



INŠTITUT
ZA NOVEJŠO ZGODOVINO



Digitalna humanistika in zgodovinopisje

Tjaša Konovšek
tjasa.konovsek@inz.si

Katja Meden
katja.meden@inz.si

Problemi zgodovinopisja (v povezavi z DH)

- obseg gradiva
- dostopnost gradiva
- jezik in pisava
- kvantitativna metoda
- ustvarjanje znanja
- cf.: [analovci](#) in [konceptualna zgodovina](#) (big data, longue durée, koncepti)



Digitalna humanistika

Definicije zagotavlja ponudnik [Oxford Languages](#) · [Več o tem](#)



digital humanities

noun

an academic field concerned with the application of computational tools and methods to traditional humanities disciplines such as literature, history, and philosophy.

"the unit will advance scholarship in both classical studies and **the digital humanities**"

Digitalna humanistika

- DH in jezikovne tehnologije
- CLARIN.SI in DARIAH-SI
 - <http://www.dariah.si/>
 - <https://www.clarin.si/info/o-projektu/>
- FF [digitalno zgodovinopisje](#)
- SI-DIH (Zbirke)
 - <https://sidih.si/20.500.12325/menu1>



Primeri DH rezultatov

Rezultati DH

- Korpusi:
 - Strukturno označeni korpusi (npr. z oznakami za stavčno analizo, semantiko, sentiment).
 - Zgodovinski ali specializirani korpusi (npr. parlamentarni, pravni, literarni).
- Digitalne znanstevnokritične izdaje:
 - Kritične izdaje besedil z digitalnimi prepisi/pripisi, faksimili in večjezičnimi prevodi.
- (Meta)Podatkovne baze
- Interaktivne vizualizacije
- Zemljevidi in GIS podatki
- Digitalni arhivi
- Ontologije in taksonomije:
 - Strukture znanja, ki kategorizirajo informacije za boljše iskanje in povezovanje
- Računalniški modeli in simulacije

Korpusi

kórp^us -a m (ô)

SSKJ²



1. *vojaška enota iz divizij in posebnih enot*: formirati korpus; poveljnik korp^usa / ekspedicijski korpus
2. jezikosl. *zbirka besedil v elektronski obliki s standardizirano obliko zapisa in izbranimi podatki*: analiza korp^usa; delo s korpusom; besedila za korpus; korpus in slovar / referenčni korpus *ki je glede na vključena besedila zvrstno kar se da uravnotežen / besedilni korpus*
3. publ., navadno s prilastkom *skupina*, *zbor*: nastopili so najboljši korpusi; baletni, godalni korpus / diplomatski korpus *diplomatski zbor*; korpus miru zlasti v ameriškem okolju *organizacija iz prostovoljcev, ki skrbi zlasti za gospodarski razvoj in pomoč nerazvitim deželam*
4. knjiž. *zbirka*: korpus srednjeveških fresk
 - ekspr. ta ženska je pa korpus *je zelo debela*

Korpusi

Attributes **ana** and **msd** both encode the morphosyntactic characteristics of the word in focus (in blue). The **ana** attribute contains Slovene-specific tags according to MULTTEXT-East Specifications, while the **msd** attribute contains the Universal Dependency tagset which greatly simplifies cross-lingual comparison.

The attribute values Q and PART both stand for *particle*. The difference between the two tags is in the tagset used.

The **word form** also known as a token from the running text in the corpus.

The attribute **lemma** indicates the basic word form of the token, that is of the word in focus (in blue).

The corpus also encodes syntactic parses and named entities but since they are not used in this tutorial they were omitted from this illustrative example.

```
<s>
<w ana="mte:Q" msd="UposTag=PART" lemma="zlasti">Zlasti</w>
<w ana="mte:Sg" msd="UposTag=ADP|Case=Gen" lemma="glede">glede</w>
<w ana="mte:NcmSg" msd="UposTag=NOUN|Case=Gen|Gender=Masc|Number=Sing"
lemma="nadzor">nadzora</w>
<w ana="mte:Va-r3s-n"
msd="UposTag=AUX|Mood=Ind|Number=Sing|Person=3|Polarity=Pos|Tense=Pres|VerbForm=Fin"
lemma="biti">je</w>
<w ana="mte:Pd-fsn" msd="UposTag=DET|Case=Nom|Gender=Fem|Number=Sing|PronType=Dem"
lemma="ta">ta</w>
<w ana="mte:NcfSg" msd="UposTag=NOUN|Case=Nom|Gender=Fem|Number=Sing"
lemma="stvar">stvar</w>
<w ana="mte:Rgp" msd="UposTag=ADV|Degree=Pos" lemma="zelo">zelo</w>
<w ana="mte:AgpfSg" msd="UposTag=ADJ|Case=Nom|Degree=Pos|Gender=Fem|Number=Sing"
lemma="kočljiv">kočljiva</w>
<pc ana="mte:Z" msd="UposTag=PUNCT">.</pc>
</s>
```

One pair of the opening and closing structural tags which, in our case, indicate punctuation (**pc**), word (**w**) or sentence (**s**).

