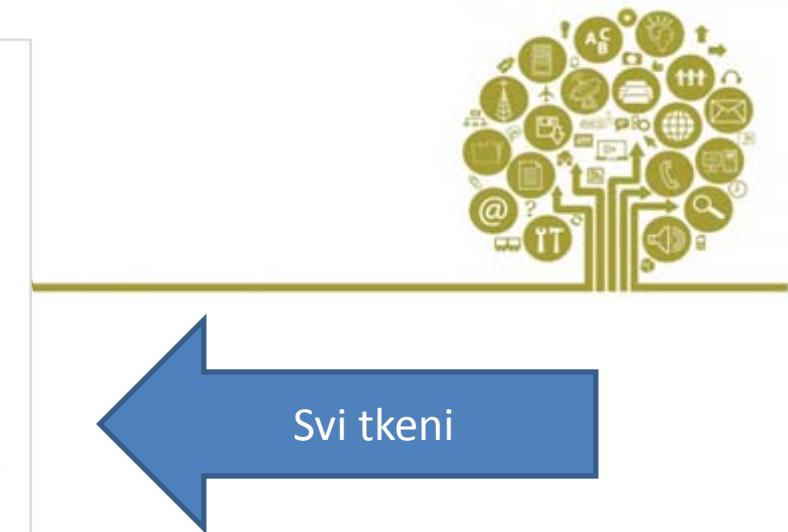
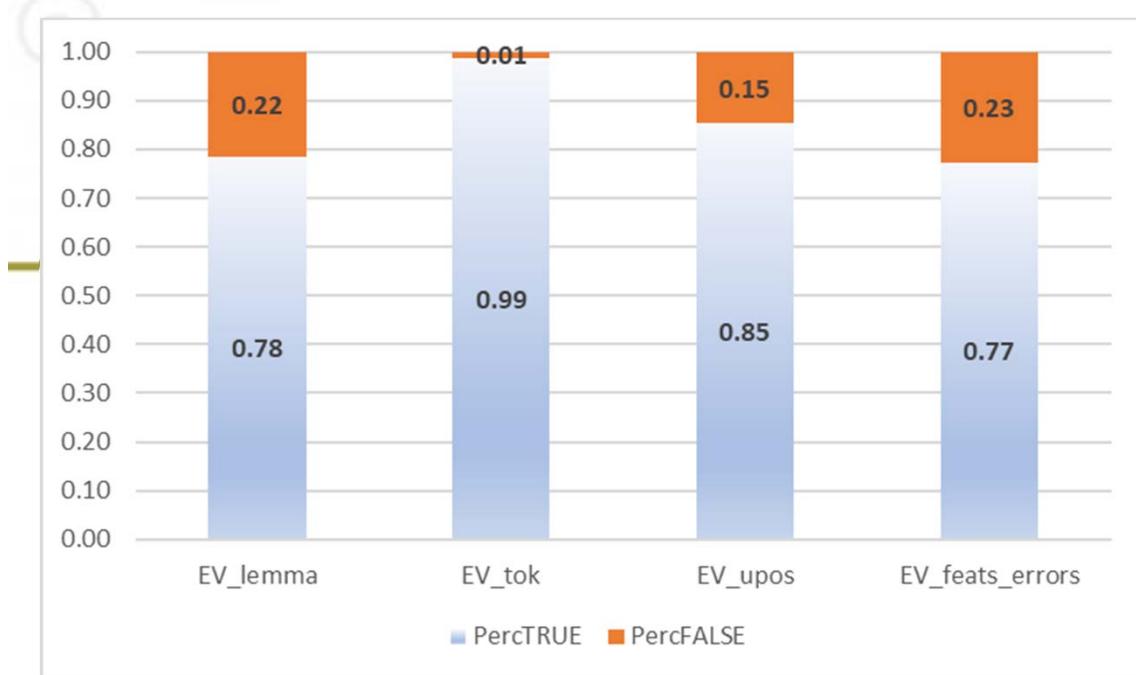
- 2.1: tokenizacija i podela na rečenice
- 2.2: morfo-sintaksička anotacija (vrsta reči, gramatičke kategorije)
- 2.3: lematizacija (u praksi spojena sa prethodnom tačkom)
- 2.4: prepoznavanje imenovanih entiteta (NER)
- 2.5: upravni govor
- 2.6: kompleksniji zadaci klasifikacije, uključujući i analizu sentimenta

Morfosintakška anotacija



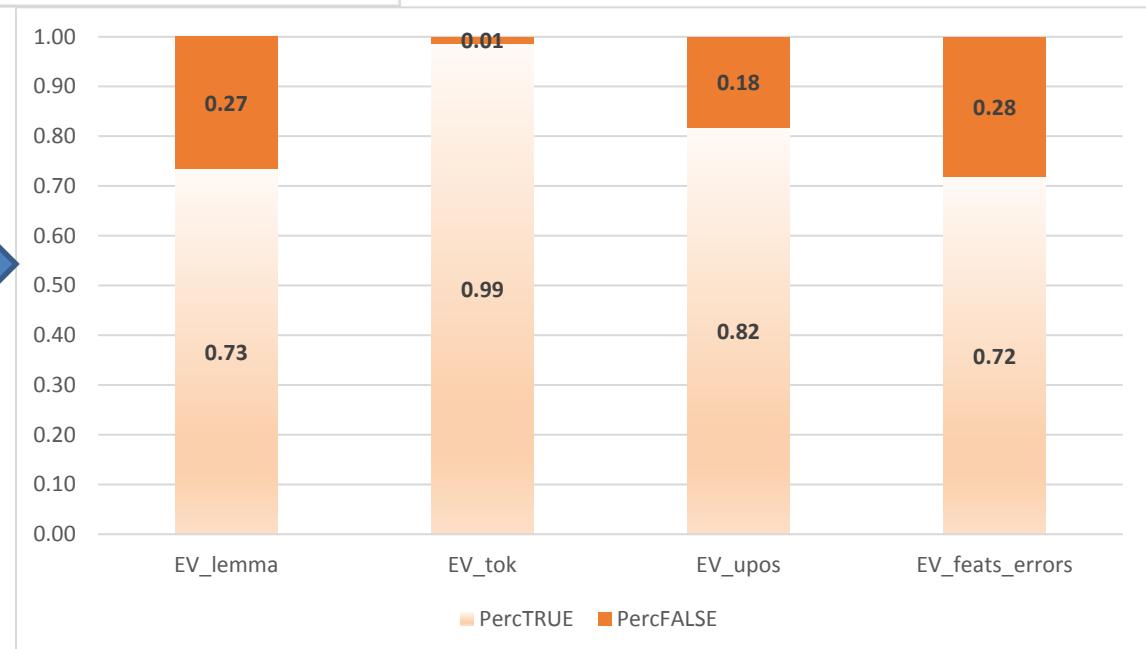
- U prvoj fazi 8 jezika: Czech, English, French, German, Hungarian, Italian, Norwegian-Nynorsk (both varieties), Portuguese, Slovene and Serbian (nedavno added).
- UD-Pipe tager, dodatno za Mađarski korišćen specijalizovani tager
- Srpski: automatski anotiran UD-pipe, evaluaciju uradila Milica Antić
- Anotirano naknadno tagerima razvijenim u okviru društva Jerteh, evaluacija u toku

