

---

# Hierarchical Recognition of Moral Foundations in Serbian Literary Corpus *SrpELTeC*

---

Milena Šošić<sup>1</sup>[\[0000-0002-9949-9686\]](#), Ranka Stanković<sup>2</sup>[\[0000-0001-5123-6273\]](#)

<sup>1</sup> Language Resources and Technologies Society – JeRTeh, Serbia,  
[milena.sosic@jerteh.rs](mailto:milena.sosic@jerteh.rs)

<sup>2</sup>University of Belgrade, Faculty of Mining and Geology, Serbia,  
[ranka.stankovic@rgf.bg.ac.rs](mailto:ranka.stankovic@rgf.bg.ac.rs)

Scientific paper

DOI: 10.18485/judig.2026.2.ch6

## Abstract

This paper presents a novel approach for the hierarchical moral analysis based on the Moral Foundations Theory (MFT) in Serbian literary corpus *SrpELTeC*. The proposed method automatically identifies and quantifies five moral foundations together with their corresponding virtue and vice categories – care/harm, fairness/cheating, loyalty/betrayal, authority/subversion and purity/degradation, within different segments of Serbian novels. The computation was performed using a developed lexicon of moral words *MFD.SR*, adapted to the linguistic and cultural characteristics of Serbian language. A key methodological contribution of this work is the recognition of moral signals across multiple levels of a hierarchically structured corpus, which enables comparable moral analysis of texts of varying sizes and characteristics. In this way, the method produces both macro and micro level representations of moral dynamics within literary narratives, that allows fine-grained tracking of shifts in moral values across different corpus segments. Preliminary results indicate that Serbian novels first published in 1840-1920. time period, exhibit distinct and measurable moral signatures across authors and literary movements, which offers new insights into the ideological and cultural structure of Serbian literature from this period.

**Keywords:** moral foundations, sentiment, lexicon, literary corpus, Serbian language

## 1. Introduction

Recent developments in computational humanities and Natural Language Processing (NLP) have enabled increasingly sophisticated research of ideological, cultural and psychological structures embedded within literary discourse (Mihalcea et al., 2024; AD et al., 2025; Ampo and Amora, 2025). Among the most influential theoretical frameworks in this domain is Moral Foundations Theory (MFT), which conceptualizes moral reasoning through several universal intuitive dimensions, including care/harm, fairness/cheating, loyalty/betrayal, authority/subversion and purity/degradation (Graham et al., 2013). Although computational applications of MFT have achieved considerable success in the English-language corpora, their implementation in morphologically rich South Slavic literary traditions remains significantly underexplored.

This paper presents a hierarchical NLP approach for the analysis of moral foundations in the corpus of Serbian novels from the middle of nineteenth-century and beginning of twentieth century (Krstev, 2021; Trtovac et al., 2021). Literary works provide a valuable resource for examining moral structures, as they encode ethical conflicts, social norms and ideological perspectives through narrative and language. This study applies a moral lexicon *MFD.SR*, adapted to Serbian, to the analysis of Serbian literary corpus in order to investigate the distribution of moral foundations across different levels of textual organization. By combining lexicon-based moral values annotation with hierarchical aggregation of lexical scores, the study aims to provide a quantitative and interpretable method for analysing moral patterns in literary texts.

The remainder of this paper is organized as follows. Section 2 describes the datasets, lexical resources, annotation schemes and analytical procedures used to investigate moral foundations in Serbian literary texts. Particular attention is devoted to the construction and application of the *MFD.SR*, the aggregation of moral scores at different textual levels and the methods employed for corpus-based analysis. Section 3 presents the quantitative and qualitative findings, including the distribution of moral foundations across the corpus, the most representative paragraphs and observed patterns in moral language use. Section 4 interprets these findings in the broader context of moral foundation theory, literary studies and computational

text analysis, while also addressing methodological limitations and potential sources of bias. Finally, Section 5 summarizes the main contributions of the study and outlines directions for future work, including the refinement of moral lexicons, evaluation of annotation quality and the application of large language models and knowledge-based approaches to moral language analysis.

## 2. Methodology

The *SrpELTeC*<sup>1</sup> corpus (Krstev et al., 2023) used in this study consists of Serbian literary novels first published in 1840-1920. time period, spanning multiple authors and narrative styles (Krstev, 2021). The corpus was developed within the COST Action *CA16204: Distant Reading for European Literary History*<sup>2</sup> (2018–2022) and formatted according to the TEI (Text Encoding Initiative) guidelines<sup>3</sup>. *SrpELTeC* is part of the European Literary Text Collection *ELTeC*, that comprises a collection of corpora of literary texts that are comparable in nature, scope and quality across several European languages.

The Level-1 TEI annotation of *ELTeC*<sup>4</sup> provides a standardized structural encoding of literary texts, including chapters, divisions, paragraphs, headings and basic bibliographic metadata, thereby ensuring interoperability and facilitating corpus-based research. With this format, in addition to the complete texts, the novels from corpus are hierarchically structured into multiple textual units, including chapters and paragraphs. In the *ELTeC*, Level-2<sup>5</sup> annotation extends the basic TEI encoding with automatically generated linguistic information, including tokenization, sentence segmentation, lemmatization, morphosyntactic tagging and named entity recognition, thereby supporting textometric, corpus-linguistic and semantic analyses (Stanković et al., 2024). Figure 1 presents a sentence (as a part of a paragraph) with annotated tokens, which enables multi-level analysis of lexical patterns across different segments of literary discourse.