Vzporedni prikazi

Vrsta prikaza (strani) ▾

Način vzporednega prikaza (fac-s-dipl-crit) ▾

Prelom vrstice:

Da

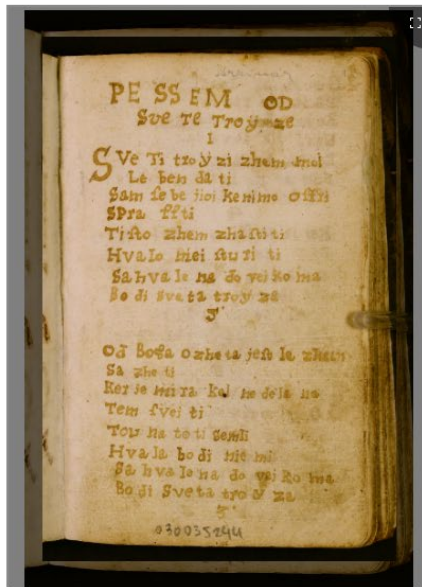
Ne

01v <<

stran

02r

>> 02v



PE SSEM OD

Sve Te Troj ze

1

Sve Ti troj zi zhem moi
Le ben da ti
Sam le be joi kenimo offri
Spra fti
Tišto zhem zha fti ti
Hvalo niei fti ri ti
Hvalo niei fti ri ti

Sahva le na do vei ko ma
Bo di sve ta troj za

2

Od Boga o zhe ta jest le zhem
Sa zhe ti
Ker je mira kel ne dela na
Tem fvei ti
Tou na to ti semli
Hva la bo di nie mi

Sa hva le na do vei ko ma
Bo di sveta tro y za

3

Pesem od Svete trojce

1

Sveti trojci čem moj leben dati,
sam sebe joj k enimo ofri sprafti.
Tisto čem častiti,
hvalo njej sturiti.

Zahvalena do vejкома
bodi Sveta trojca.

2

Od Boga Očeta jest le čem začeti,
ker je mirakelne dela na tem svejti.
Tou na to ti zemli
hvala bodi njemi.

Zahvalena do vejкома
bodi Sveta trojca.

3

Ditmajer, N., et al. (2021). *Foglarjeva pesmarica* ([Elektronska znanstvenokritična izdaja](#)).

Celotno besedilo

Celotno besedilo

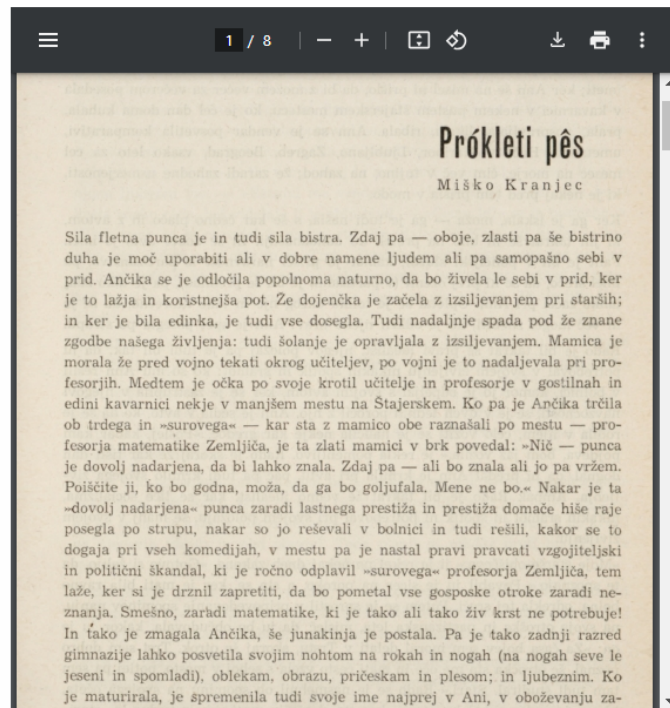
Besedilo dokumenta

Miško Kranjec (1908): Prókleti pês (1964)

>>

Juvan, M., et al. (2021). *Korpus Maj68* (Elektronska znanstvenokritična izdaja, korpus)

