
Anotacija moralnih vrednosti u srpskom onlajn diskursu: izazovi i perspektive

Jelena Redli¹[\[0000-0003-2518-5286\]](#)

¹Univerzitet u Novom Sadu, Filozofski fakultet,
Odsek za srpski jezik i lingvistiku, Srbija,
redli@ff.uns.ac.rs

Scientific paper

DOI: 10.18485/judig.2026.2.ch19

Apstrakt

Rad se bavi problemima anotacije moralnih vrednosti u srpskom onlajn diskursu na primeru dela korpusa Corpus-Mor.SR-13, koji obuhvata objave sa Tvitera i Redita i sadrži oko 17.000 anotacija. U fokusu rada nisu samo anotacione kategorije i tehnički aspekti označavanja, već pre svega lingvistički i interpretativni problemi koji se javljaju u procesu anotacije kratkih, kontekstualno zavisnih i često implicitnih iskaza.

Polazi se od pretpostavke da moralna evaluacija u srpskom jeziku često nije izražena isključivo eksplicitnim leksičkim markerima, već da u njenom oblikovanju važnu ulogu mogu imati pragmatičke implikacije, ironija, izbor sintaksičkih struktura, diskursno pozicioniranje i kulturno prepoznatljivi obrasci vrednovanja. Posebna pažnja biće posvećena primerima u kojima dolazi do neslaganja među anotatorima, jer upravo takvi slučajevi mogu ukazati na granice površinske klasifikacije i na značaj šireg diskursnog i sociokulturnog konteksta. U radu će se razmatrati i specifičnosti pojedinih domena onlajn komunikacije, poput sportskog i političkog diskursa, u kojima se moralne vrednosti često iskazuju posredno, metaforički ili kroz kolektivne identitetske obrasce. Na osnovu analiziranog materijala nastojaće se da se ukaže na potrebu za razvojem detaljnijih smernica za anotaciju moralnih vrednosti u srpskom jeziku, kao i na značaj saradnje između lingvistike i savremenih pristupa obradi prirodnog jezika.

Ključne reči: anotacija korpusa, moralne vrednosti, onlajn diskurs, sportski diskurs, pragmatika, obrada prirodnog jezika, srpski jezik.

1. Uvod

Obrada prirodnog jezika poslednjih decenija sve više se usmerava ka analizi kratkih tekstova nastalih u digitalnom okruženju. Takvi

tekstovi, među kojima su objave na društvenim mrežama, komentari korisnika, naslovi vesti i drugi kratki medijski tekstovi, predstavlja-ju važan izvor podataka za istraživanje stavova, emocija i različitih oblika vrednosnog pozicioniranja u jeziku. Njihova analiza, međutim, otvara niz metodoloških problema, naročito kada se od računarskog modela ili ljudskog anotatora očekuje da prepozna značenja koja nisu eksplicitno iskazana.

Ovaj problem posebno je izražen u jezicima s ograničenim re-sursima, među kojima je i srpski.¹ Za takve jezike razvoj pouzdanih anotiranih korpusa ima dvostruki značaj: s jedne strane, oni predsta-vljaju osnovu za izradu i evaluaciju modela obrade prirodnog jezika, a s druge, omogućavaju uvid u jezičke pojave koje se ne mogu dovoljno precizno opisati samo pomoću gotovih alata i unapred definisanih klasifikacionih shema. U istraživanjima sentimenta kratkih tekstova već je ukazano na to da se vrednosni odnos može ostvariti različitim jezičkim i retoričkim sredstvima, uključujući ambigvitet, sarkazam, subjektivnost i mešovitu evaluaciju (Batanović et al., 2020: 1–2).

Još složeniji problem predstavlja anotacija moralnih vrednosti. Za razliku od sentimenta, koji se najčešće vezuje za pozitivnu, negativnu ili neutralnu orijentaciju iskaza, moralna evaluacija podrazumeva prepoznavanje odnosa prema postupcima, akterima, normama i dru-štveno relevantnim vrednostima. U ovom radu polazi se od teorije moralnih osnova i od njene kasnije anotacione operacionalizacije u korpusnim istraživanjima, ali se posebna pažnja posvećuje upravo pro-blemima koji nastaju kada se takve kategorije primenjuju na kratke, kontekstualno ograničene i žanrovski raznovrsne tekstove.

Polazeći od takvog teorijsko-metodološkog okvira, predmet ovog rada jesu lingvistički i interpretativni problemi koji se javljaju u procesu anotacije moralnih vrednosti u srpskom onlajn diskursu. Istra-živanje se zasniva na delu korpusa Corpus-Mor.SR-13, koji obuhvata anotirane jedinice iz objava sa Tvitera i Redita.² Sportski diskurs se

¹ Pod jezicima s ograničenim resursima u ovom radu podrazumevaju se jezici za koje postoji relativno manji broj dostupnih digitalnih korpusa, anotiranih skupova podataka, specijalizovanih leksikona, evaluiranih modela i alata za obradu prirodnog jezika. Izraz se ne odnosi na strukturnu složenost ili razvijenost jezika, već na stepen njegove pokrivenosti digitalnim jezičkim resursima.

² Anotacija dela primera korišćenih u ovom radu sprovedena je u okviru radionice TESLA – Tekstometrija u praksi, održane 25. aprila 2026. godine na Rudarsko-geološkom fakultetu Univerziteta u Beogradu, na kojoj je učestvovala autorka ovog rada.

u radu posmatra kao studija slučaja, a ne kao reprezentativni model celokupnog srpskog onlajn diskursa.³ Izdvojen je zato što u kratkoj i relativno ujednačenoj građi omogućava da se jasno uoče granice između informacije, sentimenta i moralne evaluacije.

Sportski diskurs je za ovu analizu naročito pogodan zato što na prvi pogled često deluje moralno neutralno. Velik broj sportskih naslova i kratkih vesti ima prvenstveno informativnu funkciju – saopštava se rezultat, transfer, povreda, odluka saveza, izjava trenera ili ponašanje navijača. Ipak, iskazi o porazu, kazni, izbacivanju navijača, nepravednoj presudi, odanosti klubu ili sumnji u regularnost postupka mogu ostati na nivou informacije, ali mogu aktivirati i kategorije povrede, pravednosti, izdaje, autoriteta ili subverzije.

Cilj rada je da se na izabranim primerima pokaže zbog čega se anotacija moralnih vrednosti u srpskom onlajn diskursu ne može oslanjati isključivo na prepoznavanje ključnih reči. U radu se posebno razmatraju granica između moralne evaluacije i opšteg sentimenta, problem implicitnog vrednovanja, preklapanje moralnih kategorija i slučajevi neslaganja među anotatorima. Takva neslaganja ne posmatraju se samo kao slabost anotacionog postupka već i kao pokazatelj interpretativne složenosti značenja morala u jeziku.