---

<sup>1</sup> <https://github.com/COST-ELTeC/ELTeC-srp>

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

<sup>3</sup> <https://tei-c.org/>

<sup>4</sup> Encoding Guidelines for the *ELTeC*: level 1

<https://distantreading.github.io/Schema/eltec-1.html>

<sup>5</sup> Encoding Guidelines for the *ELTeC*: level 2

<https://distantreading.github.io/Schema/eltec-2.html>

```

<p xml:id="p_SRP18520_11">
  <s xml:id="s_SRP18520_13">
    <w pos="ADP" lemma="y" xml:id="w_SRP18520_174">Y</w>
    <w pos="NOUN" lemma="војводство" xml:id="w_SRP18520_175">Војводству</w>
    <w pos="VERB" lemma="имати" xml:id="w_SRP18520_176">има</w>
    <w pos="ADJ" lemma="бео" xml:id="w_SRP18520_177">белих</w>
    <w pos="NOUN" lemma="град" xml:id="w_SRP18520_178" join="right">градова</w>
    <pc pos="PUNCT" xml:id="w_SRP18520_179">,</pc>
    <w pos="ADJ" lemma="лен" xml:id="w_SRP18520_180">лепих</w>
    <w pos="CCONJ" lemma="и" xml:id="w_SRP18520_181">и</w>
    <w pos="ADJ" lemma="убав" xml:id="w_SRP18520_182">убавих</w>
    <w pos="NOUN" lemma="село" xml:id="w_SRP18520_183" join="right">села</w>
    <pc pos="PUNCT" xml:id="w_SRP18520_184">.</pc>
  </s>
</p>

```

Figure 1: The hierarchical organization of paragraphs, sentences and words available in the *SrpELTeC* corpus distribution level-2.

Tokenization and sentence segmentation were performed using the *Unitex*<sup>6</sup> corpus processing suite using language resources for Serbian (Krstev, 2008). Morphosyntactic annotation and lemmatization were carried out with a *TreeTagger* model Krstev et al. (2021) (Stanković et al., 2022). Named entity recognition was subsequently applied using *SrpNER*, a Serbian system for Named Entity Recognition (NER), implemented within the *Unitex* environment (<http://ner.jerteh.rs/>). To ensure compatibility with existing literary corpora and annotation standards, the *SrpNER* output was further simplified and mapped to the *ELTeC* NER tagset (Nešić et al., 2022) following the Annotation Guidelines for *ELTeC*.

Overall frequency statistics for each segment type is presented in Table 1 providing an overview of the structural composition of the *SrpELTeC* corpus (Krstev et al., 2023). The corpus contains more than 4.9 million word tokens distributed across 368,156 sentences and 149,402 paragraphs, making it one of the largest structurally annotated collections of Serbian literary texts currently available. The average novel contains approximately 49,000 words, 3,700 sentences and 1,500 paragraphs, indicating substantial textual diversity and sufficient volume for quantitative corpus analyses.

The TEI-based encoding captures multiple structural levels of literary texts, including 2,047 chapters and 20,831 page boundaries, enabling research not only at the lexical and sentence levels but also at higher discourse and document levels. In addition, the corpus

<sup>6</sup> <https://unitexgramlab.org/>

contains 2,873 verse segments, reflecting the presence of poetic paragraphs embedded within prose works, as well as 787 foreign-language segments and 420 references to literary works. These specialized annotations provide valuable contextual information for literary, linguistic and cultural studies. Overall, the statistics demonstrate both the richness and the complexity of the *SrpELTeC* corpus, making it a suitable resource for textometric analyses, natural language processing tasks and investigations of moral, emotional and semantic phenomena in Serbian literary texts.

Table 1: Frequency statistics of segment type appearance in the *SrpELTeC* corpus (100 novels).

| Segment type     | Frequency |
|------------------|-----------|
| Novels           | 100       |
| Chapters         | 2,047     |
| Pages            | 20,831    |
| Paragraphs       | 149,402   |
| Sentences        | 368,156   |
| Punctuations     | 1,145,701 |
| Words            | 4,931,518 |
| Verses           | 2,873     |
| Foreign segments | 787       |
| Mentioned works  | 420       |

A central contribution of this research is the application of a Serbian moral foundation lexicon *MFD.SR* (Šošić, 2026) tailored to Serbian linguistic and cultural characteristics (Šošić, 2025) to uncover moral values in the old Serbian literature. The lexicon-based approach represents one of the established methods for computationally identifying and quantifying moral language in textual data, complementing supervised and neural modeling approaches (Reinig et al., 2024; Asprino et al., 2022). Existing English-language moral foundation dictionaries (Graham et al., 2009; Hopp et al., 2021; Araque et al., 2020) cannot be directly transferred into Serbian due to substantial grammatical, semantic and cultural differences. Developing and aligning a moral lexicon to the target language and its cultural context is therefore essential for reliable lexicon-based moral analysis (Cheng and Zhang, 2023; Zúquete et al., 2023; Rizzoli et al., 2026). The moral lexicon used in this research was generated automatically

from a manually annotated textual corpus in Serbian labeled according to MFT categories. Lexicon construction combined several stages of linguistic and statistical processing, including text normalization, removal of named entities, lemmatization relying on resources developed by Krstev (2008) and sentiment estimation using the SRPOL tool (Šošić, Milena, 2024), described in (Šošić, 2022).

To identify lexical units most strongly associated with individual moral categories, the methodology for lexicon development employed the class-based TF-IDF (c-TF-IDF) algorithm, which measures the discriminative importance of a term within a specific class relative to the complete corpus (Grootendorst, 2022). Unlike standard TF-IDF, where documents are treated individually, c-TF-IDF aggregates all texts belonging to the same moral category into a single class-level document, enabling extraction of vocabulary characteristic of each moral foundation. This approach enables the extraction of morally discriminative vocabulary while reducing the influence of overly frequent or semantically generic words. Combined with sentiment estimation and linguistic normalization, the resulting lexicon captures morally salient words across morphologically rich and stylistically heterogeneous Serbian textual data.

```
"MFD.SR": [  
  {  
    "lemma": "agresor",  
    "PoS": "NOUN",  
    "weights": {  
      "authority": 0.0508,  
      "care": 0.3301,  
      "fairness": 0.1545,  
      "loyalty": 0.3956,  
      "purity": 0.0689  
    },  
    "sentiment": -0.55,  
    "moral_value": "loyalty",  
    "moral_sentiment": "loyalty.vice",  
    "moral_category": "betrayal"  
  },  
]
```

Figure 2: A lexical entry example from the lexicon *MFD.SR*.

The lexicon was therefore developed through a hybrid methodology combining manual expert annotation with automatic corpus-

driven lexical derivation. Initial seed vocabularies derived from this method were normalized and expanded through synonymic clustering, morphological derivation and frequency-based semantic analysis of the annotated corpora. In the final instance, lexicon *MFD.SR* contains approximately 4,300 words, where each lexical entry is associated with a normalized distribution of moral weights over five moral dimensions. These weights represent the relative association of a lemma with each moral foundation and are constrained to sum to one, ensuring comparability across the entries. In addition to moral weight distributions, lexicon entries include auxiliary annotations such as categorical moral labels and sentiment-related category splitting (e.g. *virtue* vs. *vice*).