Sila fletna punca je in tudi sila bistra. Zdaj pa – oboje, zlasti pa še bistrino duha je moč uporabiti ali v dobre namene ljudem ali pa samopašno sebi v prid. Ančika se je odločila popolnoma naravno, da bo živila le sebi v prid, ker je to lažja in koristnejša pot. Že dojenčka je začela z izsiljevanjem pri starših; in ker je bila edinka, je tudi vse dosegla. Tudi nadaljnje spada pod že znane zgodbe našega življenja: tudi šolanje je opravljala z izsiljevanjem. Mamica je morala že pred vojno tekati okrog učiteljev, po vojni je to nadaljevala pri profesorjih. Medtem je očka po svoje krotil učitelje in profesorje v gostilnah in edini kavarnici nekje v manjšem mestu na Štajerskem. Ko pa je Ančika trčila ob trdega in »surovega« – kar sta z mamico obe raznašali po mestu – profesorja matematike Zemljiča, je ta zlati mamici v brk povedal: »Nič – punca je dovolj nadarjena, da bi lahko znala. Zdaj pa – ali bo znala ali jo pa vržem. Poiščite ji, ko bo godna, moža, da ga bo goljufala. Mene ne bo.« Nakar je ta »dovolj nadarjena« punca zaradi lastnega prestiža in prestiža domače hiše raje posegla po strupu, nakar so jo reševali v bolnici in tudi rešili, kakor se to dogaja pri vseh komedijah, v mestu pa je nastal pravi pravcati vzgojiteljski in politični škandal, ki je ročno odplovil »surovega« profesorja Zemljiča, tem laže, ker si je drznil zapreti, da bo pometal vse gosposke otroke zaradi neznanja. Smešno, zaradi matematike, ki je tako ali tako živ krst ne potrebuje! In tako je zmaga Ančika, še junakinja je postala. Pa je tako zadnji razred gimnazije lahko posvetila svojim nohtom na rokah in nogah (na nogah se ve le jeseni in spomladi), oblekam, obrazu, pričeskam in plesom; in ljubeznim. Ko je maturirala, je spremenila tudi svoje ime najprej v Ani, v oboževanju



Podatkovne baze



[PUBLIKACIJE](#) [ŽRTVE I. SV.](#) **[ŽRTVE II. SV.](#)** [POPISI](#) [ZIC](#)

Vnesite avtorja in/ali naslov

[/ ŽRTVE II. SV. VOJNE](#) [KOPIRAJ POVEZAVO](#) [SLOJI](#) vse

AA

ANTON ANTONČIČ  Inštitut za novejšo zgodovino

priimek Antončič

ime Anton

oče Janez

mati Jožefa

priimek matere Zorec

datum rojstva 21. 01. 1911

kraj rojstva Hrib ob Temenici

kraj smrti Teharje

BA

BOGOMIR ANTONČIČ  Inštitut za novejšo zgodovino

priimek Antončič

ime Bogomir

oče Gregor

mati Frančiška

priimek matere Mahnič

datum rojstva 23. 10. 1922

kraj rojstva Dolenja vas

kraj smrti Škrlovisa

DA

DUŠAN ANTONČIČ  Inštitut za novejšo zgodovino

priimek Antončič

ime Dušan

kraj smrti Mauthausen

FA

FRANC ANTONČIČ  Inštitut za novejšo zgodovino

priimek Antončič

ime Franc

oče Anton

kraj rojstva Vrh, Cerknica

kraj smrti Nanos

Drago Antonič Inštitut za novejšo zgodovino

Osební podatki

priimek **Antonič**ime **Drago**kraj rojstva **Komen**naslov **Komen**stara občina **Komen**nova občina **Komen**pokrajina **Primorska**

zakonski stan ?

poklic (soc. status) **delavec**narodnost **Slovenec**spol **moški**

Podatki o smrti

vojni status **NOV in POS**država smrti **Slovenija**država pokopa **Slovenija**vzrok smrti **Oborožen spopad**

povzročitelj smrti ?

Viri

Inštitut za novejšo zgodovino

- Pričevanja 1941-1945, Sveto 1997, str. 25.

Opombe

Inštitut za novejšo zgodovino

- RMK Komen ga ne navaja



To delo je ponujeno pod [Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna](#)

Primeri DH študij

Primer 1: Parlament in čas

- Teza: V dobi moderne je čas postajal pomembna dobrina, z njim je bilo treba ravnati *racionalno* in zato je bilo treba delovati *učinkovito*.
- Slovenski parlamentarni korpus siParl 2.0 (NoSketch Engine), korpusno jezikoslovje
- Besede, ki izražajo dojemanje poslanskega zbora (*čim prej, hitrost, učinkovito, neučinkovito, racionalen , čas je*).
- Luščenje kolokacij; ustvaril sem seznam besed, ki se neposredno ob jedrni besedi (pred njo in za njo), povezani s časom, pojavljajo pogostejše kot naključno
- Zaključki:
 - **Neučinkovito = neracionalno = počasno**
 - V parlamentu je čas še največkrat »**skrajni**«.
 - Sama beseda čas se v vseh sklonih v parlamentu pojavi kar 262.964-krat