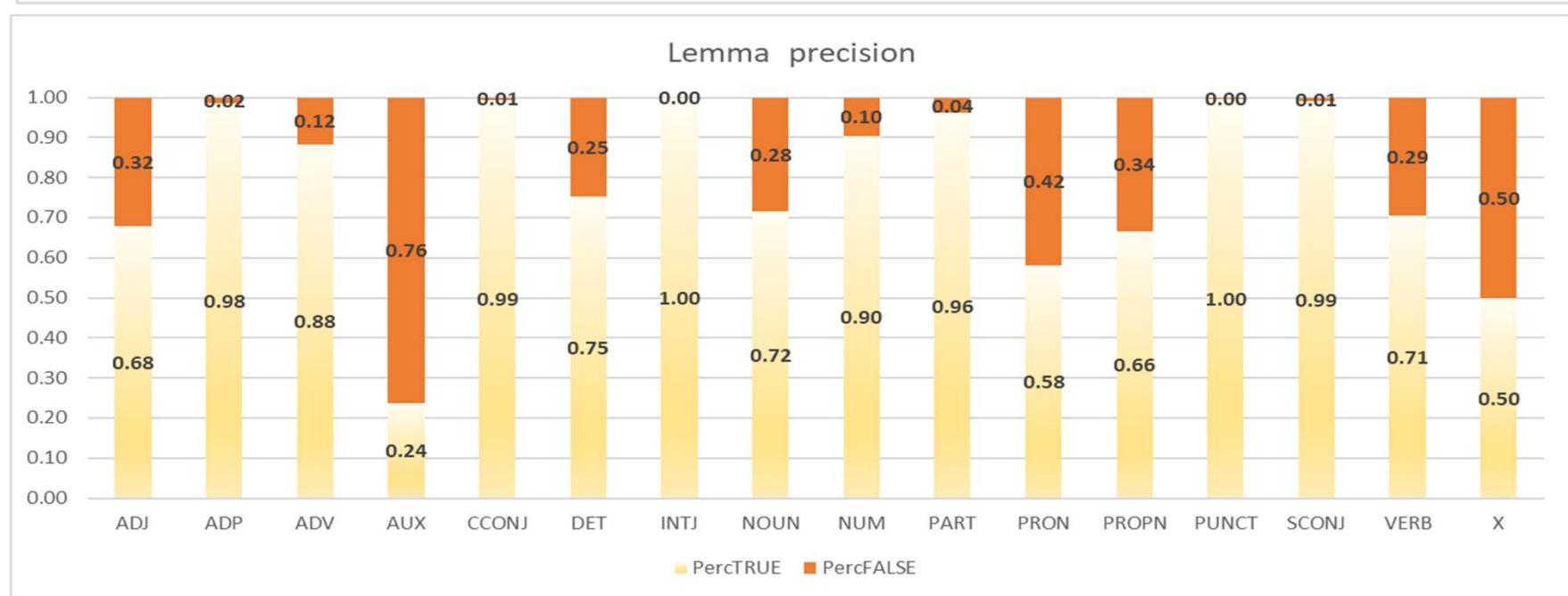
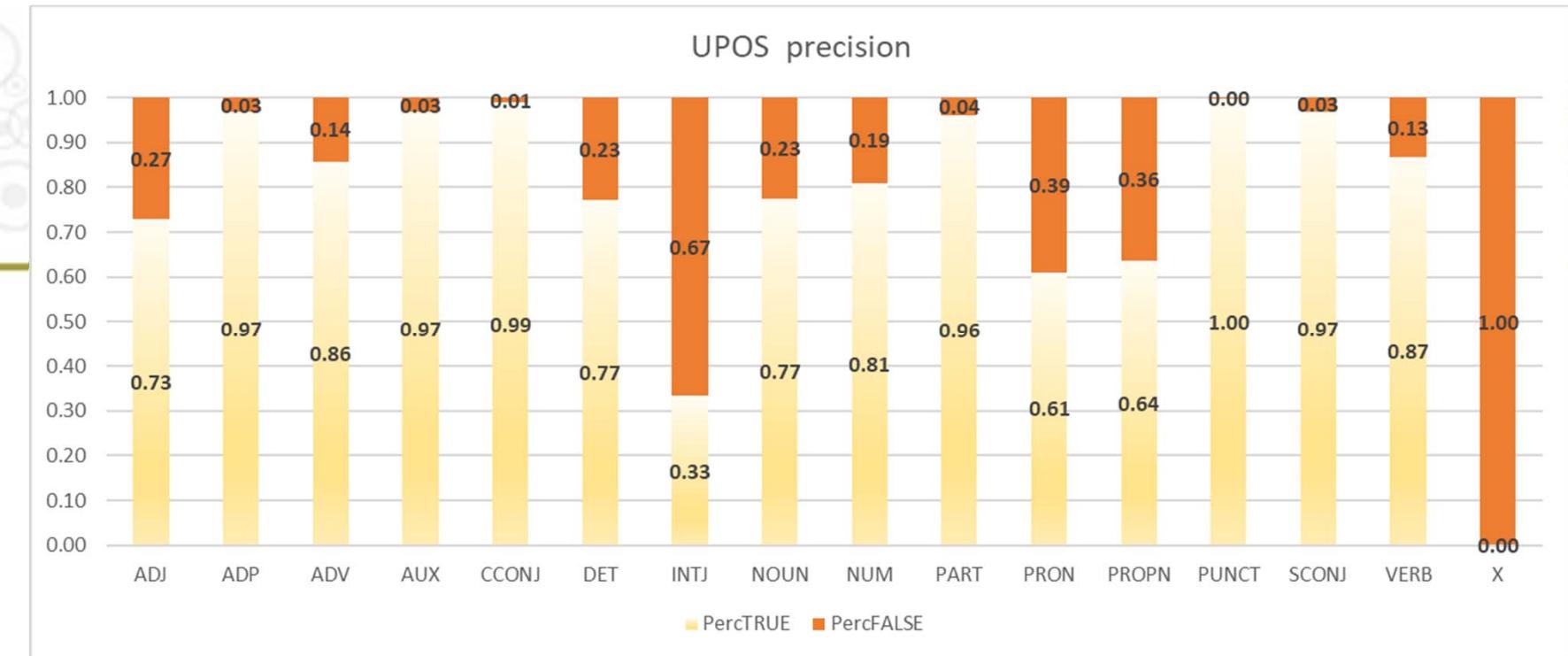
sentID	ids	form	lemma	upos	xpos	feats	EV_lemma	EV_tok	EV_upos	EV_feats_errors
1	5	ali	ali	CCONJ	Cc	_	0	0	0	0
1	6	posao	posao	NOUN	Ncmsan	Animacy=Inan Case=Acc Gender=Masc Nur	0	0	0	1
1	7	u	u	ADP	SI	Case=Loc	0	0	0	0
1	8	kući	kuća	NOUN	Ncfsl	Case=Loc Gender=Fem Number=Sing	0	0	0	0
1	9	beše	beše	VERB	Vmr3p	Mood=Ind Number=Plur Person=3 Tense=P	1	0	0	2
1	10	u	u	ADP	Sa	Case=Acc	0	0	0	0
1	11	taj	taj	DET	Pd-msan	Animacy=Inan Case=Acc Gender=Masc Nur	0	0	0	0
1	12	mah	mah	ADV	Rgp	Degree=Pos	0	0	1	1
1	13	preči	preči	VERB	Vmr3s	Mood=Ind Number=Sing Person=3 Tense=P	0	0	1	5
1	14	,	,	PUNCT	Z	_	0	0	0	0
1	15	s	sa	ADP	Sg	Case=Gen	1	0	0	0
1	16	toga	taj	DET	Pd-nsg	Case=Gen Gender=Neut Number=Sing Pron	0	0	0	0
1	17	nemade	nemada	VERB	Vmr3p	Mood=Ind Number=Plur Person=3 Tense=P	1	0	0	1
1	18	kad	kad	ADV	Rgp	Degree=Pos PronType=Int,Rel	0	0	0	2
1	19	nikom	nik	NOUN	Ncmsi	Case=Ins Gender=Masc Number=Sing	1	0	1	3
1	20	ni	ni	CCONJ	Cc	_	0	0	0	0
1	21	da	da	SCONJ	Cs	_	0	0	0	0
1	22	kaže	kazati	VERB	Vmr3s	Mood=Ind Number=Sing Person=3 Tense=P	0	0	0	0
1	23	što	što	PRON	Pi3n-n	Case=Nom Gender=Neut PronType=Int,Rel	0	0	0	0
1	24	ču	č	VERB	Vmr3p	Mood=Ind Number=Plur Person=3 Tense=P	1	0	0	2