The lexical entry “*agresor*” (eng. “aggressor”) from Figure 2 illustrates the structure of the *MFD.SR* and the distinction between sentiment polarity and moral dimensions. The lemma is associated with a strongly negative sentiment score (-0.55), reflecting its generally unfavorable evaluation in language use. At the same time, the highest moral weight is assigned to the *loyalty* foundation (0.3956), followed by *care* (0.3301). Importantly, these weights do not indicate positive moral values but rather the moral foundation to which the concept is most closely related. In MFT, foundations are represented by both *virtue* and *vice* poles. Consequently, the lemma “*agresor*” is linked to the category *betrayal* and the moral sentiment label *loyalty.vice*, indicating a violation of loyalty-related norms rather than the presence of *loyalty* as a *virtue*. Similarly, the association with the *care* foundation reflects the concept of harm and victimization, corresponding to the negative pole of the *care/harm* dimension. This example demonstrates that moral annotation captures the underlying moral frame evoked by a lexical item, while sentiment analysis captures its overall positive or negative evaluation (*virtue* or *vice*). The two dimensions are therefore complementary rather than redundant, providing a richer representation of moral meaning.

Moral annotation is performed through direct matching between corpus lemmas and lexicon entries. Since both resources share consistent lemma and Part-of-Speech (PoS) annotation schemes, no additional linguistic pre-processing is required. Each matched token inherits its corresponding moral weight vector from the lexicon. Tokens not present in the lexicon were excluded from the moral analysis. To capture moral patterns at different levels of textual structure, a

hierarchical aggregation approach is applied. Token values are progressively aggregated across increasingly larger textual units:

- **Sentence** level – Moral values are aggregated across tokens within each sentence to capture local moral expressions.
- **Paragraph** level – Sentence-level values are further aggregated to represent thematic and discourse-level moral structure.
- **Chapter** level – Paragraph level values are aggregated to represent the moral profile for each chapter.
- **Literary Work** level – Chapter-level values are aggregated to obtain a moral profile for each individual literary work.
- **Corpus** level – Work-level representations are aggregated to obtain an overall moral profile of the entire corpus.

This hierarchical design allows analysis of moral patterns at multiple levels of textual granularity, from local expressions to global narrative tendencies. After each aggregation step, values are additionally normalized to ensure numerical consistency and interpretability. This approach guarantees that all moral representations remain comparable across different textual units and hierarchical levels. The multi-layered architecture preserves local rhetorical variation while enabling direct comparison between texts differing greatly in size, narrative structure and stylistic density. For interpretability, each aggregated moral vector can be mapped to a dominant moral foundation based on the highest weight value. In cases where multiple foundations share the maximum value, multiple labels are retained to reflect ambiguity.

The analytical pipeline produces two complementary forms of representation. At the macro level, each literary work is transformed into a normalized multidimensional moral profile suitable for comparative quantitative analysis across authors, genres and historical periods. At the micro level, the framework preserves hierarchical moral annotations across sentences and paragraphs, with fine-grained analysis of intra-textual moral dynamics, which enable researchers to trace narrative transitions and shifting ethical emphases throughout the text. This approach makes it possible to identify how moral rhetoric evolves across narrative sequences, character interactions and descriptive paragraphs rather than treating novels as semantically uniform entities. Such granularity is particularly valuable in literary analysis, where ideological meaning frequently emerges through gradual rhetorical development rather than isolated lexical patterns.

Based on the token moral assignments, moral density is computed as the proportion of lexicon-matched tokens within a given textual segment:

$$MD(S) = 100 \cdot \frac{|M|}{|S|} \quad (1)$$

where  $S$  denotes a textual segment (sentence, paragraph, chapter or literary work),  $|S|$  is the total number of tokens in the segment and  $M$  is the subset of tokens matched in the moral lexicon. This measure enables the identification of textual units with concentrated moral language at the sentence, paragraph, chapter and literary work levels. In parallel, moral foundation scores are obtained by averaging token-level weight vectors within each segment:

$$F_k(S) = \frac{1}{|M|} \sum_{t_i \in M} w_k(t_i) \quad (2)$$

where  $w_k(t_i)$  denotes the lexical weight assigned to the  $k$ -th moral foundation for token  $t_i$  and  $k \in \{\text{authority, care, fairness, loyalty, purity}\}$  set of moral foundations. This aggregation provides a continuous representation of the relative prominence of the five moral foundations across the narrative flow. Finally, the foundation scores are normalized to obtain a probability distributions over the five moral foundations:

$$\hat{F}_k(S) = \frac{F_k(S)}{\sum_{j=1}^5 F_j(S)}, \quad \sum_{k=1}^5 \hat{F}_k(S) = 1 \quad (3)$$

In addition, for each moral sentiment  $ms \in \{\text{care.virtue, care.vice, \dots, purity.vice}\}$ , its relative proportion within a textual segment is calculated as

$$P_{ms}(S) = \frac{|M_{ms}|}{|M|}, \quad (4)$$

where  $|M_{ms}|$  denotes the number of tokens assigned to moral sentiment  $ms$  and  $|M|$  is the total number of moral tokens in the segment  $S$ . To rank textual units while balancing relative concentration and supporting frequency, a weighted moral sentiment score is defined as

$$W_{ms}(S) = P_{ms}(S) \ln(1 + |M_{ms}|). \quad (5)$$

### 3. Results

#### 3.1. Moral Foundation

To capture both micro and macro structures, the analysis is performed hierarchically across sentences, paragraphs, chapters, individual literary works and the entire corpus. In the analysis, segments are ranked according to a combination of moral density and aggregated moral scores, which allows identification of the most morally salient paragraphs (Table 2).