[2] Tabela 4: Čimprej – kolokacijski kandidati^[29]

		<u>Cooccurrence count</u>	<u>Candidate count</u>	<u>logDice</u>
<u>p</u> <u>N</u>	predloži	147	4.480	8,691
<u>p</u> <u>N</u>	čimprej	103	7.057	7,890
<u>p</u> <u>N</u>	pripravi	302	43.764	7,602
<u>p</u> <u>N</u>	sprejme	218	33.481	7,457
<u>p</u> <u>N</u>	sprejeti	291	47.717	7,440
<u>p</u> <u>N</u>	sprejmemo	136	20.278	7,343

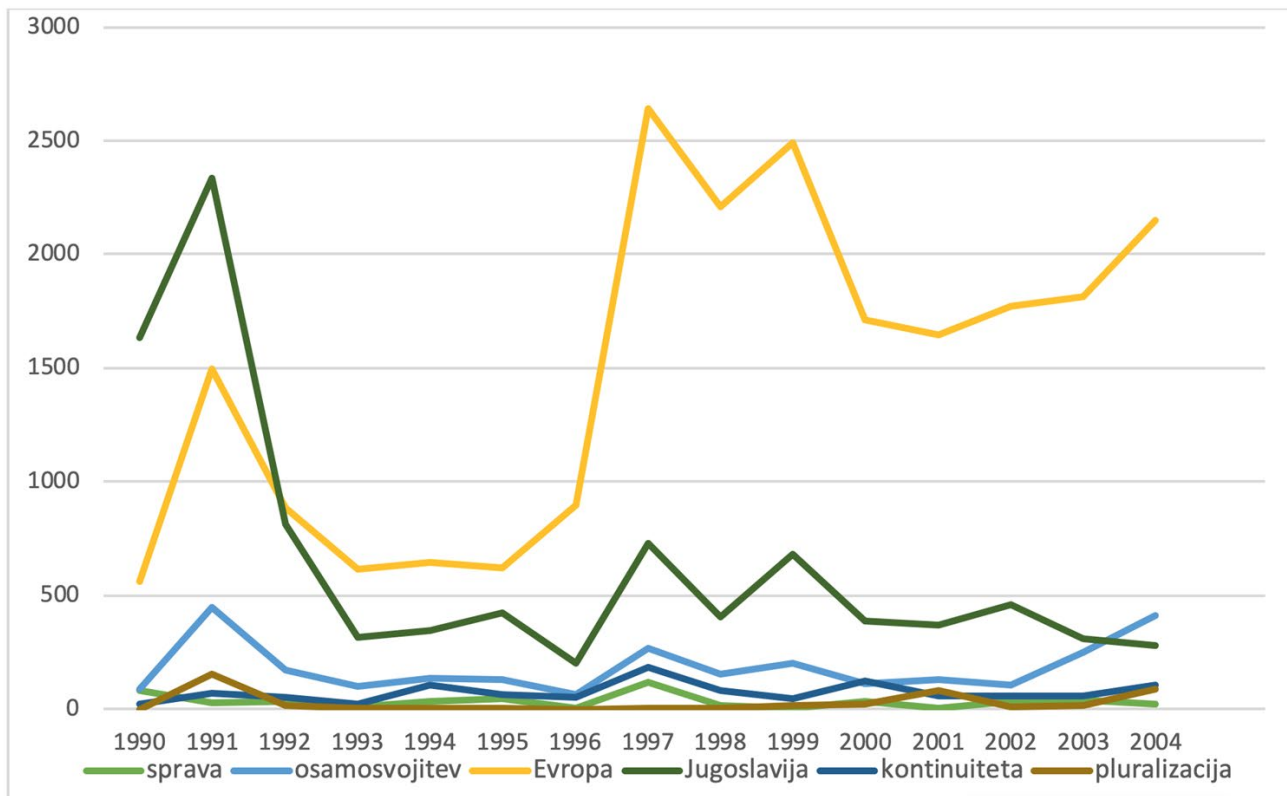
[3] Tabela 5: Učinkovito – kolokacijski kandidati^[30]

		<u>Cooccurrence count</u>	<u>Candidate count</u>	<u>logDice</u>
<u>p</u> <u>N</u>	rabo	436	3.414	9,551
<u>p</u> <u>N</u>	hitro	762	32.807	9,010
<u>p</u> <u>N</u>	delovanje	661	29.133	8,918
<u>p</u> <u>N</u>	upravljanje	395	13.813	8,780
<u>p</u> <u>N</u>	bolj	2.542	190.771	8,656
<u>p</u> <u>N</u>	izvajanje	531	28.582	8,620
<u>p</u> <u>N</u>	uspešno	363	16.469	8,533
<u>p</u> <u>N</u>	energije	328	14.149	8,495