Na taj način rad nastoji da doprinese razumevanju anotacije moralnih vrednosti kao postupka koji se nalazi na preseku semantike, pragmatike, analize diskursa i savremene obrade prirodnog jezika. Istovremeno, analiza ukazuje na potrebu da se pri razvoju anotacionih smernica za srpski jezik više pažnje posveti diskursnom domenu, kontekstu i razlikovanju moralne vrednosti od drugih oblika vrednosnog značenja.

2. Moralne vrednosti, sentiment i leksičko-semantička evaluacija

U istraživanjima digitalnog diskursa često se prepliću različiti oblici vrednosne analize: analiza sentimenta, prepoznavanje emocija, analiza uvredljivog ili toksičnog govora⁴ i analiza moralnih vredno-

³ Studija slučaja ovde označava analitički izbor tematski ujednačenog poduzorka, a ne nameru da se na osnovu sportskog diskursa izvode zaključci o svim domenima srpskog onlajn diskursa. Cilj je da se na ograničenoj građi izdvoje tipovi interpretativnih problema koji se mogu proveravati i u drugim domenima.

⁴ Termin *toksični govor* u ovom radu koristi se kao prevodni ekvivalent za *toxic language* / *toxic speech* u literaturi o obradi prirodnog jezika. Njime se obuhvataju

sti. Iako se ovi pristupi delimično dodiruju, među njima ne treba uspostavljati znak jednakosti. Sentiment se, u najopštijem smislu, odnosi na pozitivnu, negativnu ili neutralnu orijentaciju iskaza, dok moralna evaluacija podrazumeva složeniji odnos prema postupku, akteru, društvenoj normi ili vrednosti. U istraživanjima sentimenta kratkih tekstova već je ukazano na to da se i naizgled jednostavne podele komplikuju čim se u obzir uzmu neutralnost, subjektivnost, ambigviteta, sarkazam i mešoviti sentiment (Batanović et al., 2020: 1–2). Zbog toga iskaz može biti negativno obojen, a da ne bude nužno moralno relevantan, kao što može biti moralno relevantan i onda kada ne sadrži izrazito negativnu ili pozitivnu leksiku. Na sličnu potrebu za pažljivim opisom jezički specifičnih anotacionih problema ukazuje i istraživanje Sarajlić i saradnika o anotaciji citata, koreferencije i sentimenta u hrvatskim novinskim tekstovima: autori pokazuju da izrada anotiranog resursa ne podrazumeva samo primenu unapred zadatih oznaka već i beleženje pojava koje mogu predstavljati problem za automatsku obradu, naročito kada se rešenja razvijana za engleski jezik prenose na južnoslovenski jezički materijal (Sarajlić et al., 2021: 1–5).

Ovo razlikovanje je posebno važno u analizi kratkih onlajn tekstova. U njima je značenje često kondenzovano, a kontekst delimično ili potpuno odsutan. Zato se anotator ne suočava samo s pitanjem da li je određeni iskaz pozitivan ili negativan već i s pitanjem da li se u njemu uspostavlja vrednosni odnos koji se može dovesti u vezu s nekom moralnom kategorijom. Tako, informacija o porazu tima ili povredi igrača može nositi negativan sentiment, ali ne mora sadržati moralnu evaluaciju. Nasuprot tome, iskaz o nepravednoj presudi, izbacivanju navijača, izdaji kluba ili zloupotrebi autoriteta može aktivirati moralne kategorije i onda kada je formulisan kao kratak novinski naslov.

U teorijskom smislu, analiza moralnih vrednosti u ovom radu oslanja se na teoriju moralnih osnova, prema kojoj se moralno rasuđivanje može povezati s nekoliko osnovnih kategorija: brigom i povredom, pravednošću i reciprocitetom, lojalnošću grupi, autoritetom i poštovanjem, odnosno svetošću i čistotom (Graham et al., 2009: 1029–1031).

različiti oblici štetne, uvredljive, zlostavljачke ili mrzilačke onlajn komunikacije. Pavlopoulos i saradnici termin *toxic* koriste kao krovni naziv za srodne pojave koje se mogu okarakterisati kao *offensive*, *abusive*, *hateful* (Pavlopoulos et al., 2020: 4298), dok Sap i saradnici ukazuju na to da procena toksičnosti ne zavisi samo od teksta već i od socijalnog i dijalekatskog konteksta govornika (Sap et al., 2019: 1668–1669).

U anotacionoj praksi ove kategorije se često operacionalizuju preko pozitivnih i negativnih polova: *care/harm*, *fairness/cheating*, *loyalty/betrayal*, *authority/subversion* i *purity/degradation*.⁵ Takva podela omogućava preciznije označavanje, ali istovremeno otvara pitanje višestrukih oznaka, razgraničenja dominantne moralne kategorije i potrebe da se posebno obeleže iskazi bez moralnog sadržaja (Cao et al., 2024: 155–157, 164).

Zbog toga je za ovaj rad važna i razlika između leksičkog značenja i značenja koje se ostvaruje u konkretnoj upotrebi. Leksičko značenje ne iscrpljuje sve ono što određeni izraz može značiti u tekstu. Pored centralnog deskriptivnog značenja, važnu ulogu može imati i asocijativno značenje, koje obuhvata stilska, ekspresivna i konotativna obeležja, pri čemu treba razlikovati sistemsko od tekstualnog, odnosno nameravano od protumačenog značenja (Prčić, 2016: 24–36). Dražić (2022: 15–16), polazeći od kompleksne prirode značenja reči, ističe da se značenje u lingvističkoj analizi može sagledavati kroz više međusobno povezanih nivoa, među kojima izdvaja i nivo konteksta, odnosno razliku između izolovanog i kontekstualizovanog, kao i između iskazanog i neiskazanog. Takvo shvatanje značenja posebno je relevantno za analizu kratkih onlajn iskaza, u kojima je odnos između eksplicitno datog i kontekstualno protumačenog često presudan za anotaciju moralnih vrednosti. Upravo u prostoru između onoga što je leksički dato i onoga što se u kontekstu tumači nastaje veliki deo problema u anotaciji moralnih vrednosti.

