

## CODING RULES – CODES & FILTERS Working Notes

ADHOC PN-III-P1-1.1-TE-2016-0892 | Contract number: 35/2018

### Group Designators

Dimension	Ro	Hu lemma	Hu stem
Ethno Nat	tigan	cigány	cigány
	țigan		
	rrom	roma	roma
	evrei	zsidó	zsidó
	jidan	izraelita	izraelit*
	ungur		
	maghiar		
	bozgor		
	ardelean	erdélyi	erdélyi
	secui	székely	székely
		“határon túli”	“határon túli”
		román	román
		oláh	oláh
	Religion	islam	iszlám
		iszlamista	iszlam*
musulm		muzulmán	muzulm*
		muszlim	muszlim
credincio		igazhitű	igazhit*
atei		hitetlen	hitetlen
Sexual orientation	homosex	homoszexuális	homoszex*
	gay	gay	gay
	lesbi	leszbikus	leszbi*
	poponar	buzi	buzi
		buzeráns	buzer*
	LGBT	LGBT	LGBT
Social group	asistati	segélyezett	segély*
	asistați		segíj
	refugiat	menekült	menekült
	migrant	migráns	migrán*
		migráncs	migrán*
	imigrant	bevándorló	bevándor*
	pensionar	nyugdíjas	nyugdíjas
		nyugger	nyugger
	saraci		

ADHOC PN-III-P1-1.1-TE-2016-0892 | Contract number: 35/2018

## Codebook samples

Minorities	Demony m RO	Exogroup	Profession	Sexuality	Religion	Age	Economic status	Racial	Political affiliation	Historical	Migrants	Cultural	Criminal
țigani	moldoveni	englezi	cioban	homosexual	ortodocși	bătrâni	miliardar	negri	PSD-ști	daci	refugiat	maneliști	penali
țigani	oltenii	francezi	polițai	homo	preot	babe	cersetori	albi	socialiști	iobagi	imigranți	cocalari	pușcăriabili
evrei	mitici	albanezi	mineri	lesbiene	eretici	tineri	pensionar	asiatici	PNL-iști	militieni	emigranți	intelectuali	
secui	ardeleni	europeni		bulangii	sectant	moși		africani	globaliști	masoni	diaspora	analfabeți	
sasi	munteni	străini			crestini	copii			ecologiști	bastinași			
unguri	basarabeni				ateu	adulți				strămoșii			
bozgori	sudisti				satanisti					străbunii			
etnic					necredincios					cotropitorii			
minoritar					musulmani					valahi			
tătar					islamici					comuniști			
					catolici								

Imprecation /profanity	Insults																
	Animal	Abject	Intellectual	Diseases	Criminal	Sexual	Funerary	Hygienic	political	Subervient	Physical	Addiction	Status inf	Duplicity	Opportunism	Uncivilized	Incompetence
lua-te-ar	porc	Căcați	idiot	ciuma	penali	boarfe	zombi	gunoaie	sorosisti	lingăi	chior	bețiv	milogii	trădător	tupeist	ordinați	repetenți
lua-v-ar	măgar	bășinoși	tâmpit	cancer	ciorditori	perversi	cadavre	nespălat	bășiști	pupincuriști	bubos	drogat	sărănțoci	mincinos	șmecher	mârțolan	neisprăvit
morții mă-tii	jigodie	scârbe	imbecil	holera	excroci	zdreanță	hoituri	împutit	iohanști	limbiști	ciung	alcoolist	fomisți	fatarnici	parveniți	nerusnati	ratat
du-te dracului	limbric	pișaret	retarzi	răios	sarlatani	sugați	lepădătură	jeg	securiști	sclavi	olog	auroiaci	nenorociti			țărănoși	nepricepuți
plm	javră	borât	dobitor	lepra	pungaș	muiști	boșorogi	slinos	macoviști		orbi		pomanagii			golan	
muie	lipitori	scur suri			pușcăriabil	bula ngii	senili	paduc hiosi	bolșevici							primitivi	
	bestii	bali gă			infrac tori	cord itii			penel iști								
					tâlhar	găoz ari			politr uci								
						tîrfa											

## **KH Coder CODE SAMPLES for ROMANIAN**

### **\*Transylvanians**

Ardeal | ardelean | ardeleanca | ardeleni

### **\*Moldavians**

Moldova | molvodean | moldoveanca | moldoveni

### **\*Roma**

tigan | tigani | tigani | tiganeala | tiganie | tiganca | tiganul | tiganilor | rrom | rroma | rromi | rromilor

### **\*Hungarians**

maghiar | maghiara | maghiari | ungur | unguri | bozgor

### **\*Szekler**

secui | secuime | secuiesc | secuilor | secuii

### **\*Welfare\_recipients**

asistat | asistat+social | asistatii+social | asistati+social | asistati | asistatilor

### **\*LGBT**

homosexual | lesbiana | lesbi | poponar | gay | poponari | poponarul | homosecsi | lesbiene | gaylor | transsexual | transexuali

### **\*priests**

popa | preot | preoti | popi | episcop | calugar | patriarh | mitropolit

### **\*faithful**

credincios | credinciosi | crestin | crestinesc | crestineste | crestinilor | credinta

### **\*atheists**

atei | ateu | necredincios | necredinciosi | necredincio

### **\*Muslim**

musulman | musulmani | musulmana | islam | islamic

### **\*pensioner**

pensionar~ | pensionarii

### **\*Refugees**

refugiati | refugiat | refugia | refugiatilor

### **\*Migrants**

migrant | migranti | migrantilor | migratie

### **\*The\_Church**

Biserica | biserici | BOR | Biserica+Ortodoxa | Biserica+Ortodoxa+Romana | manastire | manastriri | catedrala | sfintit | moaste | mitropolit | parohie | slujba+religioasa | slujba+de+duminica

### **\*Religious\_Holidays**

Craciun | Paste | Pasti | Inviere | Iisus | Dumnezeu | Sf.+Ion | Boboteaza | Iordan | Sfantul+Ion | Sf.+Maria | Sfanta+Maria | Arhangheli | Rusalii | Sf.+Andrei | Sfantul+Andrei | Rusalii