Primer 2: Sprava

<https://sciendo.com/article/10.2478/jnmlp-2021-0006>

- sprava kot koncept (1990 - 2004)
- izbrana metodologija: konceptualna/politična zgodovina; digitalna humanistika
- postavljanje časovnice + historični kontekst
- distant reading - close reading

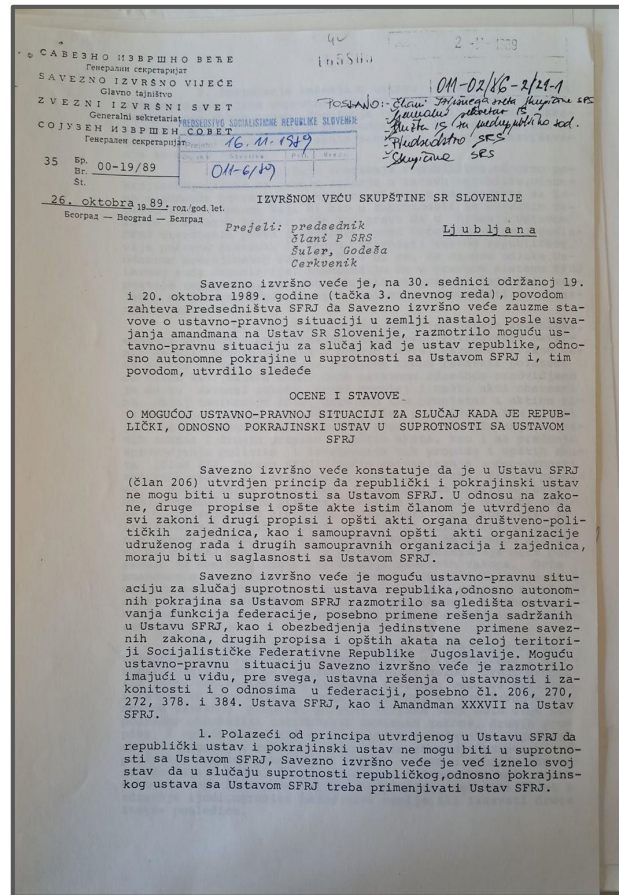
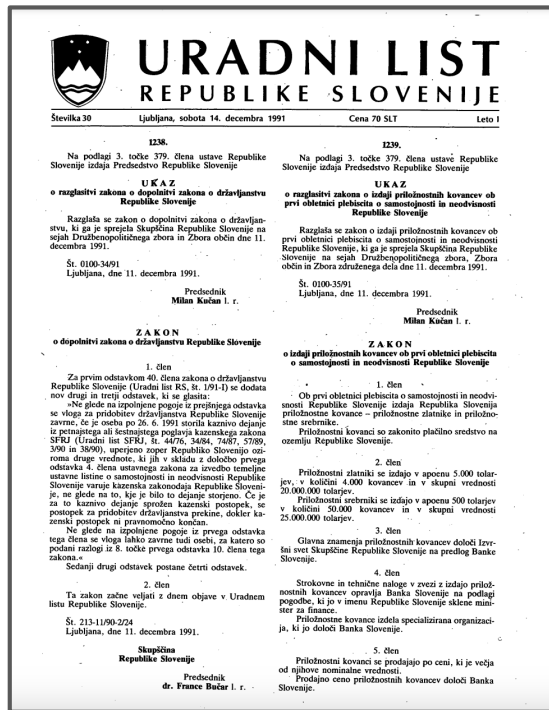


Graf 1: Sprava, osamosvojitev, Evropa, Jugoslavija (1990 – 2004).

Series "kontinuiteta" Legend Entry

Število omemb besed sprava, osamosvojitev, Evropa in Jugoslavija v skupščini oz. državnem zboru med letoma 1990 in 2004. Vir: <https://www.clarin.si/kontext/corpora/corplist> (7. 6. 2021.)

- kombiniranje z drugimi viri
(kvalitativna analiza)