Ekspresivna i evaluativna leksika dodatno pokazuje da značenje nije samo referencijalno već uključuje i stav govornika prema imenovanom sadržaju. Ristić (1996: 57–58) ekspresivnu leksiku dovodi u vezu sa specifičnošću njene konotativne semantike, a jedan njen tip opisuje preko modela „osoba + psihička ili moralna (karakterna) osobina”. U analizi pejorativa Jovanović (2021: 127–129) pokazuje da se takvim jedinicama čovek ne samo imenuje nego i kvalifikuje, pri čemu njihova konotacija može biti pogrdna, prezriva, podrugljiva ili ironična, dok se negativna ocena često uspostavlja u odnosu na kolektivna uverenja, stereotipe i etalone srpske društveno-jezičke zajednice. Posebno je važno to što se u konotaciji pejorativa prepliću evaluativna, emotivna i ekspresivna komponenta (Јовановић, 2021: 138).

⁵ Engleski nazivi anotacionih kategorija zadržani su onda kada se govori o oznakama u korpusu, jer se u tom obliku javljaju u anotacionoj shemi. Srpski prevodi koriste se opisno i ne uvode se kao zasebna terminološka norma.

U tom smislu, ključne reči i leksički markeri mogu biti korisni kao početni orijentiri, ali nisu dovoljni za pouzdanu anotaciju. Reči kao što su *nepravda*, *izdaja*, *žrtva*, *podrška*, *kazna* ili *autoritet* mogu signalizirati moralnu relevantnost, ali njihovo prisustvo ne garantuje automatski određenu moralnu oznaku. Isto tako, odsustvo takvih reči ne znači da moralna evaluacija nije prisutna. U kratkim onlajn iskazima vrednosno značenje često proizlazi iz odnosa između naslova i šireg događaja, iz pretpostavljenog znanja čitaoca, ironične distance ili načina na koji se akteri postavljaju jedni prema drugima.

Poseban problem predstavlja anotatorsko neslaganje. U tehničkim pristupima anotaciji ono se često posmatra kao pokazatelj nedovoljne pouzdanosti ili nepreciznosti smernica. Međutim, u slučaju moralne evaluacije neslaganje može imati i drugačiju vrednost. Ono može ukazati na mesta na kojima je značenje zaista granično, višeznačno ili zavisno od dodatnog konteksta. U istraživanjima toksičnosti pokazano je da kontekst može promeniti ljudsku procenu iskaza, odnosno, da isti komentar može biti drugačije ocenjen kada anotator, pored samog komentara, vidi i prethodni komentar i naslov diskusije (Pavlopoulos et al., 2020: 4296–4298). S druge strane, istraživanja pristrasnosti u detekciji uvredljivog i toksičnog govora pokazuju da procena ne zavisi samo od teksta već i od socijalnog i dijalekatskog konteksta govornika, kao i od toga da li su anotatori na taj kontekst eksplicitno upućeni (Sap et al., 2019: 1668–1669, 1671). U vezi s tim, za analizu vrednosno osetljivog onlajn diskursa relevantno je i istraživanje Lj. Šarić, koje pokazuje da se procene nepristojnosti u korisničkim komentarima ne mogu odvojiti od žanra, teme i kontekstualizacije iskaza, kao ni od kulturnog i okvirnog znanja čitalaca. Autorka pritom ukazuje i na vezu između negativno vrednovanih jezičkih sredstava i diskurzivne konstrukcije kolektivnog identiteta, odnosno granice između „nas” i „njih” (Šarić, 2022: 1–4, 10).

Slično se pokazuje i u anotaciji moralnih vrednosti. Iz anotacione organizacije koju opisuju Cao i saradnici proizlazi važno metodološko pitanje: da li se moralne osnove označavaju na širem nivou ili preko pojedinačnih pozitivnih i negativnih polova. Autori u korpusu *Moral Foundations Weibo Corpus* primenjuju drugo rešenje, odnosno anotaciju prema deset kategorija: *care/harm*, *fairness/cheating*, *loyalty/betrayal*, *authority/subversion* i *purity/degradation*, uz oznaku, *non-moral* za iskaze bez moralnog sadržaja i uz naknadno razrešavanje slučajeva u kojima jedna objava tokom anotacije dobija

više oznaka (Cao et al., 2024: 155–157, 164). Posebno je važno njihovo upozorenje da neslaganje među anotatorima ne treba uvek veštački smanjivati, jer nizak stepen saglasnosti može ukazivati na to da je veza između iskaza i određene moralne osnove slaba, granična ili zavisna od dodatnih uslova interpretacije (Cao et al., 2024: 165).

Stoga se u ovom radu anotacija moralnih vrednosti posmatra kao postupak koji se ne može svesti na mehaničko povezivanje teksta i oznake. Ona podrazumeva interpretaciju značenja na preseku leksike, konteksta, diskursnog domena i anotacionih smernica. Takav pristup posebno je važan za sportski diskurs, u kojem isti iskaz može funkcionisati kao neutralna vest, izraz afektivnog stava ili moralno relevantna procena postupka, odluke ili aktera.

3. Korpus i metodološki okvir

Istraživanje se zasniva na materijalu iz korpusa Corpus-Mor.SR-13 (Redli, 2025), koji obuhvata anotirane jedinice iz srpskog onlajn diskursa, prvenstveno s Tvitera i Redita. Korpus je nastao kao resurs za ispitivanje moralnih vrednosti u digitalnoj komunikaciji, a anotacija je sprovedena prema kategorijama zasnovanim na teoriji moralnih osnova: *care/harm*, *fairness/cheating*, *loyalty/betrayal*, *authority/subversion* i *purity/degradation*. Pored moralno relevantnih oznaka, u anotacionom okviru predviđena je i oznaka *non-moral*, namenjena primerima u kojima nije moguće utvrditi prisustvo moralne evaluacije.⁶

Za potrebe ovog rada ne analizira se ceo korpus, već radni uzorak koji obuhvata 1016 anotiranih jedinica. Svaka jedinica sadrži tekst, više anotacionih oznaka i podatak o tome koja je oznaka prihvaćena kao validna. Drugim rečima, validna oznaka nije unapred vezana za jednu fiksnu kolonu, već se određuje na osnovu posebno obeleženog izvora konačne oznake.⁷ Zbog toga je u pripremi materijala bilo potrebno rekonstruisati završnu, odnosno validnu oznaku za svaku jedinicu.

⁶ Oznaka *non-moral* u ovom radu ne znači ‘nemoralno’, već odsustvo prepoznatljive moralne evaluacije u posmatranoj tekstualnoj jedinici. Zbog toga bi doslovni prevod mogao biti pogrešan.

⁷ Termin *validna oznaka* koristi se u tehničkom smislu, prema strukturi korpusne tabele. On ne podrazumeva da je data oznaka jedino moguće tumačenje, već da je u korpusu prihvaćena kao konačna u odnosu na dostupne anotatorske oznake.