Table 2: Top 5 paragraphs ordered by aggregated moral scores for the moral foundation `loyalty`.

| Literary Work                                  | Paragraph ID | $\hat{F}_k$ |
|------------------------------------------------|--------------|-------------|
| “ <i>Patnica</i> ” (Jakov Ignjatović, 1862)    | 3571         | 0.3022      |
| “ <i>1915</i> ” (Branislav Nušić, 1921)        | 881          | 0.2593      |
| “ <i>Prvo veselje</i> ” (Milovan Glišić, 1888) | 1742         | 0.2572      |
| “ <i>Pauci</i> ” (Ivo Ćipiko, 1909)            | 612          | 0.2520      |
| “ <i>Prve žrtve</i> ” (Andra Gavrilović, 1893) | 834          | 0.2506      |

Within each paragraph, sentence-level analysis is used to detect localized peaks of moral intensity by identifying the sentence with the highest concentration of moral lexical items. Together, these procedures enable a multi-level and fine-grained examination of moral patterns in the *SrpELTeC* literary corpus. The results presented in Table 3 reveal the exact parts of the novel “*Kaluđer i hajduk*” (Stojan Novaković, 1913) where moral language is most concentrated. By using a composite score calculated as the product of the paragraph’s moral word density and the total number of detected moral words, the algorithm successfully selected paragraphs with the highest value for moral charge. Resulting paragraphs shows a strong and balanced presence of multiple moral dimensions, especially `authority`, `care` and `purity`. The detected moral lemmas, such as “*car*”, “*bog*”, “*krvav*” and “*knjiga*” indicate that the highest moral intensity in the novel is connected to themes of historical conflict, religion and social order. Furthermore, the analysis of individual sentences shows an stylistic pattern in which the author typically builds a highly moral paragraph around one short, high-impact core sentence. For example, the sentence “*Boj se otvori strašan i krvav.*” has a moral density of

Table 3: Top 3 paragraph with the highest moral charge in the containing sentences for the literary work: *“Kaluder i hajduk”* (average moral density of 28.78%).

|                         |                                                                                                                          |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------|
| <b>1. Paragraph ID:</b> | p_SRP19130_1122                                                                                                          |
| <b>MD:</b>              | 32.43 (155/478 words)                                                                                                    |
| <b>Moral Weights:</b>   | {authority: 0.2062, care: 0.2047, fairness: 0.1906, loyalty: 0.2045, purity: 0.194}                                      |
| <b>Top Sentence ID:</b> | s_SRP19130_3829 (MD=66.67)                                                                                               |
| <b>Sentence Text:</b>   | <i>“Boj se otvori strašan i krvav.”</i>                                                                                  |
| <b>Moral Lemmas:</b>    | <i>“bojiti”, “otvoriti”, “strašan”, “krvav”</i>                                                                          |
| <b>2. Paragraph ID:</b> | p_SRP19130_384                                                                                                           |
| <b>MD:</b>              | 29.79 (143/480 words)                                                                                                    |
| <b>Moral Weights:</b>   | {authority: 0.1896, care: 0.2106, fairness: 0.1975, loyalty: 0.1978, purity: 0.2046}                                     |
| <b>Top Sentence ID:</b> | s_SRP19130_1342 (MD=52.94)                                                                                               |
| <b>Sentence Text:</b>   | <i>“I ako ih je najveći deo radio zemlju, bilo ih je svuda određenih za poslove drugoga reda.”</i>                       |
| <b>Moral Lemmas:</b>    | <i>“velik”, “deo”, “raditi”, “zemlja”, “posao”, “drugi”, “red”</i>                                                       |
| <b>3. Paragraph ID:</b> | p_SRP19130_568                                                                                                           |
| <b>MD:</b>              | 33.24 (123/370 words)                                                                                                    |
| <b>Moral Weights:</b>   | {authority: 0.1691, care: 0.2323, fairness: 0.1816, loyalty: 0.2095, purity: 0.2074}                                     |
| <b>Top Sentence ID:</b> | s_SRP19130_1968 (MD=50.0)                                                                                                |
| <b>Sentence Text:</b>   | <i>“Sav istorički deo Svetoga Pisma staroga zaveta, knjige o carevima i o jevrejskoj istoriji znao je u potpunosti.”</i> |
| <b>Moral Lemmas:</b>    | <i>“deo”, “svet”, “pisma”, “star”, “knjiga”, “car”, “istorija”, “znati”, “potpunost”</i>                                 |

66.67. These short sentences act as the emotional and ethical core of the paragraph, shaping the tone for the entire surrounding text.

The normalized  $\hat{F}_k$  scores (Eq. 3) provide a continuous representation of the relative contribution of each moral foundation within a textual segment and facilitate comparisons across different segments regardless of their length. In addition, since the same aggregation procedure is applied recursively at the sentence, paragraph, chapter and literary work levels, the proposed framework produces a hierarchical representation of moral density (Eq. 1) distributions for each moral

foundation throughout the narrative structure of the literary work (Figure 3).