### **\*Religion**

religie | religia | religios

### **\*Money**

bani | lei | euro | dolari | aur | argint | RON | franci

### **\*corporations**

corporatie | corporatist | corporatisti | call-center

### **\*Government**

guvern | guvernul | ministrul | ministru | premier | premierul | primul+ministru | minister | ministerul

### **\*sexuality**

heterosexual | sexy | sexual | sex | masturbare | porno | pornografie | pornografic | fund | dezbracat | penis | penisul | chiloti | testicule | vagin | impotentia | potentia | sani | excita | anus | anal

### **\*alcohol**

alcool | bautura+alcoolica | alcoolic | alcoolica | alcoolism | bautura | tuica | vodca | spirt | bere | beri | sprit | beat | beti | baut | rachiu | imbatat | betie | mahmur | mahmureala

### **\*theft**

furt | hotie | hot | jaf | jefuit | talhar | talharie | furat | furtisag | sustras | subtilizat | hoti

### **\*stupidity**

prostie | prostia | prost | cretin | incompetent | tembel | incompetenta | prostanac | prostovan | debil

\*laziness

lenes | lene | lenea | lenevit | nepasator | lenesi | indolent | merge+si+asa

\*violence

bataie | crima | criminal | agresor | omor | omorat | viol | violata | sange | injuraturi | pumni | impuscat | sinucidere | sinucigas | ucis | ucigas

\*education

scoala | scolar | scoli | profesor | profesoara | educator | educatoare | educatie | universitate | universitar | student | studentul | studenta | studentii | invatat

\*business

afacere | afaceri | fonduri | fond | investitie | investitii | angajati | angajatii | finantare | finantat | cumpara | vinde | cumparare | cumpararea | vanzarea

\*expertise

expert | expertul | experta | experti | specialist | specializat | specialista | specialisti | analist | analisti | cercetator | cercetatorul | cercetatori | cercetatorii | om+de+stiinta | omul+de+stiinta | critic | criticul | oamenii+de+stiinta | sociolog | consultant | consultanta | consultanti | profesorul+universitar

\*vulgar\_expressions

pula | pizda | cur | fut | futu-va | va+fut | bag+pula | in+pizda | futut | fututi

\*call\_to\_action

ar+trebui | sa+ii | trebuie | ii+vom | ar+fi+trebuit | trebuie+sa

\*group\_actions

omorati | bagati | dati+afara | lasati+sa+moara | expulzati | executati | inchisi | ucisi | batuti | gazati | atacati | fututi

\*elections

alegeri | alegeri+parlamentare | alegeri+prezidentiale | alegerile+parlamentare | alegerile+prezidentiale | electorat | exit-poll | biroul+electoral | liste+electorale | listele+electorale | vot | votanti | pomana+electorala | pomeni+electorale | turism+electoral

\*political parties

partid | partid+politic | alianta+politica | PSD | USL | PDL | PNL | PC | PMP | PNTCD | UNPR | PRM | UDMR

\*corruption

DNA | Directia+Nationala+Anticoruptie | coruptie | corupt | mita | trafic+de+influenta

\*justice

judicator | avocat | procuror | tribunal | judecata | judecatorie | magistrat | instanta | sentinta

\*pension

pensie | pensii~

\*EU

UE | Uniunea+Europeana | Parlamentul+European | Comisia+Europeana | Legislatia+Europeana

## KH Coder CODE SAMPLES for HUNGARIAN

### \*Transylvanians

erdély | legerdélyibb | határon+túli | erdély

### \*Moldavians

moldva | moldvá | moldvai

### \*Roma

cigány | cigányok | cigáni | clegcigányabb | roma | romáb | romácé | romácé | romához | romái | romáj | romák | romám

### \*Romanians

román | legrománabb | roman

### \*Szekler

székely | legszékelyebb

### \*Welfare\_recipients

segély | segíly | segíj

### \*LGBT

homoszexuális | buzi | gay | meleg | homokos | leszb | transzszexuális

### \*priests

pap | pópá | pópák | lelkesz | püspök | szerzetes | pátriárka | pátriárkák | mitropolita | mitropolitá | pápa | pápá

### \*faithful

hívó | keresztény | keresztyén | hit

### \*atheists

ateisták | ateista | hitetlen | hitetlen

### \*Muslim

muzulmán | muzulmán | muszlim | muszlim | iszlám | iszlám | iszlamista

### \*pension\_er

nyugdíjas | nyugdíj

### \*Refugees

menekült | bevándorló | bevándorlás | menekültek

### \*Migrants

migráns | migráció | migráncs | migran

### \*The\_Church

egyház | kolostor | szentel | ereklye | ereklyé | mitropolita | mitropolitá | püspök | egyházkerület | mise | misé |

istentisztelet | egyházi+szertartás

### \*Religious\_Holidays

karácsony | húsvét | pünkösöd | Jézus | Isten | vízkereszt | szent+István | jordán | Szűz+Mária | Szűz+Múriá | angyalok |

boldogasszony | nagyboldogasszony | kisboldogasszony

### \*Religion

vallás

### \*Money

pénz | forint | euro | euró | dollár | arany | aranny | ezüst | HUF | frank

### \*corporations

vállalat | nagyvállalat | multinacionális | multik | call-center

### \*Government

kormány | miniszter | kormányfő | miniszterelnök | államtitkár | kormányzat | minisztérium

### \*sexuality

heteró | heteroszexuális | szex | maszturbál | pornó | pornográf | segg | meztelen | farok | farka | bugyi | heré | vagina |

vaginá | impotencia | impotens | potencia | mellek | mellei | izgat | ánus | anális

### \*alcohol

alkohol | alkoholos+ital | alkoholista | alkoholist | alkoholizmus | ital | pálinka | pálinká | vodka | vodká | szesz | bor | sör

| spricc | részeg | likőr | ittas | részeg | másnapos

### \*theft

lop | rablás | tolvaj | rabló | lenyúl | elemel | betör | fosztogat

### \*stupidity

buta | buta | hülye | hülyé | inkompetens | gyagyás | ostoba | ostobá | síkhülye | agyatlan | féleszű | idióta | idióta | eszetlen | fafej | balfék | barom | barmok | ütődött

\*laziness

lusta | lustá | naplopó | semmirekellő | munkakerülő | tétlen | ingyenélő

\*violence

verekedés | bűnügy | büntett | bűnöző | tettes | gyilkos | erőszak | megerőszakol | vér | nemi+erőszak | mocskol | gyaláz | öklöz | ló | öngyilkos | megöl