U okviru tog uzorka posebno se izdvaja poduzorak od 102 primera iz sportskog diskursa, koji je u ovom radu uzet kao osnova za kvalitativnu analizu. Razlog za izdvajanje upravo ovog poduzorka nije samo njegova tematska ujednačenost već i činjenica da sportski diskurs predstavlja pogodan teren za ispitivanje granica između neutralnog obaveštavanja, sentimenta i moralne evaluacije. Sportski naslovi često se formulišu kratko, eliptično i informativno, ali istovremeno mogu uključivati elemente vrednovanja, pripadnosti, sukoba, odgovornosti, kazne, nepravde ili institucionalne moći.

U metodološkom smislu, analiza je kvalitativna i usmerena na primere koji pokazuju interpretativne probleme u anotaciji. To znači da cilj nije kvantitativno merenje učestalosti pojedinih moralnih kategorija u sportskom diskursu, već ispitivanje slučajeva u kojima se vidi kako i zašto nastaje dilema u određivanju moralne vrednosti.⁸ Takav pristup je primeren prirodi istraživanja jer se moralna evaluacija u kratkim onlajn tekstovima često ne može pouzdano tumačiti samo na osnovu leksičkog markera ili unapred zadate liste ključnih reči.

Pri izboru primera uzeti su u obzir sledeći kriterijumi: prisustvo neslaganja među anotatorima, razlika između opšteg sentimenta i moralne vrednosti, mogućnost da se isti iskaz protumači kao moralno relevantan ili kao *non-moral*, preklapanje dveju ili više moralnih kategorija i zavisnost interpretacije od šireg konteksta, domena ili znanja o akterima.

Posebno su važni primeri u kojima se oznake anotatora razilaze. U ovom radu takva neslaganja ne posmatraju se isključivo kao nedostatak anotacionog postupka već kao lingvistički relevantni podaci. Ona pokazuju mesta na kojima je moralno značenje granično, implicitno ili zavisno od konteksta. Sličan problem prepoznat je i u istraživanjima anotacije kratkih tekstova i vrednosno osetljivog diskursa, gde se pokazuje da anotatorska procena može zavisiti od konteksta, subjektivnosti, sarkazma, dvosmislenosti, kao i od socijalnih i dijalekatskih informacija koje su anotatoru dostupne (Batanović et al., 2020: 1-2; Pavlopoulos et al., 2020: 4296–4298; Sap et al., 2019: 1668–1669).

⁸ U radu se zato ne izračunavaju mere saglasnosti anotatora. One bi bile važne za evaluaciju celog anotacionog postupka, ali ovde nisu primarni predmet analize. Težište je na kvalitativnom opisu slučajeva u kojima samo neslaganje ima interpretativnu vrednost.

U analizi se zato ne polazi od pretpostavke da je svaka razlika među oznakama posledica greške. Takvi primeri mogu ukazati na nedovoljno jasne granice između kategorija, ali i na samu prirodu moralne evaluacije u jeziku. Sportski poraz može biti emocionalno negativan, ali ne mora biti moralno relevantan; kazna ili izbacivanje navijača mogu se tumačiti kao pitanje autoriteta, pravednosti ili proceduralnog pravila, dok se izraz lojalnosti prema klubu ili navijačkoj grupi može kretati između opšte afektivne pripadnosti i moralne kategorije *loyalty*.

Izabrani primeri grupisani su prema tipu anotacionog problema koji pokazuju. U analitičkom delu rada posebno se razmatraju granica između oznake *non-moral* i moralne evaluacije, razlika između sentimenta i moralne vrednosti, implicitna moralna evaluacija, preklapanje moralnih kategorija i neslaganje anotatora kao pokazatelj interpretativne složenosti. Ovakva organizacija omogućava da se sportski diskurs ne posmatra kao tematski izolovan domen, već kao studija slučaja koja osvetljava opštije probleme moralne anotacije u srpskom onlajn diskursu.

4. Sportski diskurs kao granično područje moralne evaluacije

Sportski diskurs u ovom radu posmatra se kao granično područje moralne evaluacije. Njegova pogodnost za analizu ne proizlazi iz pretpostavke da je sport sam po sebi moralno obeležen, već upravo iz činjenice da se u sportskim naslovima često susreću informativnost, afektivnost i vrednosno tumačenje. Veliki broj naslova saopštava samo rezultat, najavu utakmice, transfer, povredu ili izjavu trenera. U takvim slučajevima ne mora postojati moralno relevantan sadržaj, čak ni onda kada je iskaz pozitivno ili negativno obojen.

Istovremeno, sportski diskurs počiva na opozicijama koje lako dobijaju vrednosnu dimenziju: pobeda i poraz, uspeh i neuspeh, poštovanje i kršenje pravila, pripadnost timu, odgovornost aktera i legitimnost odluka koje donose sportske ili druge institucije. Zbog toga se u njemu često otvara prostor za moralno tumačenje, ali taj prostor nije uvek eksplicitno obeležen. Sportski naslovi su kratki, često eliptični, oslonjeni na prethodno znanje čitaoca i na širi sportski, medijski ili navijački kontekst. Upravo zbog toga sportski diskurs omogućava da se preciznije ispita razlika između nekoliko bliskih, ali nejednakih pojava. Prva je neutralno obaveštavanje, u kojem se prenosi informacija o događaju bez vrednosnog stava. Druga je sentiment, odnosno,

pozitivan ili negativan odnos prema ishodu, akteru ili događaju. Treća je moralna evaluacija, koja nastaje tek onda kada se događaj dovodi u vezu s normom, odgovornošću, povredom, pravednošću, lojalnošću, autoritetom ili drugom moralnom kategorijom.

Za analizu su posebno značajni naslovi koji se odnose na kazne, izbacivanje navijača, sumnju u regularnost postupka, presude, povrede ili izraze pripadnosti navijačkoj zajednici. U takvim primerima moralna relevantnost može biti eksplicitna, kao kada se upotrebi izraz nepravde, ali može biti i implicitna, kada se uspostavlja kroz pretpostavljenu neregularnost, institucionalnu moć ili kolektivnu lojalnost. Isto tako, postoje primeri u kojima negativan ishod, poput poraza ili povrede, može izazvati afektivnu reakciju, ali ne mora nužno predstavljati moralnu evaluaciju.

Sportski poduzorak se zato u radu ne koristi kao ilustrativni dodatak, već kao studija slučaja koja pokazuje gde se u anotaciji moralnih vrednosti pojavljuju granični slučajevi. U njemu su naročito vidljivi problemi kratkoće iskaza, nedostatka konteksta, oslanjanja na zajedničko znanje, preklapanja moralnih kategorija i različitih anotatorskih interpretacija. Na toj osnovi analiza sportskog diskursa može doprineti opštijem razumevanju moralne evaluacije u srpskom onlajn diskursu.