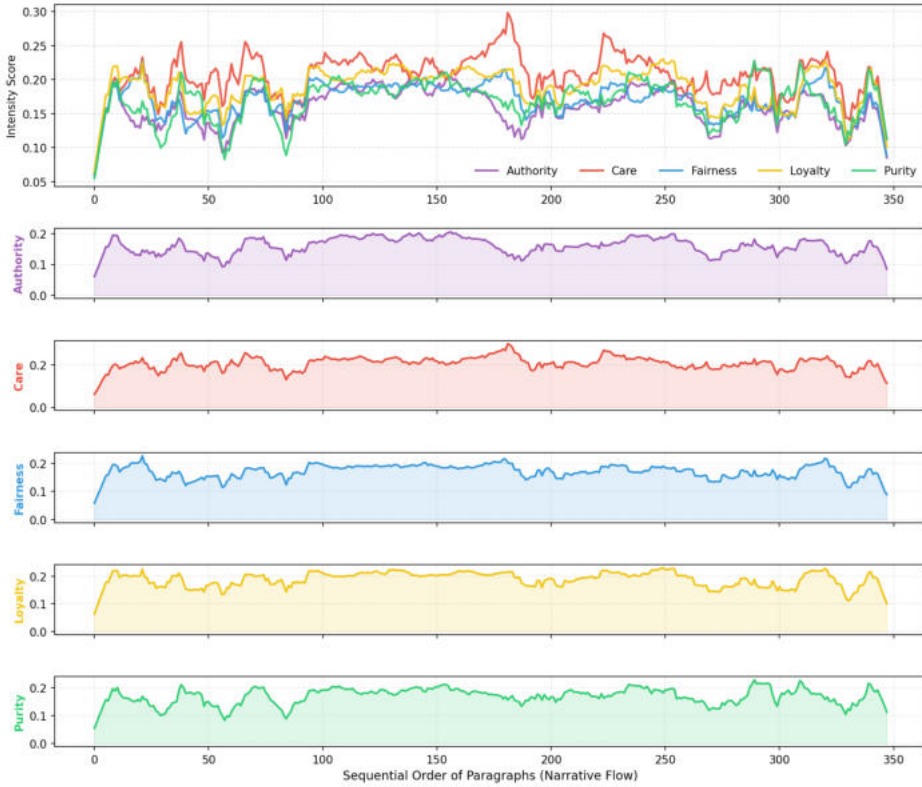


Figure 3: Comparative overview of moral foundations dynamics in the novel “*Glava šecera*” (Milovan Glišić, 1875), smoothed using a 10-paragraph sliding window moving average of moral density through the narrative flow.

The obtained results on individual literary works level indicate that distinct authors and literary movements exhibit recognizable moral signatures and narrative moral dynamics. Certain novels foreground themes associated with **authority** and **loyalty** through representations of patriarchal order, collective identity and national belonging, while others emphasize **care** and **fairness** through depictions of social injustice, suffering and interpersonal morality. Based on the individual literary works, the moral foundation **care** has been identified as a predominant for the entire *SrpELTeC* corpus.

To improve interpretability and support interdisciplinary literary analysis, the framework incorporates an interactive visualization

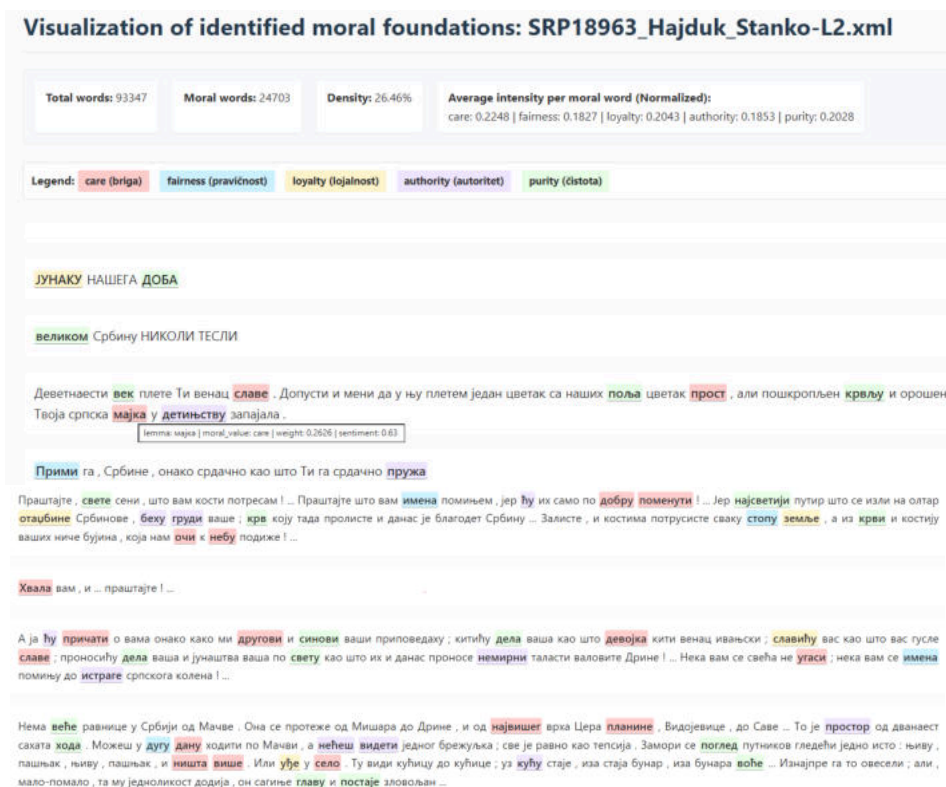


Figure 4: Visualisation of the detected moral words in the literary work “Hajduk Stanko” (Janko Veselinović, 1896).

interface in which morally classified lexical units are color-coded according to their associated moral foundations, alongside aggregated statistics of detected moral words (Figure 4). Incorporated metadata layer, available through the tooltips, provide immediate access to the lemma form, the most dominant moral weight with derived moral category and associated lexical sentiment for all moral words identified in the text. This functionality enhances transparency by enabling direct qualitative validation of computationally derived moral annotations within their original narrative context.

### 3.2. Moral Sentiment

To complement the moral foundations analysis based on the moral words scoring, we additionally examined moral sentiment, which distinguishes whether a moral foundation is expressed through its virtue or vice dimensions. While the stan-

dard analysis identifies the prevalence of the five moral foundations, moral sentiment provides a fine-grained characterization of their narrative realization across sentiment poles. For each literary work, we computed the distribution of all ten moral sentiments (care.virtue, care.vice, . . . , purity.vice) and ranked literary works according to the relative proportion (Eq. 4) of each moral sentiment in order to identify the literary texts that most strongly exemplify a particular moral framing. This analysis complements the local sentence and paragraph level investigation by identifying representative works for each moral sentiment across the entire corpus.