\*education

iskola | iskolá | tanár | tanárnő | oktató | nevelő | óvónő | oktatás | egyetem | tanul | diák | hallgató

\*business

vállalkozás | üzlet | európai+uniós+alap | befektetési+alap | befektet | alkalmaz | alkalmazott | finanszíroz | vásárol | vásárl | elad | adós

\*expertise

szakértő | szakértel | elemző | kutató | tudós | közgazdász | szociológus | politológus | tanácsadó

\*political\_parties

párt | szövetség | pártszövetség | fidesz | mszp | lmp | momentum | liberális | kommunista | fasiszta | libsi | libbant | fidesznyik

\*EU

eu | európai+Unió | Európai+Bizottság | európai+normák

## WORD EMBEDDINGS SAMPLE - ROMANIAN

### ASISTATI\_SOCIAL

[('gubernanti', 0.9458024501800537), ('asistatii\_social', 0.924261748790741), ('bir', 0.9224841594696045), ('gubernantii', 0.9146674871444702), ('alesii', 0.9130131006240845), ('ganditiva', 0.9083521366119385), ('om\_rand', 0.9066683053970337), ('distrugeti', 0.9065779447555542), ('furati', 0.889508843421936), ('lacom', 0.8888063430786133), ('alesi', 0.8884375095367432), ('saracia', 0.8871846199035645), ('realizeaza', 0.8847759366035461), ('munciti', 0.8834285140037537), ('conducatori', 0.8830870389938354)]

### ASISTATI\_SOCIAL - BAD-GOOD

[('pedofili', 0.792786717414856), ('violator', 0.7819054126739502), ('nemernicie', 0.7786158919334412), ('escroc', 0.7656158208847046), ('pedepsit', 0.7578718662261963), ('disperat', 0.7522149085998535), ('pedepsi', 0.7470794916152954), ('confisca', 0.7462253570556641), ('sclavie', 0.7373212575912476), ('psihopat', 0.7366918921470642), ('lene', 0.7340413331985474), ('confiscare\_ave', 0.7338179349899292), ('complice', 0.7305622100830078), ('instala', 0.7294549345970154), ('nenoroci', 0.7293797731399536)]

### PENSIONAR

[('pensie', 0.9579870700836182), ('pensie\_special', 0.9135664701461792), ('ajutor\_social', 0.8971158266067505), ('pensie\_salariu', 0.8925050497055054), ('salariu', 0.8814470171928406), ('bugetar', 0.8611291646957397), ('marit', 0.8605667352676392), ('salariu\_pensie', 0.8590859174728394), ('marite', 0.8294861316680908), ('leu\_luna', 0.825614869594574), ('nesimtite', 0.8226847052574158), ('marirea', 0.8199978470802307), ('salariu\_minim', 0.8189727067947388), ('marita', 0.8174207806587219), ('pensiona', 0.8136670589447021)]

### PENSIONAR - BAD-GOOD

[('mari', 0.7498512864112854), ('marire', 0.7362866401672363), ('spaga', 0.7316147089004517), ('saracie', 0.7305564880371094), ('marita', 0.7285633087158203), ('asistat\_social', 0.7177252769470215), ('taiat', 0.7144455313682556), ('gubernant', 0.7063241004943848), ('gramada', 0.6969618201255798), ('astia', 0.6917175650596619), ('siau', 0.6877564191818237), ('inchisoare', 0.6827267408370972), ('adica', 0.6766573190689087), ('incasa', 0.6722469329833984), ('batrân', 0.6696475148200989)]

### POLITICIAN

[('incapabil', 0.8835226893424988), ('clasa\_politica', 0.8772834539413452), ('interes\_personal', 0.8442164659500122), ('capabil', 0.844005823135376), ('slujba', 0.8437508344650269), ('alesi', 0.8273704051971436), ('conducere', 0.8271175622940063), ('corupti', 0.8257226347923279), ('cinstiti', 0.8230580687522888), ('incompetenti', 0.8097464442253113), ('competenti', 0.8041564226150513), ('popor', 0.8032565712928772), ('favoare', 0.8026576638221741), ('pusi', 0.7999550104141235), ('dezbina', 0.7990962266921997)]

### POLITICIAN - BAD-GOOD

[('condamnat', 0.8079996109008789), ('ticălos', 0.7890251874923706), ('condus', 0.7843425273895264), ('corupt', 0.7716017961502075), ('tolera', 0.769381046295166), ('trădător', 0.7691595554351807), ('violenta', 0.7682304978370667), ('trădare', 0.7681998014450073), ('saracie', 0.7612701058387756), ('hoție', 0.7577304244041443), ('rusine', 0.7540237307548523), ('borfaș', 0.752455921554565), ('gașcă', 0.7475906610488892), ('hoț', 0.7466064691543579), ('escroc', 0.7416002750396729)]

### IMIGRANT

[('românești', 0.9203136563301086), ('urias', 0.9118832349777222), ('proportie', 0.9097946286201477), ('graniță', 0.9082638621330261), ('potențial', 0.9057322144508362), ('întreține', 0.893783688545227), ('vândut', 0.8859002590179443), ('emigrant', 0.8855129480361938), ('cădere', 0.8854451775550842), ('armat', 0.879485011100769), ('urgenta', 0.8767713308334351), ('concurență', 0.8764800429344177), ('supraviețui', 0.8746950030326843), ('țării', 0.8740130662918091), ('uniform', 0.8739757537841797)]

### IMIGRANT - BAD-GOOD

[('țară', 0.8172794580459595), ('încheia', 0.8150277733802795), ('violenta', 0.8123819828033447), ('bătut', 0.8067965507507324), ('greși', 0.8062039613723755), ('iarăși', 0.8049556612968445), ('supraviețui', 0.8036218881607056), ('graniță', 0.8027623891830444), ('nuși', 0.8022406101226807), ('pretenție', 0.8000437021255493), ('încoace', 0.7994111776351929), ('masa', 0.7975777387619019), ('saracie', 0.7971483469009399), ('târziu', 0.7964507937431335), ('stăpân', 0.7964484691619873)]