5. Tipovi anotacionih problema

Analiza sportskog poduzorka pokazuje da se problemi u anotaciji moralnih vrednosti ne javljaju samo onda kada je značenje iskaza nejasno, već i onda kada se na prvi pogled čini da je oznaka očigledna. Upravo kratkoća sportskih naslova, njihova oslonjenost na širi kontekst i česta upotreba vrednosno obojenih izraza dovode do toga da isti iskaz može biti protumačen kao neutralna informacija, kao izraz sentimenta ili kao moralno relevantna evaluacija.

U nastavku se izdvajaju osnovni tipovi problema koji su se pokazali značajnim u analiziranom materijalu: granica između oznake *non-moral* i moralne evaluacije, razlika između sentimenta i moralne vrednosti, implicitna moralna evaluacija, preklapanje moralnih kategorija i neslaganje među anotatorima. Reprezentativni primeri iz poduzorka dati su u Tabeli 1.

Tabela 1: Reprezentativni primeri iz sportskog poduzorka.

<i>Primer</i>	<i>Konačna oznaka</i>	<i>Tip problema</i>
Gvardiola priželjkuje Silvin ostanak	non-moral	sentiment / želja nije moralna vrednost
Kerkez posle poraza od Tventea: Naivno smo primili golove, ovo je drugi nivo	non-moral	negativan ishod nije moralna vrednost
Partizan: Naši navijači izbačeni jer nisu imali pasoše	autoritet (authority)	pravila, autoritet, perspektiva
Iza Džakinog žutog kartona stoji albanska mafija?	varanje (cheating)	implicitna sumnja u regularnost
VNBA i NBA: Presuda Grajnerovoj nepravedna i za žaljenje	pravednost (fairness)	eksplicitni moralni marker
Zvezdin junak Bukari: Volim Delije! Vole i oni mene!	lojalnost (loyalty)	kolektivni identitet

5.1. Granica između oznake non-moral i moralne evaluacije

Jedan od osnovnih problema u anotaciji moralnih vrednosti jeste određivanje trenutka u kojem tekst prestaje da bude samo informativan i postaje moralno relevantan. U sportskom diskursu veliki broj naslova saopštava događaj bez eksplicitnog vrednosnog komentara. Takvi su primeri (1) *Ulaznice za meč Partizan – AEK od ponedeljka u prodaji*; i (2) *Trener AEK-a o revanšu: Idemo u Beograd po pobedu i prolaz*. U njima se ne uspostavlja vrednosni odnos prema postupku, akteru ili normi, već se prenosi informacija o događaju ili nameri. Zbog toga predstavljaju relativno jasne slučajeve oznake non-moral.

Granica postaje nejasnija u primerima u kojima naslov sadrži negativan događaj ili institucionalni postupak, ali bez eksplicitne kvalifikacije: (3) *CSKA tužio Evroligu zbog izbacivanja iz takmičenja*; (4) *Britni Grajner osuđena na devet godina zatvora*. Oba primera uključuju situacije koje se mogu povezati s pravdom, pravilima, odgovornošću ili autoritetom. Ipak, sam tekst naslova ne iznosi stav o tome da li je tužba osnovana, izbacivanje opravdano ili presuda nepravedna. U takvim slučajevima moralna relevantnost zavisi od toga da li anotator

označava samo ono što je tekstualno dato ili u tumačenje uključuje širi kontekst događaja.

Kontrastni primer pruža naslov (5) *VNBA i NBA: Presuda Grajnerovoj nepravedna i za žaljenje*. Za razliku od prethodnog primera, pridev *nepravedna* neposredno uvodi moralnu kvalifikaciju. Time se ne saopštava samo činjenica da je presuda doneta već se iskazuje vrednosni odnos prema njoj. Ovakav primer pokazuje koliko jedan eksplicitni leksički marker može promeniti anotacioni status iskaza: ono što je u primeru (4) moglo ostati u domenu neutralnog izveštavanja, u primeru (5) postaje jasno povezano s kategorijom pravednosti.

5.2. Sentiment nije isto što i moralna vrednost

Drugi čest problem odnosi se na razlikovanje sentimenta od moralne evaluacije. U sportskom diskursu negativan sentiment često je posledica poraza, povrede, eliminacije ili loše igre. Ipak, takvi iskazi nisu nužno moralno relevantni.

Naslov (6) *Kerkez posle poraza od Twentea: Naivno smo primili golove, ovo je drugi nivo* sadrži negativnu procenu igre i posledica poraza. Izraz *naivno smo primili* golove ukazuje na samokritički odnos, ali se taj odnos pre svega tiče sportskog kvaliteta, koncentracije ili iskustva, a ne moralne vrednosti. Nema jasnog upućivanja na krivicu u moralnom smislu, nepravdu, povredu drugoga, izdaju ili zloupotrebu autoriteta. Zato negativna evaluacija u ovom primeru ne mora voditi moralnoj oznaci.

Slično se vidi u primeru (7) *Spektakl u Trebinju: Delfini posle peteraca poraženi od Crne Gore*. Poraz može izazvati negativnu reakciju čitaoca, naročito ako je reč o reprezentaciji ili timu s kojim se publika identifikuje. Međutim, sam naslov ne vrednuje postupak, ne problematizuje pravila igre i ne uvodi moralno obeleženu odgovornost. On saopštava ishod. Ukoliko bi se ovakav primer označio kao *povreda* ili *izdaja*, to bi verovatno bilo posledica prenosa emocionalne reakcije na moralnu kategoriju.

Ovi primeri pokazuju da je u anotaciji neophodno razlikovati tri nivoa: nepovoljan događaj, negativan sentiment i moralnu evaluaciju. Nepovoljan događaj može biti sportski poraz, negativan sentiment može biti razočaranje zbog poraza, a moralna evaluacija nastaje tek onda kada se događaj tumači kao povreda norme, odgovornosti, pravednosti, lojalnosti ili neke druge moralne vrednosti.

5.3. Implicitna moralna evaluacija

Moralna vrednost ne mora biti izražena eksplicitnom leksičkom jedinicom. U pojedinim primerima ona se uspostavlja preko implikacije, odnosa prema pravilima ili pretpostavljene neregularnosti.

Naslov (8) *Partizan: Naši navijači izbačeni jer nisu imali pasoše* na prvi pogled saopštava razlog zbog kojeg su navijači izbačeni. Međutim, sama struktura iskaza otvara prostor za više tumačenja. Izbacivanje navijača može se posmatrati kao institucionalna mera zasnovana na pravilima, pa se može povezati s kategorijom **autoritet**. Istovremeno, ukoliko se u prvi plan stavi pitanje opravdanosti postupka, mogu se aktivirati i kategorije **pravednost** ili **varanje**. Upravo zato ovaj primer pokazuje kako se moralna evaluacija može javiti na preseku pravila, institucionalne moći i perspektive iz koje se događaj posmatra.