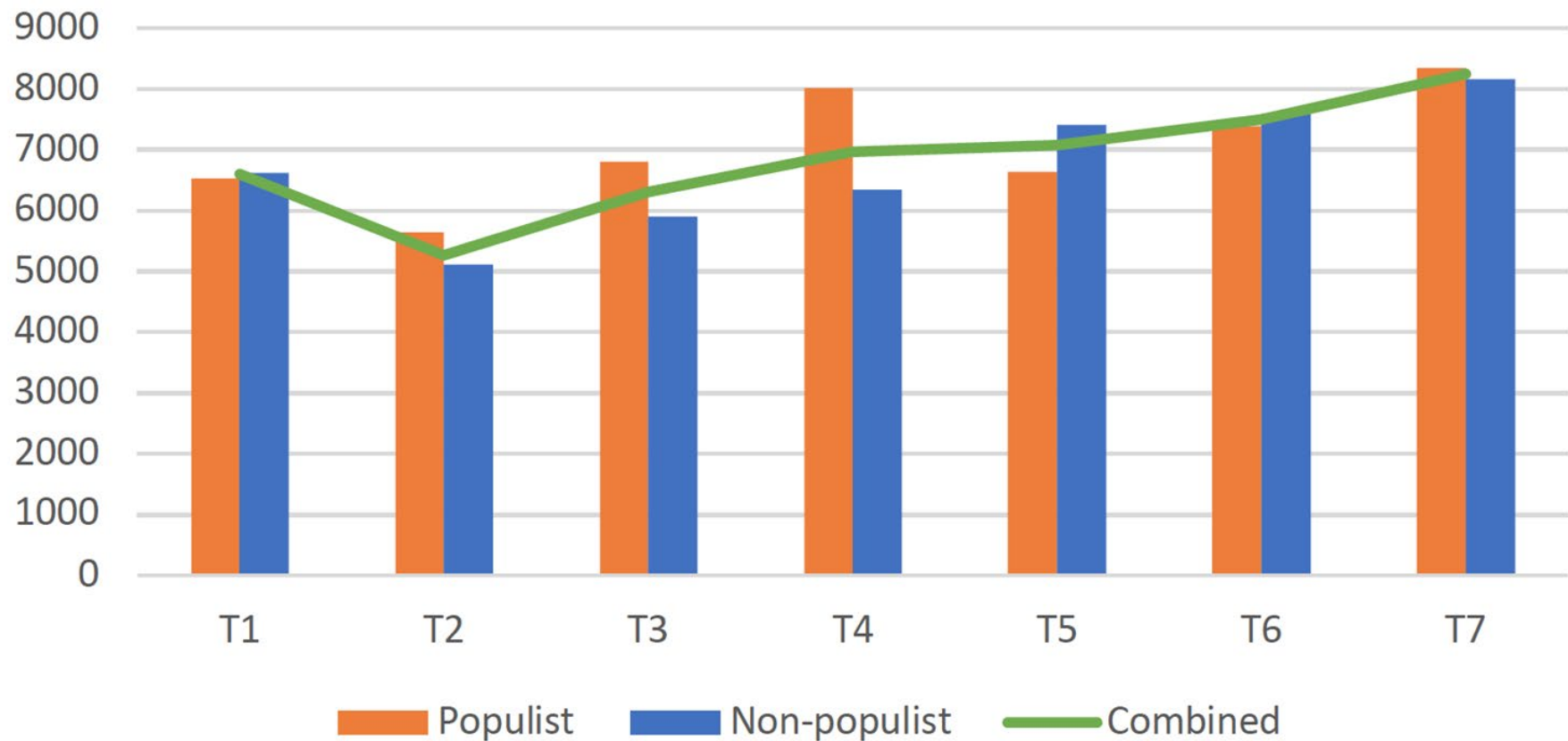
Primer 3: Populizem

<https://doi.org/10.4312/slo2.0.2023.1.69-90>

- problem: polarizacija političnega prostora, populizem
- vprašanje: populizem in naslavljanje javnosti
- izbrana metodologija: kulturna/politična zgodovina (*Lebenswelt*); digitalna humanistika
- analiza velike količine podatkov (kvantitativna analiza)

		POPULIST1-7		NON-POPULIST1-7		
	#tokens	49,070,504		77,215,381		
	#lemmas	76		74		
	LEMMA	AF	RF	AF	RF	P:N ratio
P-ONLY	Hrvat	1,341	27.33	0	0.00	/
	žena	397	8.09	0	0.00	/
	Avstrijec	318	6.48	0	0.00	/
	diplomant	300	6.11	0	0.00	/
	storilec	232	4.73	0	0.00	/
	volilec	161	3.28	0	0.00	/
	delojemalec	36	0.73	0	0.00	/
	neslovenec	31	0.63	0	0.00	/
	svojec	27	0.55	0	0.00	/
delavka	0	0.00	0	0.00	/	
N-ONLY	deležnik	0	0.00	1,784	23.10	/
	prejemnik	0	0.00	1,191	15.42	/
	najemnik	0	0.00	983	12.73	/
	dolžnik	0	0.00	752	9.74	/
	vajenec	0	0.00	444	5.75	/
	kadilec	0	0.00	290	3.76	/
krajan	0	0.00	172	2.23	/	
JOINT	oče	929	18.93	329	4.26	4.44
	obrtnik	1,187	24.19	540	6.99	3.46
	davkoplačevalec	4,762	97.04	2,178	28.21	3.44
	migrant	2,627	53.54	1,255	16.25	3.29
	vlagatelj	426	8.68	260	3.37	2.58
	podjetnik	3,880	79.07	2,671	34.59	2.29
	moški	827	16.85	619	8.02	2.10
	ljudstvo	3,089	62.95	2,376	30.77	2.05
	Italijan	272	5.54	216	2.80	1.98
	Slovenka	1,432	29.18	1,143	14.80	1.97
	pacient	1,619	32.99	1,452	18.80	1.75
	zamejstvo	1,067	21.74	966	12.51	1.74
	kmet	6,839	139.37	6,739	87.28	1.60
	prijatelj	1,024	20.87	1,012	13.11	1.59
	naročnik	517	10.54	516	6.68	1.58
	Slovenec	10,103	205.89	11,090	143.62	1.43
	dijak	2,403	48.97	2,670	34.58	1.42
	kupec	1,216	24.78	1,357	17.57	1.41
	državljan	21,570	439.57	24,828	321.54	1.37
	priča	4,061	82.76	4,701	60.88	1.36
	državljanka	6,902	140.65	8,372	108.42	1.30
	narod	4,952	100.92	6,035	78.16	1.29
	žrtev	3,945	80.39	4,810	62.29	1.29
	sosed	738	15.04	928	12.02	1.25
	človek	68,517	1,396.30	86,824	1,124.44	1.24