### REFUGIATI

[('vanzarea', 0.9491820335388184), ('multinationale', 0.9314088225364685), ('agentii', 0.9294372797012329), ('razboaie', 0.9281362295150757), ('canadian', 0.9208742380142212), ('produca', 0.918961763381958), ('strain', 0.9173069596290588), ('primitiv', 0.9140066504478455), ('bancilor', 0.9115343689918518), ('agenti', 0.906425416469574), ('vanduta', 0.9058310389518738), ('conditie', 0.9047649502754211), ('agresiune', 0.9038773775100708), ('britanic', 0.9035707116127014), ('energetic', 0.9028104543685913)]

### REFUGIATI - BAD-GOOD

[('armament', 0.7930762767791748), ('exterminare', 0.7902316451072693), ('sovietică', 0.7801975011825562), ('terrorist', 0.7768800258636475), ('atentat', 0.7739458084106445), ('tanc', 0.7685271501541138), ('fabrica', 0.7645453810691833), ('rusesc', 0.7579615116119385), ('african', 0.7557963132858276), ('rus', 0.751880407333374), ('abuza', 0.7454137802124023), ('viol', 0.7426174879074097), ('rachetă', 0.7425710558891296), ('hitler', 0.7424502372741699), ('chinez', 0.7418605089187622)]

### TÂNĂR

[('viitor', 0.8542914390563965), ('pasă', 0.8007388114929199), ('generatiile', 0.7978280782699585), ('traie', 0.7866458892822266), ('decent', 0.779647707939148), ('dle\_tariceanu', 0.7794591784477234), ('tineret', 0.7744258642196655), ('noștrii', 0.7654154300689697), ('bine', 0.7573164701461792), ('mentalitate', 0.756335973739624), ('educa', 0.7523830533027649), ('normalitate', 0.747164785861969), ('educati', 0.7467024326324463), ('lipsiti', 0.7406916618347168), ('rind', 0.7382382154464722)]

#### TÂNĂR - BAD-GOOD

[('sărăcie', 0.8389119505882263), ('șiau', 0.7775505185127258), ('păsa', 0.7661556601524353), ('plecat', 0.753673791885376), ('revoluție', 0.7521789073944092), ('nație', 0.7462567090988159), ('căpăta', 0.7450152635574341), ('bătrân', 0.744123637676239), ('manipulat', 0.7416713237762451), ('nimeni', 0.737675666809082), ('nesimțire', 0.7374688982963562), ('violență', 0.7322877049446106), ('as', 0.7284497022628784), ('bătu', 0.7243400812149048), ('democrație', 0.7240249514579773)]

#### BĂTRÂN

[('sărac', 0.9150033593177795), ('bătu', 0.9045209884643555), ('sărăcie', 0.8913739919662476), ('strămoș', 0.8842618465423584), ('înghiți', 0.8815957903862), ('pământ', 0.8732643127441406), ('nepot', 0.8695558905601501), ('pune\_mână', 0.8640463352203369), ('șiau', 0.862718939781189), ('noștrii', 0.8578217625617981), ('sărman', 0.8578143119812012), ('sânt', 0.8574120998382568), ('hrăni', 0.8558924794197083), ('plecat', 0.8540191650390625), ('casa', 0.8502804636955261)]

#### BĂTRÂN - BAD-GOOD

[('ceaușescu', 0.8316394090652466), ('bătu', 0.8229302167892456), ('încet', 0.8191518783569336), ('sărăcie', 0.8178056478500366), ('înjura', 0.8148131370544434), ('sânge', 0.809837818145752), ('târziu', 0.8083467483520508), ('pune\_mână', 0.7997657656669617), ('nuși', 0.7959142923355103), ('stăpân', 0.7949123978614807), ('ăla', 0.7943681478500366), ('pământ', 0.7935063242912292), ('nesimțire', 0.7884999513626099), ('gât', 0.7869915962219238), ('greși', 0.7861291170120239)]

#### MULTINATIONALE

[('bancilor', 0.9489991664886475), ('multinationalelor', 0.9488940238952637), ('produca', 0.9383494853973389), ('investeasca', 0.9349924921989441), ('refugiatii', 0.9314088225364685), ('vanzarea', 0.9290977716445923), ('bir', 0.9233738780021667), ('vanduta', 0.9229246973991394), ('forta\_muncă', 0.9147508144378662), ('controleaza', 0.9141678214073181), ('guvernantii', 0.9126054048538208), ('functionarii', 0.9122922420501709), ('vanda', 0.9120272994041443), ('investiti', 0.9112946391105652), ('implicat', 0.9105358123779297)]

#### MULTINATIONALE - BAD-GOOD

[('sovietică', 0.7870234251022339), ('armament', 0.7859616875648499), ('fabrica', 0.7809151411056519), ('legaliza', 0.7760068774223328), ('exterminare', 0.7748576998710632), ('abuza', 0.7700355052947998), ('viol', 0.7683317065238953), ('instala', 0.7675287127494812), ('servi', 0.7617487907409668), ('pedepsit', 0.7590837478637695), ('scos', 0.7560423612594604), ('mercenar', 0.7555158138275146), ('aplica\_lege', 0.7552838325500488), ('psihopat', 0.7544395923614502), ('rusesc', 0.7538145780563354)]

#### OM\_RAND

[('obligati', 0.9170184135437012), ('realizeaza', 0.9141116738319397), ('guvernantii', 0.9136497974395752), ('mare\_majoritate', 0.9100871086120605), ('inchisi', 0.9072742462158203), ('smecherii', 0.9068117141723633), ('asistati\_social', 0.9066684246063232), ('pusi', 0.9064017534255981), ('bir', 0.9052985310554504), ('asistatii\_social', 0.9023016691207886), ('conducatorii', 0.9001432657241821), ('mintiti', 0.8974342942237854), ('alesii', 0.8974063992500305), ('munceasca', 0.8973536491394043), ('bataie\_juca', 0.8957957029342651)]

#### OM\_RAND - BAD-GOOD

[('pedofili', 0.7709088325500488), ('pedepsit', 0.764062762260437), ('psihopat', 0.7546213865280151), ('pedepsi', 0.7411772012710571), ('confisca', 0.7380597591400146), ('instala', 0.7368574142456055), ('tentativă', 0.7330840826034546), ('lene', 0.7300658226013184), ('nemernicie', 0.7266241312026978), ('inuman', 0.7265316247940063), ('disperat', 0.7249712347984314), ('violator', 0.7246606945991516), ('pumn\_gură', 0.7246478796005249), ('iresponsabil', 0.7207268476486206), ('abuza', 0.719113826751709)]