Još jasniji primer implicitne moralne evaluacije jeste (9) *Iza Džakinog žutog kartona stoji albanska mafija?*. U ovom iskazu nema eksplicitne reči *prevara*, *manipulacija* ili *nepravda*. Ipak, pitanje sugerise mogućnost skrivenog uticaja i neregularnosti. Moralna relevantnost proizlazi iz pretpostavke da sportska odluka nije rezultat igre ili pravila, već delovanja nelegitimnog aktera. Zbog toga se ovaj primer može povezati s kategorijom **varanje**, iako odgovarajući leksički marker nije neposredno prisutan. Istovremeno, upitna forma zahteva oprez: naslov ne tvrdi otvoreno da je do manipulacije došlo, nego uvodi sumnju.

5.4. Preklapanje moralnih kategorija

Pojedini primeri pokazuju da ista tekstualna jedinica može aktivirati više moralnih kategorija. To se naročito javlja u naslovima koji uključuju pravila, kaznu, institucije, povredu ili kolektivnu pripadnost.

Naslov (10) *Većiti plaćaju zbog pirotehнике, Kragujevčani i Pazarci u problemu zbog navijača* može biti tumačen kroz kategoriju **autoritet**, jer se pominju kazna, pravila i posledice navijačkog ponašanja. Međutim, u zavisnosti od ugla posmatranja, moguće je uključiti i pitanje odgovornosti, krivice ili pravednosti. Ako je akcenat na instituciji koja sankcioniše, aktivira se **autoritet**; ako je akcenat na navijačima koji su svojim ponašanjem doveli klub u problem, moguće

je govoriti o povredi kolektivnog interesa; ako je akcenat na tome da li je kazna opravdana, otvara se domen pravednosti.

Drugačiji tip preklapanja vidi se u primeru (11) *Pavlović se povredio na prvom treningu sa ekipom Partizana*. Ovaj naslov prirodno se povezuje s kategorijom **povređa**, jer je reč o povredi. Međutim, u anotaciji se može javiti i **briga**, naročito ako anotator u prvi plan stavi saosećanje ili brigu za povređenog sportistu. Razlika između povrede i brige ovde nije razlika između dva potpuno različita sadržaja, već između fokusa na negativni događaj i fokusa na reakciju prema osobi koja je tim događajem pogođena.

Treći tip preklapanja odnosi se na kolektivni identitet. Primer (12) *Zvezdin junak Bukari: Volim Delije! Vole i oni mene!* sadrži očigledno pozitivan afektivni odnos, ali se ne iscrpljuje u pozitivnom sentimentu. Izjava uspostavlja uzajamnu vezu između igrača i navijačke grupe, pri čemu se aktivira pripadnost zajednici. Zbog toga se može tumačiti kroz kategoriju **lojalnost**. Međutim, granica ni ovde nije potpuno jednostavna: ljubav prema navijačima može biti shvaćena kao emotivna naklonost, kao medijska formula ili kao izraz lojalnosti prema kolektivnom identitetu kluba.

5.5. Neslaganje anotatora kao lingvistički podatak

U analiziranom materijalu posebno su značajni primeri u kojima anotatori nisu došli do iste oznake. Takva neslaganja ne treba posmatrati samo kao grešku ili kao slabost anotacionog postupka. Ona često pokazuju da se tekst nalazi na granici između više mogućih interpretacija.

U naslovu *Gvardiola priželjkuje Silvin ostanak* anotacije se kreću između moralnih oznaka i **non-moral**. Problem nastaje zato što glagol *priželjkivati* izražava stav, ali taj stav nije moralni. On upućuje na želju, očekivanje ili sportski interes, a ne na procenu postupka kao dobrog, lošeg, pravednog, nepravednog, lojalnog ili nelojalnog. Neslaganje u takvom primeru pokazuje koliko lako prisustvo subjektivnosti može biti pogrešno prepoznato kao moralna relevantnost.

Slično se vidi u primerima sportskih poraza. Naslovi poput *Spektakl u Trebinju: Delfini posle peteraca poraženi od Crne Gore* ili *Kerkez posle poraza od Twentea: Naivno smo primili golove, ovo je drugi nivo* mogu izazvati negativan sentiment, ali sam negativan ishod ne mora biti moralno obojen. Ako anotator negativnost događaja tumači kao povredu ili izdaju, moralna oznaka se zapravo oslanja na

emocionalno čitanje događaja, a ne na moralnu evaluaciju prisutnu u tekstu.

Nasuprot tome, kod primera *Partizan: Naši navijači izbačeni jer nisu imali pasoše* neslaganje je drugačije prirode. Tu ne dolazi samo do mešanja sentimenta i moralnosti već do različitog određivanja centralnog aspekta iskaza. Jedan anotator može u prvi plan staviti institucionalnu odluku, drugi posledicu po navijače, treći pitanje opravdanosti pravila. Zato su moguće oznake iz domena *autoritet*, *pravednost/varanje* ili *non-moral*, u zavisnosti od toga šta se smatra dominantnim značenjem naslova.

Ovakvi primeri potvrđuju da anotacija moralnih vrednosti ne zavisi samo od prisustva ili odsustva ključnih reči nego od odluke o tome koji element iskaza nosi vrednosnu relevantnost. U tom smislu, neslaganje među anotatorima može biti korisno upravo zato što pokazuje gde su granice kategorija nestabilne. Ono ukazuje na mesta na kojima je potrebno preciznije razgraničiti sentiment, moralnu evaluaciju i neutralnu informaciju, ali i na mesta na kojima je sama priroda iskaza zaista višeznačna.

Zato se u ovom radu anotatorsko neslaganje posmatra kao deo analitičkog materijala. Ono ne ukida vrednost anotacije, već pokazuje da moralno značenje u kratkom onlajn diskursu često nastaje u prostoru između leksičkog sadržaja, diskursnog domena, šireg konteksta i interpretativnih očekivanja anotatora.

6. Diskusija

Analiza izabranih primera pokazuje da se moralna anotacija kratkih sportskih tekstova ne može svesti na prepoznavanje moralno obeleženih reči. Leksički markeri mogu biti važan oslonac, naročito u primerima u kojima se eksplicitno javlja izraz poput *nepravedna*, ali oni nisu dovoljni za stabilno određivanje moralne kategorije. S druge strane, odsustvo takvih markera ne znači da moralna evaluacija nije prisutna, jer se ona može uspostaviti posredno, kroz kontekst, pretpostavljenu neregularnost, odnos prema autoritetu ili pripadnost kolektivu.