Table 4: Top-ranked literary work for each moral sentiment. The ranking is based on the relative share of the corresponding moral sentiment within each work.

| Moral Sentiment                 | Literary Work                                          | $P_{ms}$ | $MD$  |
|---------------------------------|--------------------------------------------------------|----------|-------|
| authority.vice<br>(subversion)  | <i>“Pastir kralj”</i><br>(Kosta Brančić, 1909)         | 5.84     | 25.43 |
| authority.virtue<br>(authority) | <i>“Stradija”</i><br>(Radoje Domanović, 1902)          | 17.01    | 32.97 |
| care.vice<br>(harm)             | <i>“Đul Marika”</i><br>(Jelena Dimitrijević, 1903)     | 16.08    | 12.93 |
| care.virtue<br>(care)           | <i>“Jedna ženidba”</i><br>(Milan Savić, 1898)          | 41.48    | 23.45 |
| fairness.vice<br>(cheating)     | <i>“Zločin svekrve”</i><br>(Kosta Jovanović, 1912)     | 5.79     | 26.52 |
| fairness.virtue<br>(fairness)   | <i>“Zločin svekrve”</i><br>(Kosta Jovanović, 1912)     | 10.56    | 26.52 |
| loyalty.vice<br>(betrayal)      | <i>“Neimari sa krajem”</i><br>(Kosta Petrović, 1911)   | 5.01     | 26.21 |
| loyalty.virtue<br>(loyalty)     | <i>“Pastir kralj”</i><br>(Kosta Brančić, 1909)         | 9.74     | 25.43 |
| purity.vice<br>(degradation)    | <i>“Sinovac”</i><br>(Dimitrije Radović, 1905)          | 8.79     | 24.54 |
| purity.virtue<br>(purity)       | <i>“Ljubav i patriotizam”</i><br>(Jovan Jeremić, 1879) | 22.99    | 25.13 |

Table 4 presents the highest-ranked literary work for each moral sentiment. The results indicate that different works exhibit distinct moral profiles despite the overall predominance of the

Table 5: Highest-ranked representative sentences for virtue-oriented moral sentiments.

| Moral Sentiment                 | Literary Work                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | $W_{ms}$ |
|---------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| authority.virtue<br>(authority) | “ <i>Stradija</i> ”<br>(Radoje Domanović, 1902)<br>“ <i>Eto vidite, danas mi je ministar prosvete poslao jedan svoj raspis, da ga i ja potpomognem i naredim svima policijskim organima u područnom mi ministarstvu da se strogo pridržavaju toga raspisa ministra prosvete.</i> ”                                                                                                                                                                                                                                                                                           | 1.82     |
| care.virtue<br>(care)           | “ <i>Radiša</i> ”<br>(Janko Veselinović, 1894)<br>“ <i>Već sam vam kazao, da ste vi došli u školu da naučite da čitate knjige, da pišete, da računate i da naučite još mnogo što šta što u knjigama ima, a prosti ljudi ne znaju to zato, što ne umeju da čitaju.</i> ”                                                                                                                                                                                                                                                                                                      | 2.64     |
| fairness.virtue<br>(fairness)   | “ <i>Dragocena ogrlica</i> ”<br>(Milutin Uskoković, 1910)<br>“ <i>Ono je pitanje bilo i prošlo, i — predlog bi se primio — sedanjem i ustajanjem.</i> ”                                                                                                                                                                                                                                                                                                                                                                                                                      | 1.61     |
| loyalty.virtue<br>(loyalty)     | “ <i>Đurađ Branković</i> ”<br>(Jakov Ignjatović, 1855)<br>“ <i>Bajazit, da bi Marka i njegove privrženike utaložio, paštio se uveriti ji, da on to zato čini da posle oborene Brankovićeve i Lazareve porodice svojim prijateljima srpsku zemlju preda.</i> ”                                                                                                                                                                                                                                                                                                                | 1.79     |
| purity.virtue<br>(purity)       | “ <i>Radiša</i> ”<br>(Janko Veselinović, 1894)<br>“ <i>Sveti Stevane, sveto Bogojavljenje, sveti Jovane Krstitelju, sveti Savo učitelju, sveto Sretenije, sveti Arsenije i Maksime, sveto Voskresenije, sveti Đurđe, sveta Trojice živa pomoćnice, sveti Vrtolomeja, sveti Petre i Pavle, sveti Vrači, sveto Prokoplje, sveta Ognjena Marija, sveta Mando i Mandalena, sveta Ano, sveta Paraskevo, sveti Panteleja, sveto Preobraženje, sveta Bogorodice, sveti Jovane, sveti Dimitrije, sveti Arandele, sveti Nikola i svi sveci i blagi danci, da ste nam na pomoći!</i> ” | 3.02     |

care foundation and virtue sentiment dimension in the corpus (e.g. care.virtue). For example, literary work “*Jedna Ženidba*” exhibits the highest proportion of care.virtue (41.48%), whereas “*Đul Marika*” contains the highest proportion of care.vice (16.08%). Similarly, literary “*Stradija*” is characterized by the strongest authority.virtue dimension (17.01%), while “*Pastir Kralj*” exhibits the highest proportion of authority.vice (5.84%). The results also

demonstrate that moral sentiment analysis captures qualitative differences in the moral sentiment framing of the literary works as predominant **virtue** over **vice** dimension for each foundation.

Table 6: Highest-ranked representative sentences for **vice**-oriented moral sentiments.

| Moral Sentiment                | Literary Work                                                                                                                                       | $W_{ms}$ |
|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| authority.vice<br>(subversion) | <i>“Pastir kralj”</i><br>(Kosta Brančić, 1907)<br><i>“Sila je sila, jes, sila je sila!...”</i>                                                      | 1.61     |
| care.vice<br>(harm)            | <i>“Đul Marika”</i><br>(Jelena Dimitrijević, 1903)<br><i>“Vatra je na ognjištu gorela i dogorela pa se, malo po malo, pokrila debelim puhorom.”</i> | 1.79     |
| fairness.vice<br>(cheating)    | <i>“Zločin svekrve”</i><br>(Kosta J. Trifković, 1895)<br><i>“Ja predlažem da se kazni, a i kaznu sam izmislio.”</i>                                 | 1.39     |
| loyalty.vice<br>(betrayal)     | <i>“Milan Narandžić”</i><br>(Jakov Ignjatović, 1860)<br><i>“Ti tvoje isteruj, a mene ostavi, pa kud puklo da puklo.”</i>                            | 1.39     |
| purity.vice<br>(degradation)   | <i>“Pred zoru”</i><br>(Veljko Milićević, 1912)<br><i>“Neće biti gospodara i robova, rasipanja i gladi; nestaće nasilja i patnje.”</i>               | 1.95     |

In addition to the literary work level analysis, we identified the highest-ranked sentence for each of the five moral foundations across **virtue** and **vice** dimensions. Sentences were ranked using a weighted score that combines the relative proportion of the target sentiment among all moral words in the sentence with the number of its occurrences (Eq. 5). This approach favors sentences that are both strongly concentrated around a particular moral sentiment and supported by multiple relevant lexical items, thereby reducing the influence of very short sentences containing only one or two moral words. Tables 5 and 6 present the highest-ranked representative sentences oriented to **virtue** and **vice** moral sentiments, respectively. The examples illustrate how the **virtue** and **vice** dimensions of the five moral foundations are realized in concrete textual contexts. The **virtue**-oriented sentences frequently foreground education, institutional order, collective affiliation, justice and religious devotion (Table 5), whereas **vice**-

oriented examples evoke harm, deception, betrayal, coercive power and degradation (Table 6). These sentence examples provide an interpretable qualitative complement to the quantitative distributions obtained at the literary work and corpus levels. The selected examples should be interpreted as lexicon-based representative sentences rather than manually validated prototypes of moral sentiment, since the ranking is determined by the aggregated moral values of individual lexical items rather than the overall contextual meaning of the sentence.