#### MARE\_MAJORITATE

[('realizeaza', 0.9108270406723022), ('om\_rand', 0.910086989402771), ('obligati', 0.8754496574401855), ('inconstienti', 0.875029444694519), ('iluzie', 0.872748613357544), ('pretentii', 0.8701125383377075), ('indolent', 0.8654606342315674), ('afecteaza', 0.8640365600585938), ('nevoiti', 0.8639065027236938), ('cetateanul', 0.8632792830467224), ('lene', 0.8611098527908325), ('caresi', 0.8605377674102783), ('ganditiva', 0.8603001832962036), ('preocupa', 0.8597087860107422), ('smecherii', 0.857853353023529)]

#### MARE\_MAJORITATE - BAD-GOOD

[('pedepsit', 0.7650017738342285), ('violență', 0.7620230317115784), ('sărăcie', 0.7581123113632202), ('abuza', 0.7535328269004822), ('lene', 0.7534250020980835), ('nau', 0.7481329441070557), ('pedepsi', 0.7468647956848145), ('iresponsabil', 0.7422599792480469), ('tolera', 0.7396954298019409), ('inuman', 0.7351740598678589), ('supraviețui', 0.7326334714889526), ('instala', 0.7300089597702026), ('oare', 0.7279176712036133), ('pedofili', 0.7255831956863403), ('tentativă', 0.7254626750946045)]

#### POPOR\_ROMAN

[('luptă\_impotriva', 0.8713907599449158), ('neam\_romanesc', 0.8668399453163147), ('urmasii', 0.862590491771698), ('bataia\_juca', 0.8538187742233276), ('condusa', 0.8481744527816772), ('clasă\_politică', 0.8417403101921082), ('hotiei',

0.8325186967849731), ('conducatori', 0.8304949998855591), ('tradat', 0.8270711898803711), ('cinstiti', 0.8218766450881958), ('comunisti', 0.8199633359909058), ('dusmanii', 0.8174299597740173), ('acesti\_om', 0.8168131113052368), ('coruptiei', 0.8081251382827759), ('intimpla', 0.8071325421333313)]

#### POPOR\_ROMAN - BAD-GOOD

[('dictator', 0.7651103734970093), ('pedepsi', 0.7648961544036865), ('ticălos', 0.7619315385818481), ('dictatură', 0.7565089464187622), ('condamnat', 0.7474435567855835), ('tolera', 0.7469678521156311), ('lovitură\_sta', 0.743923544883728), ('criminal', 0.7418686151504517), ('violență', 0.7418291568756104), ('comunist', 0.7286434769630432), ('instiga', 0.7242999076843262), ('terorist', 0.7209417223930359), ('tâlhar', 0.709347128868103), ('pedofili', 0.7079564332962036), ('condus', 0.7057318687438965)]

#### CETATEANUL

[('ganditiva', 0.942978024482727), ('realizeaza', 0.9424018859863281), ('obligati', 0.940012514591217), ('caresi', 0.9178447723388672), ('pusa', 0.9170831441879272), ('afecteaza', 0.9169032573699951), ('aflati', 0.9161624312400818), ('cetateanului', 0.9123848676681519), ('politia\_roman', 0.9121848940849304), ('luata', 0.9109233617782593), ('rand', 0.9101993441581726), ('solutia', 0.9093015789985657), ('interesati', 0.9059988260269165), ('inafara', 0.9050782322883606), ('autoritatile', 0.9030358195304871)]

#### CETATEANUL - BAD-GOOD

[('tentativă', 0.8036259412765503), ('instala', 0.791975736618042), ('incita', 0.7745465040206909), ('viol', 0.7736126184463501), ('pedepsit', 0.7719885110855103), ('aplica\_lege', 0.771672248840332), ('confisca', 0.7699155807495117), ('pedofili', 0.7687695622444153), ('psihopat', 0.7681449055671692), ('iresponsabil', 0.7593169212341309), ('abuza', 0.7551000118255615), ('comite', 0.7517954707145691), ('diversiune', 0.7497290372848511), ('servi', 0.7494356632232666), ('paralelă', 0.7461020350456238)]

#### INSTITUTIILE\_STAT

[('condamnati', 0.9534891843795776), ('protejeaza', 0.9344924688339233), ('securistilor', 0.9337416887283325), ('coruptilor', 0.9311769008636475), ('pusi', 0.9307410717010498), ('complicitate', 0.9275962710380554), ('consecintele', 0.9261168241500854), ('hotie', 0.923491358757019), ('hotiei', 0.9234906435012817), ('puscarii', 0.9233237504959106), ('smecherii', 0.9208527207374573), ('trasi', 0.9204639792442322), ('judecatori', 0.9190720915794373), ('jaf', 0.9190013408660889), ('organizata', 0.9184091091156006)]

#### INSTITUTIILE\_STAT - BAD-GOOD

[('pedepsi', 0.7981997132301331), ('lovitură\_sta', 0.7937189340591431), ('condamnat', 0.7923922538757324), ('crimă', 0.7917894124984741), ('pedofili', 0.7889018058776855), ('complice', 0.7856364250183105), ('psihopat', 0.7764703035354614), ('pedepsit', 0.775912880897522), ('terorist', 0.7696206569671631), ('violență', 0.7667388916015625), ('dictator', 0.7636623382568359), ('tentativă', 0.7618290781974792), ('condamnare', 0.7593826055526733), ('disperat', 0.7560453414916992), ('comite', 0.754707932472229)]

#### POPULATIA

[('populatiei', 0.9303591251373291), ('cetateni', 0.9290416240692139), ('interesati', 0.9165771007537842), ('populatie', 0.9153863191604614), ('alesi', 0.9100358486175537), ('realizeaza', 0.9018360376358032), ('milion\_roman', 0.897861897945404), ('cetateanului', 0.8943526744842529), ('societatii', 0.889641523361206), ('cetateanul', 0.8894945979118347), ('devina', 0.8874274492263794), ('intro\_tară', 0.8859214782714844), ('distruga', 0.8856103420257568), ('multinationale', 0.8852055668830872), ('cetatenii', 0.8823457360267639)]