Svi tkeni

Bez interpunkcije

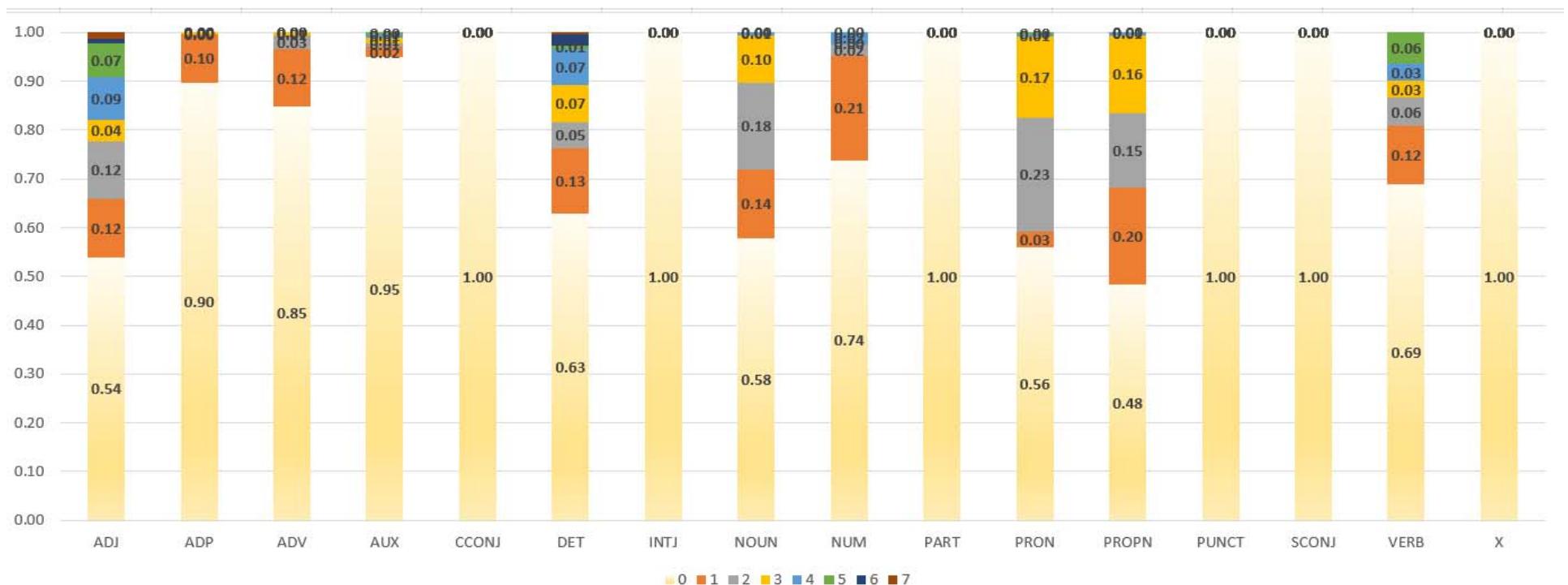




Gramatičke kategorije



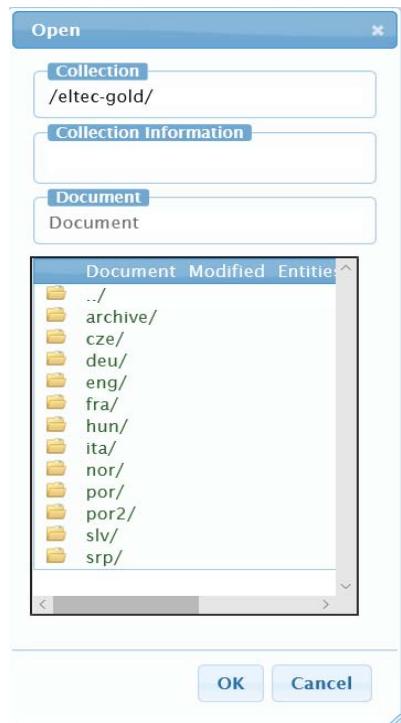
- Broj evidentiranih grešak u dodeljivanju kategorija
(samo preciznost, nije odziv uzet u obzir)



Ručna anotacija imenovanih entiteta



- Jezici: Czech, English, French, Hungarian, Norwegian, Portuguese, Serbian, Slovene + German
- demonimi, profesije i druge role, zanimanja, mesta, organizacije
- 5 pasusa po 400 tokena iz 20 novela za svaki jezik, ručno anotirano u bratu dostupno na <http://brat.jerteh.rs/>



Entity Type	Value
ORG: ?	ORG: ?
	ORG: Government
	ORG: Administrative
	ORG: Community
	ORG: Political
	ORG: Workplace
	ORG: Education
	ORG: Religious
EVENT: ?	EVENT: ?
	EVENT: Bussiness
	EVENT: Natural_disaster
	EVENT: Political
	EVENT: Culture
	EVENT: Sport
	EVENT: Religious
	EVENT: War
	PLACE: ?
PLACE: Mountain	
PLACE: City	
PLACE: Country	
PLACE: Region	
PLACE: Waterbody	
PLACE: Village	
PLACE: Astronym	
PLACE: Continent	

Primer: Jurmuša Fatima



NER Guidelines | NER&Beyond | Path: /eltec-gold/srp/SRP18790_MilanDjM_JurmusaFatima

14 <p n="SRP18790448">Ономадне ноћи, у глуво доба, Дели-топ је у три маха рикао: Ay! ay! ay! И наши се сад боје да Власи, пошто су устали на

DEMO

PLACE PLACE PLACE
Крушевац и на Алексинац, не ударе и на Ниш!..</p>

15 <p n="SRP18790449">— Та несрећа може и бити, рече Зелида, жена тефтердарева: — Цар сада има велику муку с Мисирлијама, па овамо не може ни послати војске</p>