Ova činjenica posebno je važna za anotaciju kratkih onlajn iskaza. Kod dužih tekstova anotator često ima širi kontekst na osnovu kojeg može rekonstruisati stav autora, odnos među akterima i vrednosni okvir iskaza. U naslovima i kratkim objavama taj kontekst je uglavnom sveden na minimum. Zato se anotator nalazi u situaciji da

odlučuje da li treba da označi samo ono što je tekstualno dato ili i ono što je u datom diskursnom okviru verovatno podrazumevano. Upravo na toj granici nastaju mnoga neslaganja.

U sportskom diskursu ovaj problem dodatno je pojačan time što se velik broj negativnih događaja lako doživljava kao vrednosno značajan. Poraz, povreda, eliminacija, kazna ili izbacivanje navijača najčešće aktiviraju afektivnu reakciju, ali ta reakcija nije uvek moralna. Moralna relevantnost nastaje tek onda kada se događaj predstavi kao kršenje pravila, povreda očekivane odgovornosti, nelojalnost, zloupotreba moći ili narušavanje neke druge vrednosti.

Slično važi i za pozitivne iskaze. Izrazi podrške, zahvalnosti ili želje ne moraju sami po sebi biti moralno relevantni. Ove kategorije postaju relevantnije onda kada se odnos između aktera ne svodi samo na afektivnu naklonost nego se predstavlja kao obaveza, odanost, solidarnost ili pripadnost grupi. Zato je u anotacionim smernicama važno jasnije razlikovati emotivnu naklonost od moralno shvaćene lojalnosti ili brige.

Poseban izazov predstavljaju primeri u kojima se prepliću pravila, kazne i institucije. U takvim slučajevima moguće je više interpretativnih pravaca. Ako se u prvi plan stavi institucija koja sprovodi pravilo, primer se može povezati sa kategorijom *autoritet*. Ako se u prvi plan stavi opravdanost odluke, otvara se domen *pravednost/varanje*. Ako se posmatra posledica po grupu, može se aktivirati i odnos prema kolektivnoj odgovornosti ili povredi interesa zajednice. To pokazuje da oznaka ne zavisi samo od teksta, nego i od toga koji se element iskaza smatra semantički i pragmatički centralnim.

Iz toga proizlazi da anotatorsko neslaganje u ovom tipu materijala ne treba tumačiti isključivo kao neuspeh anotacije. Deo neslaganja može biti posledica nedovoljno preciznih smernica, neujednačenog razumevanja kategorija ili različitog stepena poznavanja sportskog konteksta. Međutim, deo neslaganja ima dublju lingvističku osnovu: ono pokazuje da se moralna vrednost često uspostavlja na granici između leksičkog značenja, diskursnog okvira i interpretacije. Takvi slučajevi mogu biti posebno korisni jer ukazuju na mesta na kojima je potrebno razviti preciznije kriterijume za razlikovanje moralno relevantnih i moralno neutralnih iskaza.

Na osnovu analiziranog materijala mogu se izdvojiti najmanje tri metodološke posledice za dalji rad na anotaciji moralnih vrednosti u srpskom jeziku. Prvo, smernice bi morale eksplicitno da razgraniče

negativan sentiment, nepovoljan događaj i moralnu evaluaciju. Drugo, potrebno je jasnije opisati kada se pojedina kategorija aktivira na osnovu samog teksta, a kada tek na osnovu šireg konteksta. Treće, trebalo bi predvideti poseban tretman graničnih primera, naročito onih u kojima se moralna relevantnost može tumačiti samo kao implicitna ili slaba.

Ovo je važno i za razvoj računarskih modela. Ako ljudski anotatori nisu saglasni oko toga da li je određen iskaz moralno relevantan, teško je očekivati da će model pouzdano naučiti stabilnu vezu između površinskih jezičkih obeležja i moralnih kategorija. U tom smislu, kvalitativna analiza anotatorskih neslaganja nije sporedna dopuna kvantitativnoj obradi, već može imati ključnu ulogu u izradi boljih anotacionih smernica i pouzdanijih korpusa.

Sportski diskurs je u ovom radu poslužio kao studija slučaja, ali problemi koji su u njemu uočeni nisu ograničeni samo na sportske tekstove. Slične dileme mogu se očekivati i u političkom, medijskom, konfliktnom ili svakodnevnom onlajn diskursu, posebno onda kada su iskazi kratki, eliptični i oslonjeni na zajedničko znanje učesnika komunikacije. Zbog toga rezultati ove analize ukazuju na šire pitanje: kako u anotaciji moralnih vrednosti uspostaviti ravnotežu između doslednosti oznaka i očuvanja interpretativne složenosti jezika.

7. Zaključak

U radu su razmatrani lingvistički i interpretativni problemi koji se javljaju u anotaciji moralnih vrednosti u srpskom onlajn diskursu, na primeru sportskog poduzorka iz korpusa Corpus-Mor.SR-13. Polazna pretpostavka bila je da moralna evaluacija u kratkim digitalnim tekstovima ne može biti pouzdano određena samo na osnovu eksplicitnih leksičkih markera već da zahteva uzimanje u obzir konteksta, diskursnog domena, implicitnih značenja i mogućeg preklapanja anotacionih kategorija.

Analiza je pokazala da sportski diskurs predstavlja pogodno područje za ispitivanje ovih problema upravo zato što često stoji na granici između informacije, sentimenta i moralnog vrednovanja. Poraz, povreda, kazna, izbacivanje navijača, institucionalna odluka, sumnja u regularnost ili pripadnost timu mogu aktivirati moralnu interpretaciju, ali ne znače automatski da je iskaz moralno relevantan. Zato je u anotaciji važno razlikovati nepovoljan događaj, negativan ili pozitivan sentiment i moralnu evaluaciju.

Posebno se pokazalo da oznaka *non-moral* nije samo rezidualna kategorija za primere koji ne sadrže moralne vrednosti već važan analitički instrument. Ona omogućava da se razgraniče neutralni ili samo afektivno obojeni iskazi od onih u kojima se zaista uspostavlja vrednosni odnos prema normi, postupku, akteru ili zajednici. U tom smislu, precizno određivanje oznake *non-moral* jednako je važno kao i prepoznavanje pojedinačnih moralnih kategorija.

Rezultati analize ukazuju i na to da anotatorsko neslaganje ne treba posmatrati isključivo kao tehnički problem ili kao dokaz neujednačenosti anotacije. U mnogim slučajevima ono otkriva mesta na kojima je samo značenje granično, implicitno ili zavisno od šireg konteksta. Takvi primeri mogu biti posebno korisni za razvoj anotacionih smernica, jer pokazuju gde je potrebno preciznije razgraničiti sentiment, moralnu relevantnost i neutralnu informaciju, kao i kako tretirati slučajeve u kojima se prepliću dve ili više moralnih kategorija.