#### POPULATIA - BAD-GOOD

[('crimă', 0.7454257607460022), ('dictatură', 0.7383333444595337), ('lovitură\_sta', 0.7344359755516052), ('instala', 0.7265175580978394), ('legaliza', 0.7193338871002197), ('instaura', 0.71802818775177), ('regim', 0.7173874378204346), ('diversiune', 0.7155460119247437), ('sovietică', 0.7154717445373535), ('genocid', 0.7124373912811279), ('pedepsit', 0.7114064693450928), ('confisca', 0.7088450193405151), ('furt', 0.7036449909210205), ('pedepsi', 0.7030752897262573), ('victimă', 0.7022203207015991)]

## WORD EMBEDDINGS SAMPLE - HUNGARIAN

### SEGÉLY

[('befizet', 0.9964795708656311), ('termel', 0.9933181405067444), ('roma', 0.9916682243347168), ('minimál', 0.991542398929596), ('nos', 0.9910553693771362), ('sehol', 0.9907693862915039), ('pénzéből', 0.989710807800293), ('multik', 0.9895225763320923), ('üzlet', 0.9895050525665283), ('munkahely', 0.9885160326957703), ('nyugdíjat', 0.98847496509552), ('tudta', 0.9883910417556763), ('kezel', 0.9883216619491577), ('simán', 0.9880029559135437), ('gond', 0.9878680109977722)]

### SEGÉLY - BAD-GOOD

[('pl', 0.979894757270813), ('fizet', 0.9766671657562256), ('stb', 0.975071370601654), ('kifizet', 0.9683970212936401), ('szint', 0.9649806022644043), ('felvesz', 0.960861325263977), ('vesz', 0.9586732387542725), ('alacsony', 0.9518368244171143), ('minimum', 0.9439988136291504), ('ár', 0.9429068565368652), ('haverok', 0.9423019289970398), ('plusz', 0.9398895502090454), ('keres', 0.9366135597229004), ('amennyi', 0.9364519119262695), ('számol', 0.9360474348068237)]

### NYUGDÍJAS

[('jár', 0.9827567338943481), ('tanár', 0.976457953453064), ('ér', 0.9744547605514526), ('külföld', 0.9735674858093262), ('munkás', 0.970892608165741), ('rokkant', 0.9706131815910339), ('több', 0.9703662395477295), ('föld', 0.9700229167938232), ('mégis', 0.9681249260902405), ('idős', 0.9680908918380737), ('magan', 0.9675830602645874), ('ért', 0.967445969581604), ('lehet', 0.9662864208221436), ('másik', 0.9660863280296326), ('normális', 0.966044545173645)]

### NYUGDÍJAS - BAD-GOOD

[('fizetés', 0.9773486852645874), ('pénz', 0.9769889712333679), ('adó', 0.9717791676521301), ('számol', 0.9713433980941772), ('bér', 0.971062183380127), ('alacsony', 0.9695886373519897), ('nyugdíj', 0.9674087762832642), ('vesz', 0.9641155004501343), ('kifizet', 0.9633578062057495), ('keres', 0.9623952507972717), ('emel', 0.9622533321380615), ('magas', 0.959191620349884), ('millió', 0.9577244520187378), ('kerül', 0.9575941562652588), ('emelés', 0.9574719071388245)]

### POLITIKUS

[('vált', 0.9728645086288452), ('negatív', 0.9669613838195801), ('esély', 0.965411901473999), ('fideszt', 0.963817834854126), ('pár', 0.962005615234375), ('megélhetési', 0.961502730846405), ('baloldali', 0.9600820541381836), ('megért', 0.959047794342041), ('látzik', 0.9577229022979736), ('pláne', 0.9566382765769958), ('szól', 0.9566155076026917), ('év\_múlva', 0.9557974934577942), ('jobbiknak', 0.9557254314422607), ('érthető', 0.9553513526916504), ('értelmes', 0.9547390937805176)]

### POLITIKUS - BAD-GOOD

[('zseb', 0.9319007396697998), ('számol', 0.9311119318008423), ('hibás', 0.9283143281936646), ('vesz', 0.9272298812866211), ('sincs', 0.9267213940620422), ('válaszol', 0.9250159859657288), ('szint', 0.9237338304519653), ('hazugság', 0.922394871711731), ('túlóra', 0.9215600490570068), ('ukrán', 0.9208651185035706), ('kifizet', 0.9208410382270813), ('munkaerő', 0.9201341867446899), ('hazudik', 0.9199192523956299), ('os', 0.9191167950630188), ('azonnal', 0.9186109900474548)]

### MIGRÁNS

[('menekül', 0.9896173477172852), ('brüsszel', 0.9845329523086548), ('nevez', 0.9829716682434082), ('ukrán', 0.981995701789856), ('különbség', 0.9810731410980225), ('olcsó', 0.9768914580345154), ('migránsok', 0.9763962626457214), ('felesleges', 0.9758162498474121), ('ellenes', 0.9756962060928345), ('maffia', 0.9749469757080078), ('francia', 0.9739416837692261), ('nos', 0.9738143682479858), ('osztrák', 0.973713755607605), ('támogat', 0.973220705986023), ('mielőtt', 0.9726724028587341)]

### MIGRÁNS - BAD-GOOD

[('pl', 0.9590458273887634), ('stb', 0.9490034580230713), ('fizet', 0.9402649998664856), ('szint', 0.9353664517402649), ('német', 0.934739351272583), ('kifizet', 0.9252980947494507), ('vesz', 0.9237666130065918), ('amennyi', 0.9188946485519409), ('menekül', 0.9188041687011719), ('nyugati', 0.9143260717391968), ('felvesz', 0.9094538688659668), ('alacsony', 0.9090906977653503), ('fegyver', 0.9090579748153687), ('cég', 0.9066920280456543), ('propaganda', 0.9049376249313354)]

### MENEKÜLT

[('kártérítés', 0.9863619804382324), ('kvóta', 0.9845219850540161), ('szervez', 0.9844057559967041), ('irányít', 0.9838656187057495), ('euból', 0.9835662841796875), ('módszer', 0.9830222725868225), ('muszlim', 0.9826362133026123), ('kormányzati', 0.98258376121521), ('folyamat', 0.9823716878890991), ('részlet', 0.9821391701698303), ('jogos', 0.9817702770233154), ('megjegyzem', 0.9815737009048462), ('terrorista', 0.9815372228622437), ('migráció', 0.981421172618866), ('egységes', 0.981016993522644)]