DEMO

16 <p n="SRP18790450">Узалуд је стара Френчевићка намигивала оком, узалуд је и ногом назгизила Тевтердаревицу, да прекине тај разговор,

PERS [Female] PERS [Female] DEMO PERS [Female]
Марија и Милана већ чуше да су Срби устали на оружје, и да скоро може бити „што горе то доле“, како вељаху Нишлије.</p>

DEMO

17 <p n="SRP18790451">Кад би једнога дана после по дне, а Акила призове снахе своје, и овако им рекне :</p>

DEMO

18 <p n="SRP18790452">— Ваше ће мајке доћи у Ниш, и ви ћете ићи пред пашу да кажете: хоћете ли да останете туркиње, или волите да се вратите у

Србе ? За то упamtите што вам сада кажем: паша је брат мој, а моји су синови, ви знате, и сами силни. Ако кажете да сте силом потурчене, и ако

пожелите да се вратите у Србе, моји ће вас синови дочекати на Чамурлији, исећи ће вас на комаде, и ваше ће режњеве повешати по трњу.“</p>

19 <p n="SRP18790453">И доиста Смиљана и Магдалена дођу, суоче се пред пашом са својим ћерима , али Марија и Милана, у страху,

кажу пред пашом овако :</p>



[←](#) [→](#) [C](#) [Није безбеди](#) [brat.jerteh.rs](#)

NER Guidelines | NER&Beyond | Path: /eltec-gold

| no use trying to sneak past. Be bold! </p>

83 <p n="ENG19060702">Robert answered this appeal by



unexpectedly bursting into "The British Grenadiers," and to its

quick-step they approached the gates of Babylon.</p>

посла по письмам труда дочитали и все письмами брос, нау

ROLEMISC

буде , угодна; ако му не буду по волји њени јунаци, или с

PERS [Fiction][Male]

то Мојсије и Соломун!...</p>

чи разбуде у читаоца така осећања каква не би изазвала нава мисао; мисао осветљава мисао: мисао је мајка мисли; мисао

Edit Annotation

Text

Мојсије Link

Search

Google, Wikipedia, UniProt, EntrezGene, GeneOntology, ALC

Entity type

- PERS
- ROLE
- DEMO
- ORG
- PLACE
- WORK
- EVENT
- FAC
- ROLEMISC
- MISC
 - TEMPORAL
 - QUANTITY
 - OTHER

Entity attributes

PERS_REALITY: Fiction ▾ PERS_GENDER: Male ▾

Notes

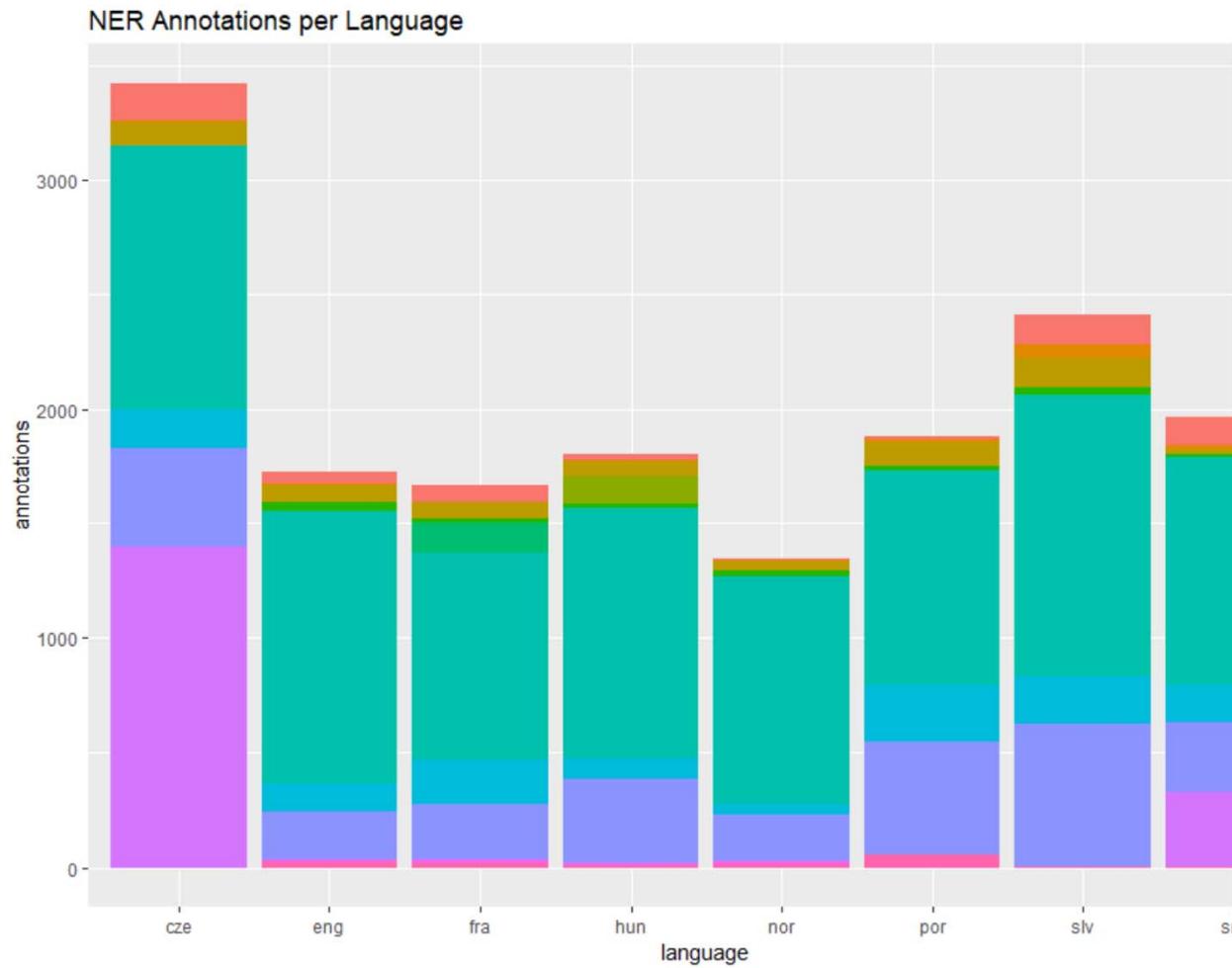
Add Frag. Delete Move OK Cancel



Rezultati ručne anotacije

	eng	fra	por	hun	nor	slv	srp	cze
Words	53093	52031	53958	58825	52900	53348	67232	52650
NE	1744	1669	1883	1805	1344	2412	1962	3424
DEMO	56	77	17	29	4	133	121	163
EVENT	7	3	9	7	8	54	18	5
FAC	75	70	103	65	41	128	23	108
ORG	37	22	19	20	25	37	11	0
PERS	1186	900	940	1091	990	1230	995	1150
PLACE	126	192	248	87	42	208	163	167
ROLE	203	244	490	367	201	620	301	454
WORK	25	18	54	7	10	2	4	0