#### 4. Discussion

The results of this study demonstrate that moral dimensions in Serbian literary texts can be modeled using a lexicon-based approach grounded in MFT. Since no formal evaluation has yet been conducted, this pioneering effort can be seen as a proof of concept that opens new research directions. While the initial results are promising, a thorough evaluation will be required to assess its performance and reliability in future studies. By applying a structured moral lexicon consisting of the five primary moral dimensions, it is possible to quantify the distribution of moral content across lexical items and textual units. The analysis shows that aggregating moral scores at the lemma level improves stability and reduces noise introduced by individual annotations and contextual variation. Normalization of moral weights ensures comparability across lexical entries by constraining values to a unified scale. However, this procedure may also reduce sensitivity to differences in absolute moral intensity, which should be considered when interpreting results.

The findings indicate that certain moral foundations tend to dominate in specific textual contexts. In particular, *care* and *loyalty* frequently emerge as prominent dimensions, while *fairness*, *authority* and *purity* exhibit more context-dependent behaviour, often reflecting narrative structure, character relations or ideological framing within the text (Figure 3). These patterns suggest that moral language usage is not uniformly distributed, but instead closely tied to thematic and stylistic properties of literary works. A key limitation of this approach lies in its reliance on static lexical resources, which cannot fully capture contextual dependencies, irony or shifts in meaning across different narrative styles. Additionally, lexical polysemy may introduce uncertainty in the assignment of moral values at

the lemma level. The lexicon also does not include multi-word expressions (MWEs), even though moral meaning in texts is frequently realized through fixed phrases, idiomatic constructions and collocations rather than isolated lemmas. Furthermore, the lexicon has limited coverage of archaic and historical language forms, which are characteristic of older Serbian literary texts in the *SrpELTeC* corpus. This may lead to reduced recall for morally salient expressions expressed through outdated or non-standard vocabulary, thereby affecting the completeness of the moral annotation in diachronic literary data.

## 5. Conclusion

This study demonstrates that lexicon-based methods provide an effective and interpretable framework for analyzing moral structures in Serbian literary texts. By leveraging a moral foundations lexicon, it is possible to obtain a quantitative representation of moral content and compare its distribution across different literary works and authors. The application of normalized moral weights ensures consistency and comparability across lexical units, while aggregation at the lemma level enhances robustness and reduces annotation noise.

Future work will focus on extending the proposed hierarchical framework to additional structural units of literary texts, including verses and other narrative components, which would enable more detailed analysis of moral dynamics across different levels of textual organization. Next, additional validation and expansion of the moral lexicon are necessary in order to improve lexical coverage, particularly for multi-word expressions, archaic vocabulary, historically specific expressions and morally relevant lexical units not currently represented in the resource. Furthermore, future research may explore the integration of contextual language models and hybrid approaches that combine lexicon-based methods with machine learning techniques to improve the detection of implicit and context-dependent moral expressions. Finally, expanding the methodology to other literary periods, authors, genres may provide broader insights into the evolution of moral discourse in Serbian literature.

The use of moral lexicons to identify moral signals in literary corpora represents a promising direction for computational literary analysis, that enables systematic exploration of ethical and ideological patterns in textual data while maintaining interpretability and linguistic grounding.

## Acknowledgements

This research was supported by the Science Fund of the Republic of Serbia, #7276, Text Embeddings - Serbian Language Applications - TESLA <https://tesla.rgf.bg.ac.rs/>.

## Language Resources

Krstev, Cvetana and Stanković, Ranka and Trtovac, Aleksandra and Milnović, Vasilije Vitas, Duško and Obradović, Ivan. 2023. *The Serbian novel collection for the ELTeC: The European Literary Text Collection*. European Language Grid (ELG). Available through GitHub and the European Language Grid: <https://live.european-language-grid.eu/catalogue/corpus/23621>.

Krstev, Cvetana and Vitas, Duško and Stanković, Ranka and Škorić, Mihailo. 2021. *SrpMD4Tagging - Serbian Morphological Dictionaries for Tagging*. ELG, <https://live.european-language-grid.eu/catalogue/lcr/9294>, 1.0.

Šošić, Milena. 2026. *MFD.SR – Lexicon of Serbian Moral Words*. European Language Grid (ELG). Forthcoming language resource.

Šošić, Milena. 2024. *SentiWords.SR – A Lexicon of Sentiment-Scored Serbian Words*. European Language Grid (ELG).

## References

Mahit Nandan AD, Ishan Godbole, Pranav M Kapparad, and Shrutilipi Bhattacharjee. 2025. *Comparative Analysis of Religious Texts: NLP Approaches to the Bible, Quran, and Bhagavad Gita*. In *Proceedings of the new horizons in computational linguistics for religious texts*, pages 1–10.

Weena Mae G Ampo and Epifelward Niño Amora. 2025. *Exploring Sentiment and Moral Dynamics in Children’s Literature Using Machine Learning: A Sentiment Analysis Approach*. In *2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL)*, pages 199–208. IEEE.