Rom	627	12.78	808	10.46	1.22
bolnik	1,279	26.06	1,717	22.24	1.17
prošilec	343	6.99	468	6.06	1.15
javnost	16,248	331.12	22,367	289.67	1.14
starš	5,732	116.81	7,893	102.22	1.14
oseba	16,836	343.10	23,762	307.74	1.11
subjekt	3,406	69.41	4,866	63.02	1.10
družina	11,120	226.61	16,298	211.07	1.07
otrok	18,205	371.00	26,762	346.59	1.07
gost	966	19.69	1,438	18.62	1.06
begunec	1,247	25.41	1,879	24.33	1.04
mladina	1,384	28.20	2,101	27.21	1.04
delničar	444	9.05	684	8.86	1.02
tujec	3,169	64.58	4,908	63.56	1.02
zavarovanec	896	18.26	1,394	18.05	1.01
volivec	3,478	70.88	5,544	71.80	0.99
lastnik	8,031	163.66	12,814	165.95	0.99
mati	320	6.52	512	6.63	0.98
družba	23,431	477.50	38,532	499.02	0.96
študent	4,973	101.34	8,202	106.22	0.95
posameznik	7,367	150.13	12,307	159.39	0.94
zavezanec	2,437	49.66	4,096	53.05	0.94
uporabnik	3,441	70.12	5,866	75.97	0.92
nosilec	2,211	45.06	3,812	49.37	0.91
občan	1,558	31.75	2,688	34.81	0.91
prebivalec	5,318	108.37	9,404	121.79	0.89
partner	4,580	93.34	8,312	107.65	0.87
potrošnik	1,657	33.77	3,060	39.63	0.85
generacija	2,279	46.44	4,215	54.59	0.85
delavec	10,768	219.44	20,055	259.73	0.84
invalid	3,032	61.79	5,760	74.60	0.83
prebivalstvo	2,727	55.57	5,452	70.61	0.79
manjšina	2,742	55.88	5,518	71.46	0.78
učenec	1,437	29.28	3,071	39.77	0.74
ženska	2,941	59.93	6,517	84.40	0.71
upokojenec	3,547	72.28	8,097	104.86	0.69
skupnost	16,208	330.30	38,163	494.24	0.67
priladnik	1,375	28.02	3,238	41.93	0.67
upravičenec	1,673	34.09	4,523	58.58	0.58
upnik	566	11.53	1,725	22.34	0.52
podpisnik	465	9.48	1,460	18.91	0.50
udeleženec	500	10.19	1,685	21.82	0.47
porabnik	129	2.63	540	6.99	0.38
populacija	480	9.78	2,179	28.22	0.35
Total	359,320	7,322.53	524,195	6,788.74	1.08



Primer 4: European Politics

Behind Closed Doors <https://ojs.inz.si/pnz/article/view/4174>

- problema: avtomatično in rutinsko sprejemanje regulativ in politik EU in umanjanje javnih debat; nezmožnost vlade, da bi o svojih odločitvah dosledno informirala zainteresirano javnost in svoje odločitve približala ljudem
- “democracy is threatened by politicians delivering more than they promise; decisions that were neither promised nor communicated by the government”

Figure 1:
Right-left
party position
based on rile
variable
(source:
MARPOR)