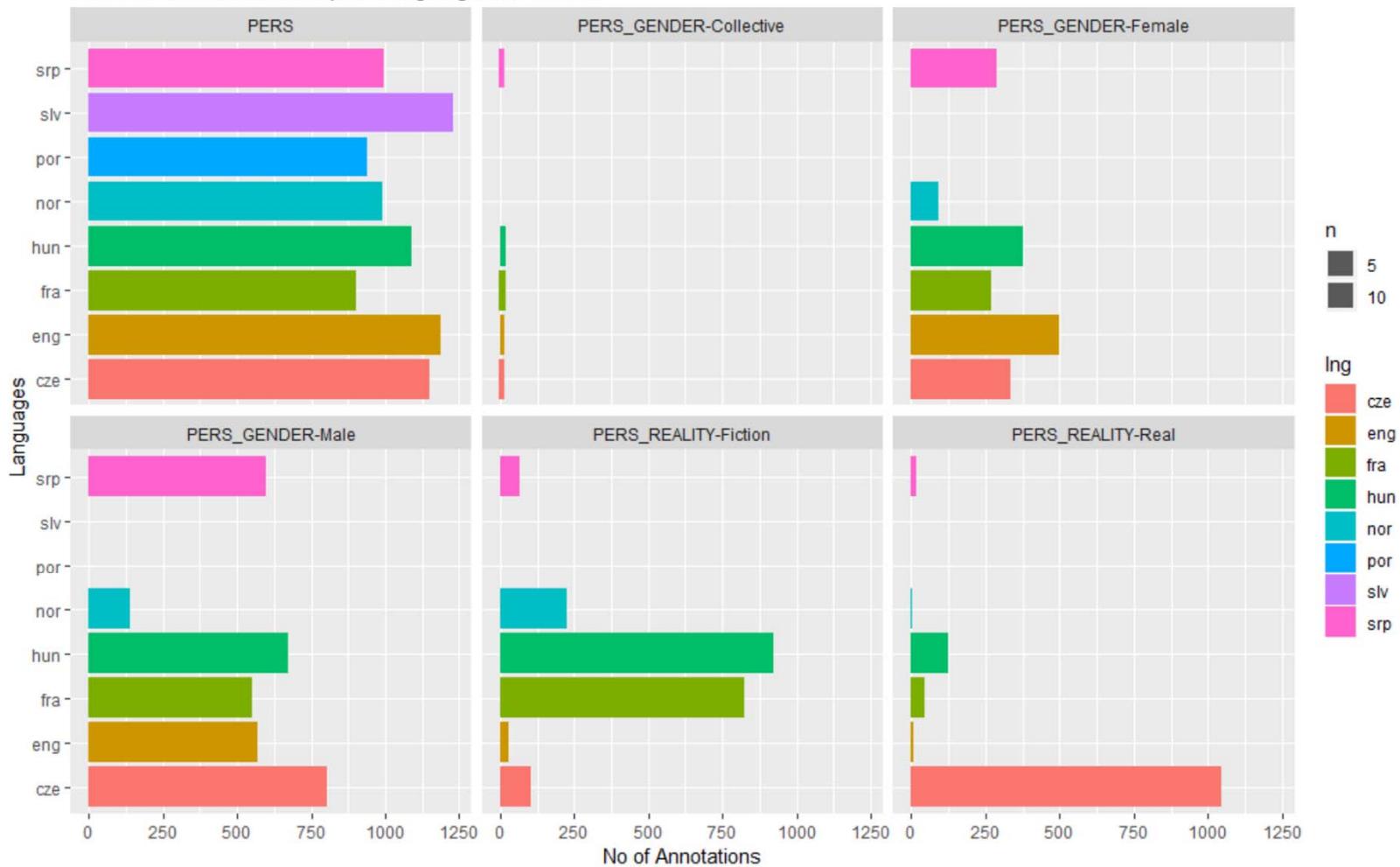
NER anotacija po jezicima



Anotacija PERS



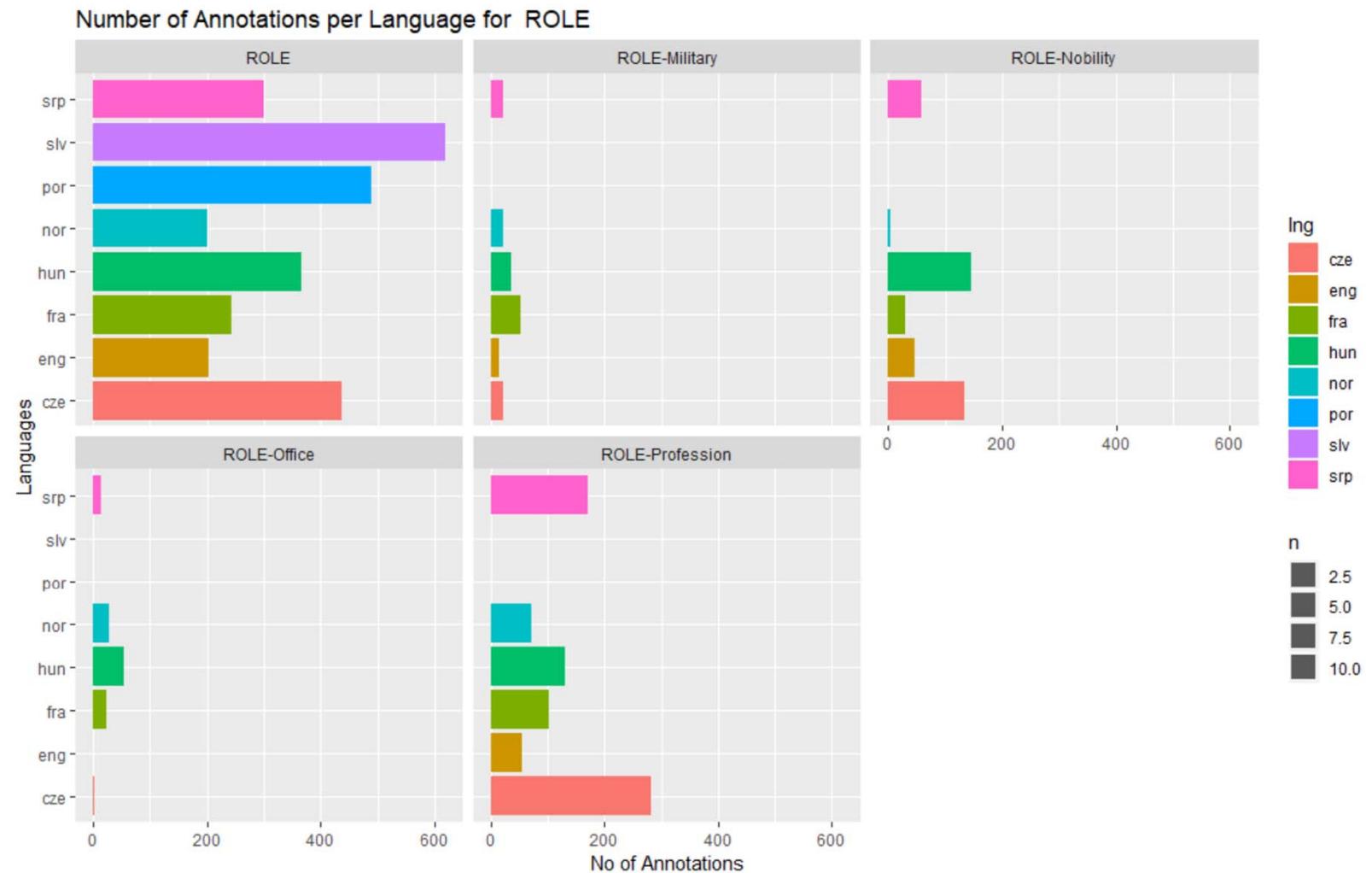
Number of Annotations per Language for PERS