### MENEKÜLT - BAD-GOOD

[('pl', 0.9544789791107178), ('fegyver', 0.9531878232955933), ('letelepedési', 0.9449105858802795), ('eus', 0.9418460130691528), ('német', 0.9402451515197754), ('amennyi', 0.9359430074691772), ('költés', 0.9358898401260376), ('alkalmaz', 0.9351613521575928), ('stb', 0.9348182678222656), ('orosz', 0.9321978092193604), ('hazai', 0.9310862421989441), ('cég', 0.9277596473693848), ('rendszer', 0.9276227355003357), ('azaz', 0.9272397756576538), ('deviza', 0.9264777898788452)]

### FIATAL

[('amily', 0.9873179197311401), ('felno', 0.9867473244667053), ('innen', 0.9851793646812439), ('borzasztó', 0.9824458360671997), ('sajnál', 0.9803687334060669), ('szomorú', 0.9803208112716675), ('jó', 0.9795825481414795), ('☹️', 0.9791819453239441), ('sir', 0.9791580438613892), ('hamar', 0.9764140844345093), ('csoda', 0.9762790203094482), ('igaz', 0.975926399230957), ('hihetetlen', 0.9753219485282898), ('nézem', 0.9750592112541199), ('tisza', 0.9748178720474243)]

### FIATAL - BAD-GOOD

[('ennyi', 0.9902950525283813), ('komoly', 0.986472487449646), ('annyi', 0.9847056865692139), ('senki', 0.9841752648353577), ('semmi', 0.9840937852859497), ('agy', 0.9815686941146851), ('tud', 0.9814083576202393), ('lop', 0.980213463306427), ('ostoba', 0.9797435402870178), ('dolog', 0.9787287712097168), ('ugyan', 0.9776083827018738), ('hülye', 0.9771636128425598), ('veszik', 0.9768283367156982), ('mennyi', 0.9767385721206665), ('☹️ ☹️ ☹️'), 0.9767345786094666]]

#### IDŐS

[('szül', 0.988369345664978), ('másik', 0.9881293773651123), ('szegény', 0.9867796301841736), ('könnyű', 0.9843586683273315), ('rengeteg', 0.984015703201294), ('ér', 0.9827845096588135), ('jár', 0.9827306866645813), ('nő', 0.9817833304405212), ('normális', 0.9813567399978638), ('szegénység', 0.9810811281204224), ('tanár', 0.9799216985702515), ('máshol', 0.9798823595046997), ('belől', 0.9782147407531738), ('élt', 0.9777272343635559), ('☹️'), 0.9772478938102722]]

#### IDŐS - BAD-GOOD

[('keres', 0.9740265607833862), ('vesz', 0.9706095457077026), ('kifizet', 0.9681114554405212), ('munkahely', 0.9663801193237305), ('szint', 0.9652548432350159), ('pénz', 0.9630608558654785), ('kap', 0.9619745016098022), ('lop', 0.9601537585258484), ('kell', 0.9588170051574707), ('fizetés', 0.9572212100028992), ('fél', 0.9561841487884521), ('olcsó', 0.9561098217964172), ('számol', 0.9560365676879883), ('stb', 0.9537966251373291), ('alacsony', 0.9535227417945862)]

#### MULTIK

[('fenntart', 0.9948912858963013), ('sehol', 0.9938469529151917), ('rendez', 0.9925217628479004), ('csere', 0.9924191832542419), ('kezel', 0.9921736121177673), ('segély', 0.9895225763320923), ('közmunka', 0.9892310500144958), ('befizet', 0.9891842603683472), ('osztrák', 0.989115297794342), ('átlagos', 0.988545835018158), ('nos', 0.9882597923278809), ('tudta', 0.9880518317222595), ('ukrán', 0.987721860408783), ('teher', 0.9875139594078064), ('adófizetők', 0.987504243850708)]

#### MULTIK - BAD-GOOD

[('pl', 0.9793303608894348), ('stb', 0.9682095050811768), ('fizet', 0.959976315498352), ('szint', 0.9592046737670898), ('kifizet', 0.9584957361221313), ('felvesz', 0.9545263648033142), ('vesz', 0.9457994103431702), ('amennyi', 0.943178653717041), ('alacsony', 0.942719042301178), ('haverok', 0.9411132335662842), ('vagyon', 0.9365120530128479), ('fegyver', 0.936214029788971), ('plusz', 0.9337010383605957), ('orosz', 0.9335525035858154), ('nyugati', 0.933249831199646)]

#### KISEMBER

[('fejlik', 0.9976538419723511), ('gyilkol', 0.9963574409484863), ('logika', 0.9963330030441284), ('kötél', 0.9962769746780396), ('dögöl', 0.995693027973175), ('bárkit', 0.9955500960350037), ('értet', 0.9954558610916138), ('rettenetes', 0.9953542947769165), ('lövét', 0.9952352046966553), ('végülis', 0.9950718283653259), ('szenny', 0.9949747323989868), ('mindég', 0.9949442744255066), ('közül', 0.9948891401290894), ('zsiros', 0.9948003888130188), ('utánna', 0.9947935342788696)]

#### KISEMBER - BAD-GOOD

[('onnan', 0.9665862917900085), ('zavar', 0.9664359092712402), ('imf', 0.9651199579238892), ('tudta', 0.9650737047195435), ('mell', 0.9644185304641724), ('csatorna', 0.9641637802124023), ('szint', 0.9635823965072632), ('fél', 0.9625597596168518), ('kamu', 0.9619926810264587), ('hol', 0.9615140557289124), ('melyik', 0.9603248238563538), ('nulla', 0.9599454402923584), ('mind', 0.9595141410827637), ('nyugat', 0.9573478698730469), ('multik', 0.9571181535720825)]

#### TÖBBSÉG

[('ellenék', 0.9904894828796387), ('képes', 0.9889593124389648), ('tisztességes', 0.987847089767456), ('politika', 0.987639844417572), ('választó', 0.9853896498680115), ('vált', 0.9835695028305054), ('többi', 0.982427716255188), ('vélemény', 0.9816241264343262), ('képvisel', 0.9812849760055542), ('párt', 0.9807696342468262), ('létez', 0.9797870516777039), ('ország', 0.9792999029159546), ('lényeg', 0.9783433079719543), ('valóság', 0.9763423800468445), ('hazugság', 0.9763413071632385)]