Iz toga proizlazi da moralna anotacija srpskog onlajn diskursa zahteva kombinovanje leksičko-semantičke, pragmatičke i diskursne analize. Ključne reči mogu biti početni signal, ali ne mogu zameniti interpretaciju iskaza u konkretnom domenu. Za dalji razvoj resursa namenjenih obradi prirodnog jezika posebno je važno da se anotacione smernice ne zasnivaju samo na listama primera i oznaka, već i na opisu graničnih slučajeva, mogućih izvora neslaganja i kriterijuma za razlikovanje moralne evaluacije od drugih oblika vrednosnog značenja.

Iako je analiza u ovom radu ograničena na sportski poduzorak, uočeni problemi imaju širi značaj za anotaciju moralnih vrednosti u srpskom jeziku. Kratkoća iskaza, zavisnost od konteksta, implicitnost i preklapanje kategorija nisu svojstveni samo sportskom diskursu, već predstavljaju opšte izazove u analizi onlajn komunikacije. Zbog toga kvalitativna analiza anotatorskih dilema može doprineti ne samo boljem razumevanju moralne evaluacije u jeziku već i izradi pouzdanih anotiranih korpusa za savremenu obradu prirodnog jezika.

Zahvalnica

Ovo istraživanje finansira Ministarstvo nauke, tehnološkog razvoja i inovacija Republike Srbije na osnovu ugovora br. 451-03-34/2026-03/200166, sklopljenog s Filozofskim fakultetom Univerziteta u Novom Sadu.

Language Resources

Jelena Redli. 2025. *Podkorpus Corpus-Mor.SR-13-Redli korpusa Corpus-Mor.SR* (Šošić, Stanković). JERTEH, https://docs.google.com/spreadsheets/d/1rGsxM0D58o6i-ChIVNz-pqn35FDQfkBKKCb_zRMhrzE/edit?usp=sharing, 1.0.

References

- Vuk Batanović, Miloš Cvetanović, and Boško Nikolić. 2020. *A Versatile Framework for Resource-limited Sentiment Articulation, Annotation, and Analysis of Short Texts*. *PLOS ONE*, 15(11):e0242050.
- Renjie Cao, Miaoyan Hu, Jiahao Wei, and Baha Ihnaini. 2024. *The Moral Foundations Weibo Corpus*. In *Proceedings of the 1st Workshop on NLP for Science (NLP4Science)*, pages 155–165. Association for Computational Linguistics.
- Jasmina Dražić. 2022. *Pristupi proučavanju značenja reči*. Filozofski fakultet, Novi Sad.
- Jesse Graham, Jonathan Haidt, and Brian A. Nosek. 2009. *Liberals and Conservatives Rely on Different Sets of Moral Foundations*. *Journal of Personality and Social Psychology*, 96(5):1029–1046.
- John Pavlopoulos, Jeffrey Sorensen, Lucas Dixon, Nithum Thain, and Ion Androutsopoulos. 2020. *Toxicity Detection: Does Context Really Matter?* In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4296–4305. Association for Computational Linguistics.
- Tvrtko Prčić. 2016. *Semantika i pragmatika reči*, treće elektronsko izdanje edition. Filozofski fakultet, Novi Sad.
- Maarten Sap, Dallas Card, Saadia Gabriel, Yejin Choi, and Noah A. Smith. 2019. *The Risk of Racial Bias in Hate Speech Detection*. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1668–1678. Association for Computational Linguistics.
- Jelena Sarajlić, Gaurish Thakkar, Diego Alves, and Nives Mikelić Pre-radović. 2021. *Quotations, Coreference Resolution, and Sentiment*

- Annotations in Croatian News Articles: An Exploratory Study. In *Proceedings of the Conference on Digital Curation Technologies (Qurator 2021)*, page 16, Berlin, Germany.
- Ljiljana Šarić. 2022. Impoliteness Strategies in Croatian and Serbian User Comments on Online News Articles: A Study Based on Readers' Perceptions. *Jezikoslovlje*, 23(1):1–34.
- Јована Б. Јовановић. 2021. Пејоративи у номинацији човека у српском језику. *Јужнословенски филолог*, 77(1):127–161.
- Стана Ристић. 1996. Типови експресивне лексике у савременом српском језику: модел „особа + психичка или морална особина”. *Јужнословенски филолог*, 52:57–78.

Moral Value Annotation in Serbian Online Discourse: Challenges and Perspectives

Jelena Redli

Summary

This paper examines linguistic and interpretative challenges in the annotation of moral values in Serbian online discourse. The analysis is based on a portion of the Corpus-Mor.SR-13 dataset, which contains annotated units from Twitter and Reddit, while the sports discourse subset is treated as a case study. The aim is not to provide a quantitative account of moral categories, but to show why moral value annotation cannot be reduced to explicit lexical markers or predefined keyword lists.

The theoretical framework combines Moral Foundations Theory with lexical-semantic, pragmatic and discourse-oriented approaches to meaning. The paper distinguishes moral evaluation from related forms of value-oriented analysis, especially sentiment analysis, emotion detection and the analysis of offensive or toxic online speech. While sentiment usually refers to the positive, negative or neutral orientation of an utterance, moral evaluation involves a more complex relation to actions, actors, social norms and values. This distinction is particularly important in short online texts, where context is often limited and evaluative meanings may remain implicit.

The sports discourse subset is analysed as a borderline domain in which neutral information, affective reaction and moral evaluation frequently overlap. The analysis focuses on several types of annotation problems: the boundary between the label *non-moral* and moral evaluation, the difference between sentiment and moral value, implicit moral evaluation, overlap between moral categories, and annotator disagreement.

The examples show that negative events such as defeat, injury or elimination may trigger an affective response without necessarily activating a moral category. Conversely, moral relevance may be established indirectly, through references to rules, institutional authority, suspected irregularity, injustice or collective loyalty. Annotator disagreement is therefore not treated only as a weakness of the annotation procedure, but also as linguistically relevant evidence of semantic and pragmatic instability.

The results indicate that Serbian annotation guidelines should more clearly distinguish unfavourable events, sentiment and moral evaluation. Although limited to sports discourse, the analysis is relevant for broader Serbian moral value annotation and for more reliable annotated resources.

Keywords: corpus annotation, moral values, online discourse, sports discourse, pragmatics, natural language processing, Serbian language

Works are licensed under the terms and conditions of the Creative Commons Attribution 4.0 International License.

<https://creativecommons.org/licenses/by/4.0/>