n
5
10

Ing
cze
eng
fra
hun
nor
por
slv
srp

Antacija ROLE





Lang	Types	Professions (word forms)
cze	195	kněz (priest) 5, rychtáři (magistrate) 5, plavec (swimmer) 6, knížete (knight) 6, rychtář (magistrate) 7, knížecí (knight) 11, kníže (knight) 12
eng	91	Sir 6, Dr. 6, Secretary 6, Colonel 7, General 8, President 10, Professor 13, Queen 13, Lord 16
fra	147	domestique (servant) 5 colonel (colonel) 6 comte (count) 6 capitaine (captain) 9 général (general) 12
hun	249	tanár (teacher, professor) 7 tiszttartó (steward, curator, bailiff) 7 gróf (count) 8 huszár (hussar, cavalry man) 10 úr (sir, lord, mr.) 11 grófnő (countess) 11 ur (sir, lord, mr.) 15
nor	116	Kapellanan (chaplain) 6, Kaptein (captain) 6, Professoren (professor) 6, Unge-Konsulen (consul) 6, Styrmanden (manager) 7, Konsulen (consul) 10, Kapteinen (captain) 11
por	221	carcereiro (jailer) 9, conselheiro (counsellor) 9, pagem (page) 9, tenente (lieutenant) 9, bispo (bishop) 10, doutor (doctor) 13, padre (priest) 17, fidalgo (nobleman) 18
slv	386	kapitan (captain) 7, sodnik (judge) 7, kmetje (peasants) 7, grajščak (nobleman) 7, grof (count) 8, vitez (knight) 13, doktor (doctor) 15, profesor (professor) 16
srp	301	лекар (physician) 5, ђата (junior clerk) 5, учитељ (teacher) 5, кмет (village leader) 6, паша (pasha) 6, биров (police-man) 6, Газда (boss) 6, дућанџија (owner of the small shop) 6, поп (priest) 7

Tekstometrija nad srpELTeC



textometrie.jerteh.rs/

SRPROMAN Admin My profile Logout Help Contact en

Text: SRP18800_DragocenaOgrlica SRPROMAN:[srpos="A"][[srlemma="село"]]

Query : [srpos="A"][[srlemma="село"]]

Reference	Context
1 SRP18800_DragocenaOgrlica	да вам напоменем, да је она била поцелана на два три завађена табора. Ови су, чим се који дочећавао власти, један другог гонили и прогонили. Кад су међусобни раздори и сваковрсна подметања прешла сваку меру, кад су таласи злоупотреба и насиља почели запљускавати у питома села и у мирне вароши, — онда букуну револуција. У тој земљи потече братска крв. Тешки ударац револуције заљуљаше, из темеља потресоше и саме друштвене установе. Шта је невиних живота у овој борби пропало?... Гром је ударао у стогодишњи дуб. Овај
2 SRP18910_CyedomirM_Ik...	очита, те Гавран, плакајући у глас као мало дете, одмиле до пред царске двери и причестише га Одоше гости одабрани да под вењак седну за совру. Посадише игумана дечанској у врх совре, па онда калуђере и попове све по старешинству, па онда кнезове од суседних села и друге главне људе. Навалише и игуман и калуђери и сви гости на Иконију да седне, али их она смерно мольаше да је пусте да њих, као најмилије госте, које је икада у своме животу имала, двори на овај велики дан за њу. И
3 SRP18920_RajkoOdRasine	. ја вас, а Бог све нас. Нека му је слава! " 15 У вече тога дана, и кад се већ добро смркло, велика једна гомила сувог грања горела је на највишо и најширо зидини од развалиног града. То је стари пустиник давао знак најближем селу под брдом, да му се пошће један или два поуздана человека. Требали су му да проведу Рајка и Маргиту кроз Венчац до Космаја. Маргита је сама спавала у пећини, а старапац и Рајко седели су горе на градском зиду поред ватре. „ Ти си ми
4 SRP18920_RajkoOdRasine	свима потрама, као да ће сваког часа да се разглave у све саставке своје. Разбудише се и такнуже срдио гуске и патке, што, дуж брлога испред кућа, беху задремале; излетеше пси из авлија и кидишаце на кола и коње што само луњавом будише мир успаваног села . Маргите се поче да стеже лед око срца, кад опази да нигде у селу нема свеће или ватре у кући, ни живе человека на улици. Тек на самом излазу из главне, и управо једине, улице сеоске у алеју од јабланова, опазише једног сељака
5 SRP18920_RajkoOdRasine	кад су стигли на вис на оно место, где су недавно Рајко и Маргита тако лепе снове о лепијој будућности сновали. Онде, у присоју под оном стеном, ископаше гроб за Рајка. Кад већ гроб би доста дубок, Маргита нареди да српски сељаци побију на најближе село , где има српски поп, да га доведу да очита молитву и гроб препије. Па онда рече Ружи, Мартину и свима својим маџарским слугама: сад мени за љубав пођите овуда у околину и наберите польска цвећа. Хоћу да сваки од вас, који сте моји
6 SRP18960_TadijaK_Gosp...	лука. Кад је Татомир дотрчко и јавио да је дошло „званично“, скучило се све живо из варошице, сем десетак затуцаних приврженика. Продана Жмурића, који се нису помањали из својих кућа. Озбиљнији људи ушли су у механу и поседали, а остали набавили из облизијег села гајдаша и цигане са зурлама и бубњевима, па направили весеље око ватре. Прво катранјаво буре запаљено је пред трговином Петра Дрндаревића, а друго пред магазом Дине Цекића, (коме је онај народ око ватре намеравао да викне: Yai — али пошто је запалио буре
7 SRP19000_MilutinU_Dosljaci	сва зелена. Од ње се пуштао бескрајан видик, као над океаном, који је гутао широку траку Дунава, његова пошумљена острва, мирне белуџаве рукаве и ритове, обрасле у бледу трску. На само ивици хоризонта, далеко тамо на северу, смешило се једно чисто банатско село , а доле у дну, камо је текла сва дунавска вода, назирала се у сивини даљине два готска цркве у Панчеву. Обухвативши једним погледом цео тај видик, Милош уздизао се напрштиени Грот, као каква горостасна старицу у блузи од кадифе са шеширом од црних
8 SRP19000_MilutinU_Dosljaci	терасе на Калимегдану. Висока брда, обрасла ситним или крупним зеленилом, проширене сивим стенама и засађене питомим пропланцима, долине, клисуре, урвине, попла, реке, њиве усред шуме, шљивац, белина брезе у мркој растовој гори, пустиња, питомина, варош, растурено село , све то вешто измешано међу собом руком неког тајанственог уметника, даје чудну и живописну слику која се ретко другде може наћи. Више Милоша уздизао се напрштиени Грот, као каква горостасна средњевековна кула; на западу, према њему равњао се Бистош; а према овоме чамци
9 SRP19000_MilutinU_Dosljaci	, већ је продужио пут ноћним возом. Уморан од вожње у неудобним поштанским колима, заспао је одмах после Лапова. Већ је било уврлико свануло кад се пробудио. Око њега сасвим други предео. Није било више дивних ужиčких брда обраслих у шуму и бујад, ни расутих села где се сељаци довикују с брда на брдо, ни посних њива где мршава стока крви леђа оруђи камен и прљушу. Воз је јурио кроз таласасте равнице Шумадије. Унаоколо је било све обрађено и питомо. Огромне њиве бујног кукуруза таласале се око железничке пруге као надаше простране воде. Подмлађена трава у ливадама одисала је польским мирисимима. Из рогоза се дизала магла. Живе ограде, начињене од ситног багрења, биле су се урешиле сребрним ѡерданцима росе. С времена на време угледало би се какво богато ушорено село , где се јасно распознавала школа и судница. Овде онде видели су се радници, људи и жене, који су остављали рад и зачујено гледали у ватрена кола што су хуктала пред њима. Поред сваког сумњичења које га је мучило, Милош је осећао струјање једног пријатног
10 SRP19000_MilutinU_Dosljaci	таласале се око железничке пруге као надаше простране воде. Подмлађена трава у ливадама одисала је польским мирисимима. Из рогоза се дизала магла. Живе ограде, начињене од ситног багрења, биле су се урешиле сребрним ѡерданцима росе. С времена на време угледало би се какво богато ушорено село , где се јасно распознавала школа и судница. Овде онде видели су се радници, људи и жене, који су остављали рад и зачујено гледали у ватрена кола што су хуктала пред њима. Поред сваког сумњичења које га је мучило,