Oscar Araque, Lorenzo Gatti, and Kyriaki Kalimeri. 2020. *Moral-Strength: Exploiting a Moral Lexicon and Embedding Similar-*

- ity for Moral Foundations Prediction. *Knowledge-based systems*, 191:105184.
- Luigi Asprino, Luana Bulla, Stefano De Giorgis, Aldo Gangemi, Ludovica Marinucci, et al. 2022. [Uncovering Values: Detecting Latent Moral Content from Natural Language with Explainable and Non-trained Methods](#). In *Proceedings of deep learning inside out (DeeLIO 2022): The 3rd workshop on knowledge extraction and integration for deep learning architectures*, pages 33–41.
- Calvin Yixiang Cheng and Weiyu Zhang. 2023. [C-mfd 2.0: Developing a Chinese Moral Foundation Dictionary](#). *Computational Communication Research*, 5(2):1.
- Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P Wojcik, and Peter H Ditto. 2013. [Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism](#). In *Advances in experimental social psychology*, volume 47, pages 55–130. Elsevier.
- 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.
- Maarten Grootendorst. 2022. [BERTopic: Neural Topic Modeling with a Class-based TF-IDF Procedure](#). *arXiv preprint arXiv:2203.05794*.
- Frederic R Hopp, Jacob T Fisher, Devin Cornell, Richard Huskey, and René Weber. 2021. [The extended Moral Foundations Dictionary \(eMFD\): Development and applications of a crowd-sourced approach to extracting moral intuitions from text](#). *Behavior research methods*, 53(1):232–246.
- Cvetana Krstev. 2008. *Processing of Serbian – Automata, Texts and Electronic Dictionaries*. University of Belgrade, Faculty of Philology, Belgrade.
- Cvetana Krstev. 2021. [The Serbian Part of the ELTeC Collection through the Magnifying Glass of Metadata](#). *Infotheca-Journal for Digital Humanities*, 21(2):26–42.
- Rada Mihalcea, Laura Biester, Ryan L Boyd, Zhijing Jin, Veronica Perez-Rosas, Steven Wilson, and James W Pennebaker. 2024. [How Developments in Natural Language Processing Help us in Understanding Human Behaviour](#). *Nature Human Behaviour*, 8(10):1877–1889.

- Milica Ikonić Nešić, Ranka Stanković, Christof Schöch, and Mihailo Škorić. 2022. [From ELTeC Text Collection Metadata and Named Entities to Linked-data \(and Back\)](#). In *Proceedings of the 8th Workshop on Linked Data in Linguistics within the 13th Language Resources and Evaluation Conference*, pages 7–16.
- Ines Reinig, Maria Becker, Ines Rehbein, and Simone Paolo Ponzetto. 2024. [A Survey on Modelling Morality for Text Analysis](#). In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 4136–4155.
- Valentina Rizzoli, Mauro Sarrica, Andrea Bobbio, Luciana Carraro, and Francesca Romana Alparone. 2026. [The Words of Morality: An Italian Adaptation of the Moral Foundations Dictionary 2.0 for the Linguistic Inquiry Word Count](#). *Journal of Language and Social Psychology*, 45(2):293–315.
- Milena Šošić. 2022. [SRPOL–A Lexicon Based Framework for Sentiment Strength of Serbian Texts](#). *Review of the NCD*, 41:58–73.
- Ranka Stanković, Cvetana Krstev, Branislava Šandrih Todorović, Duško Vitas, Mihailo Škorić, and Milica Ikonić Nešić. 2022. [Distant Reading in Digital Humanities: Case Study on the Serbian Part of the ELTeC Collection](#). In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 3337–3345, Marseille, France. ELRA.
- Ranka Stanković, Cvetana Krstev, and Duško Vitas. 2024. [SrpEL-TeC: A Serbian Literary Corpus for Distant Reading](#). *Primerjalna književnost*, 47(2):45–63.
- Aleksandra Trtovac, Vasilije Milnović, and Cvetana Krstev. 2021. [The Serbian Part of the ELTeC - from the Empty List to the 100 Novels Collection](#). *Infotheca - Journal for Digital Humanities*, 21(2):7–25.
- Mafalda Zúquete, Diana Orghian, and Flávio L Pinheiro. 2023. [A Moral Foundations Dictionary for the European Portuguese Language: The Case of Portuguese Parliamentary Debates](#). In *International Conference on Computational Science*, pages 421–434. Springer.
- Milena Šošić. 2025. [Modeling Moral and Emotional Language Aspects in Conversational Texts Classification](#). Phd thesis, University of Belgrade, Faculty of Mathematics. In Serbian.

---

## Hijerarhijsko prepoznavanje moralnih vrednosti u književnom korpusu *SrpELTeC*

---

Milena Šošić, Ranka Stanković

### Sažetak

Ovaj rad predstavlja novi pristup za hijerarhijsku analizu morala zasnovanoj na Teoriji moralnih vrednosti (*Moral Foundations Theory, MFT*) u okviru korpusa srpske književnosti *SrpELTeC*. Predloženi metod automatski identifikuje i kvantifikuje pet osnovnih moralnih vrednosti zajedno sa njima odgovarajućim kategorijama sentimenta – briga/šteta (eng. care/harm), pravičnost/prevara (eng. fairness/cheating), lojalnost/izdaja (eng. loyalty/betrayal), autoritet/subverzija (eng. authority/subversion) i čistota/degradacija (eng. purity/degradation), u okviru različitih segmenata srpskih romana. Izračunavanje je izvršeno korišćenjem razvijenog leksikona moralnih reči *MFD.SR*, prilagođenog lingvističkim i kulturološkim karakteristikama srpskog jezika. Ključni metodološki doprinos ovog rada predstavlja prepoznavanje moralnih signala na više nivoa hijerarhijski strukturisanog korpusa, što omogućava uporedivu analizu tekstova različitih veličina i karakteristika. Na ovaj način, predloženi metod generiše reprezentacije moralne dinamike na makro i mikro nivou u okviru književnih narativa, čime se omogućava granularno praćenje promena moralnih vrednosti kroz različite segmente korpusa. Preliminarni rezultati ukazuju da srpski romani prvi put objavljeni u 1840-1920. periodu, ispoljavaju prepoznatljive i merljive moralne obrasce među autorima i književnim pravcima, čime se otvaraju novi uvidi u ideološku i kulturološku strukturu srpske književnosti ovog vremenskog perioda.

**Ključne reči:** moralne vrednosti, sentiment, leksikon, književni korpus, srpski jezik

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

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