#### TÖBBSÉG - BAD-GOOD

[('köt', 0.9526351690292358), ('számol', 0.9516737461090088), ('uniós', 0.9504960775375366), ('gazdaság', 0.9495774507522583), ('mindez', 0.9486262798309326), ('nyugati', 0.9477055072784424), ('kerül', 0.947591245174408), ('lopás', 0.9471541047096252), ('használ', 0.9463500380516052), ('nyilván', 0.9460496306419373), ('menekül', 0.9459968209266663), ('egyenlő', 0.9452779293060303), ('eleve', 0.9444825053215027), ('támogat', 0.9443178176879883), ('illegális', 0.9443110227584839)]

#### MAGYAR\_NÉP

[('képvisel', 0.9910854697227478), ('való', 0.9862314462661743), ('választó', 0.9857317209243774), ('elérlik', 0.985166688919067), ('létez', 0.9831606149673462), ('kormányváltás', 0.9823354482650757), ('erős', 0.9821696281433105), ('szükség', 0.9811556339263916), ('józan\_ész', 0.9802818298339844), ('ellenzéki', 0.9800633788108826), ('világ', 0.9775245189666748), ('nép', 0.9775204062461853), ('politika', 0.9772107601165771), ('független', 0.9766783714294434), ('kampány', 0.9758689403533936)]

#### MAGYAR\_NÉP - BAD-GOOD

[('hatalom', 0.9756602048873901), ('bűnöző', 0.9753808975219727), ('nyugati', 0.9736859202384949), ('illegális', 0.9735440015792847), ('mindez', 0.9691349267959595), ('helyzet', 0.9679440259933472), ('migránsok', 0.9676874876022339), ('köt', 0.9675809144973755), ('jelenlegi', 0.9674315452575684), ('uniós', 0.9672225713729858), ('eleve', 0.9669259786605835), ('menekül', 0.9661960601806641), ('média', 0.9648346900939941), ('lopás', 0.9644514322280884), ('amennyi', 0.964382529258728)]

#### ÁLLAMPOLGÁR

[('lét', 0.9893238544464111), ('alap', 0.9889664649963379), ('szabály', 0.9871371984481812), ('szempont', 0.9863201379776001), ('működő', 0.9861294031143188), ('alkotmányos', 0.9858375787734985), ('biztosít', 0.9857165813446045), ('bírótság', 0.9855266213417053), ('korábbi', 0.9854303002357483), ('gazdasági', 0.984413743019104), ('kötelező', 0.9843853116035461), ('csalás', 0.9842357635498047), ('helyez', 0.9840295910835266), ('amely', 0.9831374883651733), ('rendszer', 0.9830665588378906)]

#### ÁLLAMPOLGÁR - BAD-GOOD

[('eset', 0.9562257528305054), ('szerződés', 0.9542709589004517), ('deviza', 0.9500458240509033), ('rendszer', 0.9463747143745422), ('alkalmaz', 0.938920259475708), ('lakosság', 0.9375226497650146), ('jelenleg', 0.9355698823928833), ('árfolyam', 0.9344792366027832), ('költség', 0.9334183931350708), ('azaz', 0.9328534603118896), ('alap', 0.9328299760818481), ('kötelező', 0.9296565055847168), ('szabály', 0.9294371604919434), ('cég', 0.9261809587478638), ('bírótság', 0.9241361618041992)]

#### ÁLLAMI

[('ingatlan', 0.9689355492591858), ('szerződés', 0.9615383148193359), ('ellátás', 0.9564598798751831), ('alkalmaz', 0.9503467082977295), ('ügyészség', 0.9500595331192017), ('feltétel', 0.949870228767395), ('helyi', 0.9470981359481812), ('intézmény', 0.9463422894477844), ('jelentős', 0.9453338384628296), ('terület', 0.9452798366546631), ('deviza', 0.9436874985694885), ('árfolyam', 0.9436051845550537), ('rendőrség', 0.9434266090393066), ('rendelkezik', 0.942379355430603), ('pénzügyi', 0.940453052520752)]

#### ÁLLAMI - BAD-GOOD

[('szerződés', 0.9274961948394775), ('költség', 0.9257142543792725), ('árfolyam', 0.9184534549713135), ('hitel', 0.9133884310722351), ('alkalmaz', 0.912676990032196), ('eset', 0.9096122980117798), ('ingatlan', 0.9075223803520203), ('deviza', 0.9038982391357422), ('összeg', 0.894797682762146), ('cég', 0.89432692527771), ('eus', 0.8901242017745972), ('mnb', 0.8796185255050659), ('ellátás', 0.8771591782569885), ('feltétel', 0.8727456331253052), ('lakosság', 0.8722596764564514)]

#### LAKOSSÁG

[('alap', 0.986609935760498), ('gazdasági', 0.9865520000457764), ('szabály', 0.9855003952980042), ('jelentős', 0.9850366115570068), ('feltétel', 0.9829990863800049), ('állampolgár', 0.9824466705322266), ('szempont', 0.9823504686355591), ('rendszer', 0.9804409742355347), ('bírótság', 0.980380654335022), ('alkotmányos', 0.9791594743728638), ('törvénytelen', 0.9789927005767822), ('lét', 0.9779316782951355), ('kötelező', 0.9777095317840576), ('rendelkezik', 0.977019727230072), ('csalás', 0.976833701133728)]

#### LAKOSSÁG - BAD-GOOD

[('szerződés', 0.9456784725189209), ('eset', 0.9440294504165649), ('deviza', 0.9354002475738525), ('alkalmaz', 0.9327237606048584), ('költség', 0.9311181902885437), ('rendszer', 0.9274798631668091), ('árfolyam', 0.9215385317802429), ('azaz', 0.9203047156333923), ('cég', 0.914757490158081), ('alap', 0.9143778681755066), ('hitel', 0.9139434099197388), ('jelenleg', 0.9122763872146606), ('szabály', 0.9115813970565796), ('kötelező', 0.908046543598175), ('feltétel', 0.9076741933822632)]