56 occurrence(s)

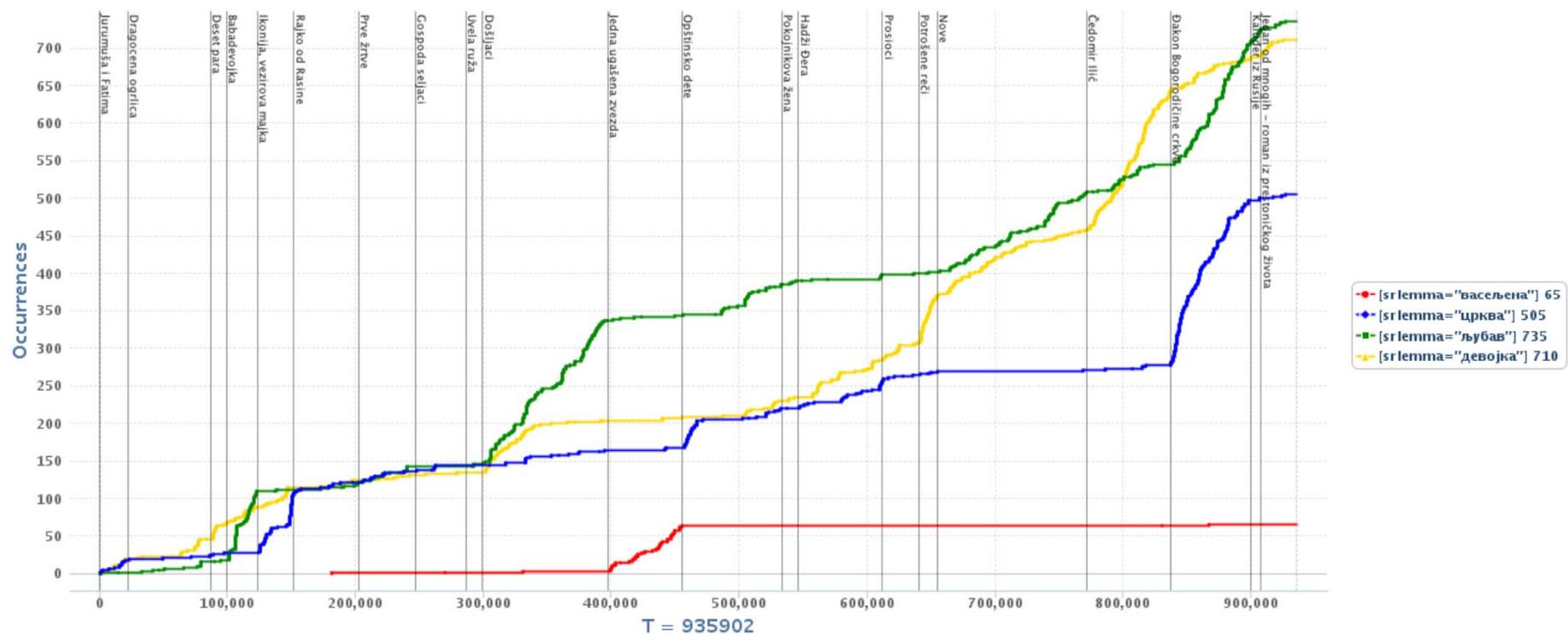
1-56/56 Pagination 100 Export

Tekstometrija nad srpELTeC



- Više detalja u radu Jelene Jaćimović u Infoteci

Progression of [srlemma="васељена"], [srlemma="црква"], [srlemma="љубав"], [srlemma="девојка"] in srpELTeC (structure: text, property: title, filtering REGEX:)



Mera specifičnosti pojedinih vrsta reči prema autorima



Author	f_N	S_N	f_V	S_V	f_A	S_A	f_{ADV}	S_{ADV}
Uskokovic M	35331	-11.2	35446	52.5	13502	48.4	9577	-0.8
Komarcic L	31869	6.2	25443	-104.4	13193	205.6	7481	-23.1
Dimitrijevic J	22330	-37.4	21938	0.5	7066	-29.3	5687	-21.2
Nusic B	16926	16.6	14675	0.6	3812	-120.5	4952	28.9
Mijatovic C	15985	-1.2	13862	-16.3	6259	37.9	4276	-0.4
Ilic D	13728	2.4	12740	6.1	3319	-82.4	3684	1.6
Sekulic I	12497	-3.3	11826	1.1	4772	16.4	4183	42.7
Gavrilovic A	10879	83.6	8125	-3.8	3215	2.7	2356	-1.7
Kostic T	10276	162.7	6606	-24.0	1983	-50.0	1411	-69.5
Milicevic M	7464	8.1	6140	-4.6	1981	-15.3	1680	-5.9
Gavrilovic D	4105	-39.0	5197	32.8	1417	-7.9	1635	20.1
Stankovic B	3867	-108.0	5126	8.5	1268	-34.4	1731	20.0
Siskovic D	3937	-18.4	4838	40.6	981	-46.3	1445	14.1
Jankovic M	1371	-19.8	1860	17.2	539	-0.9	530	4.0

The frequency of nouns (N), verbs (V), adjectives (A) and adverbs (ADV) by authors and their *specificity score* for the whole corpus

проза

Веселиновић, Јанко

Хајдук Станко



Вук Милићевић

Беспуће



Гавриловић, Андра

Прве жртве: приповетка из српске прошлости



Гавриловић, Драга

Бабадевојка



Глишић, Милован Ђ.

Приповетке

Приповетке, књига 1

Приповетке, књига 2

Приповетке, свеска 1



Димитријевић, Јелена

Нове : роман



Домановић, Радоје М.