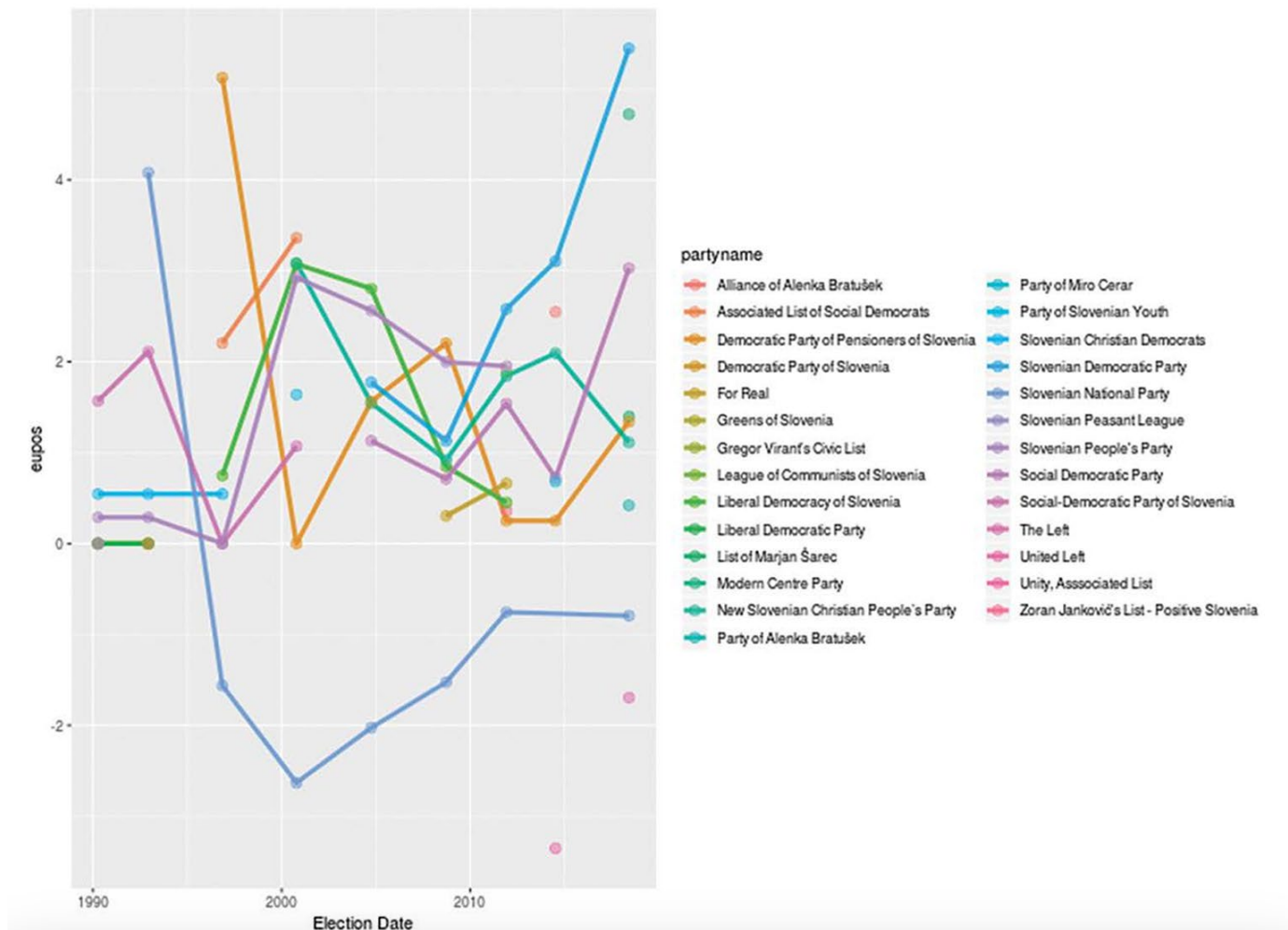


Table 11: Press releases about the EU legislation and other issues in 2003 and 2004

Agenda items	2003				2004			
	Releases		Words		Releases		Words	
	No.	Percentage	No.	Percentage	No.	Percentage	No.	Percentage
EU legislation	29	4.47%	11,267	5.18%	25	4.91%	10,954	6.61%
Other	620	95.53%	206,167	94.82%	484	95.09%	154,883	93.39%
Total	649		217,434		509		165,837	

- automatic and routine adoption of EU regulations without serious public debate
- Slovenia's accession to the EU is seen as inevitable and politicians would like to show that they are prepared to do whatever is necessary for integration
- the government failed to communicate relevant matters in a timely manner
- politicians are working on issues and actions that were not included in their manifestos and therefore constituents are simply not aware



institut_za_novejso_zgodovino



Inštitut za novejšo zgodovino



History on the Edge / Zgodovina na Špici



INŠTITUT
ZA NOVEJŠO ZGODOVINO

Tjaša Konovšek

tjasa.konovsek@inz.si

Katja Meden

katja.meden@inz.si