Приповетке Радоје М. Домановића



<http://aurora.jerteh.rs/>

aurora.jerteh.rs наводи:

OCR и корекција текста : Ранка Станковић, Цветана Крстев

Кодирање за ELTeC : Цветана Крстев

Aurora

Игњатовић, Јаков

Вечити младожења

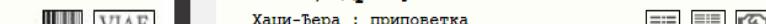
Приповетке

Приповетке



Илић, Драгутин Ј.

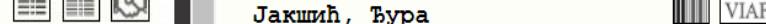
Хаши-Ћера : приповетка



Јакшић, Ђура

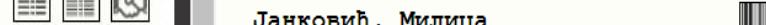
Приповетке, свеска 1

Сирота Банаћанка :
приповијетка



Јанковић, Милица

Калуђер из Русије

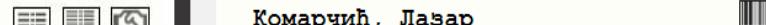


Комарчић, Лазар

Драгоценна огрилица: прича у своје време

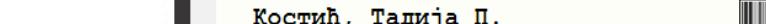
Једна угашена звезда :
илюстровани роман

Просоци



Костић, Тадија П.

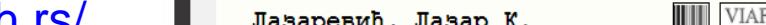
Господе сељаци: приповетка



Лазаревић, Лазар К.

Приповетке, свеска 1

Приповетке, свеска 1
Школска икона



Мијатовић, Чедомил

Рајко од Расине: приповетка с краја XVII века



Милићевић, Милан Ђ.

Десет пара: прича из живота у вароши

Јурмуса и Фатима или
Турска сила сама себе
једе: прича о ослобођењу шест округа 1832-1834



Нушић, Бранислав

Општинско дете



Светолик Ранковић

Горски цар

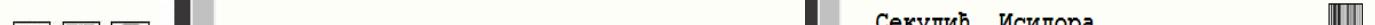
Јесење слике

Сеоска учитељица



Секулић, Исидора

Ђакон Богородичине цркве



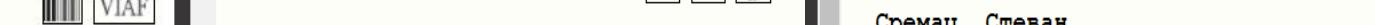
Сремац, Стеван

Две приповетке

Ивкова слава

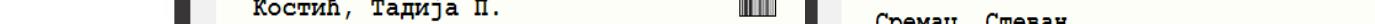
Лимунација у
селу

Поп Ђира и поп Спира



Сремац, Стеван

Божићна печеница



• I još mnogi
drugi



Баро 3
Баром 3
Бару 11
барут 12
барута 11
барутног 1
барутом 1
баруштина 1
Батали 2
батли 2
бато 2
бату 1
бахат 1
Бахат 1
Бахне 1
бахне 2
бахну 7
бахнуше 1
баца 2
бацали 1
бацао 3
бацати 1
бацаху 1
баци 20
Баци 3

</p> <p rend="Tekst">џебана – муниција; **барут** и олово </p> <p rend="Tekst">џумбус – забава, игре</p> <p rend="Tekst">— Па?!{S} А пушке, а **барут**?{S} Кад то покисне, шта онда мислиш?{S} Мислиш да "И дође...{S} Логором прође глас да је **барут** закисао и да немају ни зрна сува.{S} Јуди шапутах {S} У то буре Лазар је бацио варницу, и **барут** је букнуо... </p> <p rend="Tekst">Сад – не зауста...{S} Тамо ја идем где се крв пролева и **барут** мирише, мене ће овде угушити дим у кући!...{S} Ја лио...{S} Ти ваљда мислиш да мени смрди **барут**, шта ли?... </p> <p rend="Tekst">— Опрости, дево! – рупица на пушци или топу где се пали **барут** </p> <p rend="Tekst">ведити – венути, нестајати <, онда за ножеве.{S} Грехота је трошити **барут** без потребе.{S} Сад сваки на своје место и – таја вако потраје још два дана, неће се моћи **барут** сачувати! – рече један. </p> <p rend="Tekst">— Бо се прибијаху један уз другог, чувајући **барут** и оружје да не закисне.{S} Небом се гањаху сиви о о бостан!...{S} Него... шаљи што пре по **барут**. </p> <p rend="Tekst">Чупић се мишљаше кога да по </p> <p rend="Tekst">— Ја сам послao по **барут**. </p> <p rend="Tekst">— Послаo!{S} А хоћe ли наћi обро. </p> <p rend="Tekst">— Нама треба **барута**.{S} Нађи где знаш и што више можеш...{S} А ево тeзврти пут. </p> <p rend="Tekst">— Нема **барута**!...{S} Да јуришамо! – виче Заврзан. </p> <p rend="Tekst">И показа му два бурета **барута** што беху у једном ћошку шатора. </p> <p rend="Tekst">— Ала тај бежи од **барута**, као горско зверче!... – насмеја се Заврзан. </p> дарнута је клица зла.{S} А она је буре **барута**.{S} У то буре Лазар је бацио варницу, и барут је то би учинила ова шака соколова да беше **барута**; али барута нестаде. </p> <p rend="Tekst">Турци и му Рудничанин. – Да није због пушке и **барута**, па готово боља ова хладовина!... </p> <p rend="Tekst"> а ова шака соколова да беше барута; али **барута** нестаде. </p> <p rend="Tekst">Турци баш у тај маош. </p> <p rend="Tekst">— Ваља тражити **барута** – одговори Чупић. </p> <p rend="Tekst">— Где? </p> <p rend="Tekst">— Послаo!{S} А хоћe ли наћi **барута** колико треба?...{S} А Турцима може доћи ћеф да брече: </p> <p rend="Tekst">— Неста нам **барута**...{S} Шта да радимо? </p> <p rend="Tekst">Све обје ниси могао дахнути чиста ваздуха од **барутног** дима и прашине што је дизају копите коњске...

сторожнице... и да сада... —
познају...

— Онда, побратиме, добро ми дошао! – И Зека издаде налог да се крећу...

Пропиšта црна земља под грдним теретом... Турски буљуци ницаху са свију страна, као печурке. Град шабачки и поља око Шапца беху притиснута војском... Дигла се сила да уништи слободу; дигао се цар из Стамбола да згаси шаку јада, дојучерање робове своје...

Последњих дана јула и првих дана августа 1806. године само је пушка пуцала... Ту ниси могао дахнути чиста ваздуха од **барутног** дима и прашине што је дизају копите коњске...

И борило се, тукло се на све стране; оживеше честе и жбунови; иза сваког трна пушка је пуцала...

На једној страни заповедник врховни вожд, Кара-Ђорђе, с војводама: Луком Лазаревићем, Стојаном Чупићем, Милошем Потерцем, Јаком Ненадовићем, Цинцар-Јанком, ЦинцарМарком, Јанком Катићем из Рогаче и још многима... На другој Сулејман-паша, што га звају Кулин-капетан, с делијама босанским: Мехмед-капетаном из Зворника, Синан-пашом Херцеговцем, Мулом Сарајлијом, Беширевић Асом и другима.

Кара-Ђорђе је имао мало људи према турској сили. Оно што је пред Турке ставио, било је свега

Izazovi



- Lepo al nepraktično
- Evaluacije: različiti formati, razičite tokenizacije
- Prikazivanje u tabelarnim formatima struktuiranih anotacija i obrnuto
- Poravnanje sa drugačijim tokenizacijama
- XML –TEI format (stendof-inline poravnanje)
- Atributi u drugim sistemima retko moguće mapirati