
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ/ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

DOI: <https://doi.org/10.60797/IRJ.2025.162.15> EDN: GDJBAA**КОРПУСНЫЙ АНАЛИЗ ФАКТОРОВ ВЛИЯНИЯ НА ТОНАЛЬНОСТЬ СЛОВА «ПАТРИОТИЗМ»**

Научная статья

Шарнин М.М.^{1,*}¹ ORCID : 0000-0003-0450-5156;¹ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Российская Федерация

* Корреспондирующий автор (mc[at]keywen.com)

Предложена: 06.11.2025; Принята: 10.12.2025; Опубликовано: 17.12.2025

Аннотация

В работе представлен эффективный метод выявления позитивных и негативных факторов, влияющих на эмоциональную тональность слова «патриотизм», в корпусе текстов объемом 10 ГБ из сообщений в телеграм-каналах. Факторы, влияющие на тональность, выявляются среди семантических связей, полученных с помощью нейросети Word2Vec и тональных словарей. Метод использует статистические закономерности между словами из тональных словарей, входящими в ассоциативные группы из близких по семантике слов. Наиболее сильными позитивными факторами в корпусе являются: нравственность, духовность, любовь к Родине, гуманизм. Метод позволяет осуществлять мониторинг факторов, влияющих на тональность слова «патриотизм», в большом динамическом корпусе текстов, что может способствовать повышению эффективности работы по продвижению патриотизма как традиционной ценности в информационной среде.

Ключевые слова: корпусный анализ, факторы влияния на тональность, патриотизм, семантические связи, нейронная сеть, тональные словари.

CORPUS ANALYSIS OF FACTORS INFLUENCING THE TONE OF THE WORD "PATRIOTISM"

Research article

Sharnin M.M.^{1,*}¹ ORCID : 0000-0003-0450-5156;¹ Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russian Federation

* Corresponding author (mc[at]keywen.com)

Suggested: 06.11.2025; Accepted: 10.12.2025; Published: 17.12.2025

Abstract

The work presents an effective method for identifying positive and negative factors influencing the emotional tone of the word "patriotism" in a 10 GB corpus of texts from Telegram channel messages. Factors influencing tone are identified among semantic connections obtained using the Word2Vec neural network and tonal dictionaries. The method uses statistical patterns between words from tonal dictionaries that are part of associative groups of semantically similar words. The strongest positive factors in the corpus are morality, spirituality, love for the motherland, and humanism. The method allows to monitor the factors influencing the tone of the word "patriotism" in a large, dynamic corpus of texts, which can contribute to increasing the effectiveness of promoting patriotism as a traditional value in the information environment.

Keywords: corpus analysis, factors influencing tone, patriotism, semantic connections, neural network, tonal dictionaries.

Введение

Целью настоящей работы является выявление факторов, влияющих на отношение к патриотизму в современном обществе. Патриотизм — это одна из традиционных ценностей, которые согласно Стратегии национальной безопасности Российской Федерации необходимо защищать, а также принимать меры по их сохранению и укреплению [1]. Эмоциональная тональность слова «патриотизм» в большом корпусе текстов отражает отношение авторов этих текстов к патриотизму. Если слово «патриотизм» имеет позитивную эмоциональную тональность в популярных текстах, то это способствует сохранению и укреплению патриотизма как ценности.

Эмоциональная тональность терминов/ценностей может меняться со временем благодаря появлению новых популярных текстов, в которых эти термины окружает соответствующий позитивный или негативный эмоциональный фон, и создаются соответствующие семантические связи (в том числе ассоциативные). Связи терминов с позитивными словами, такими как комфорт, успех, семейное счастье, создают позитивный/привлекательный образ и повышают эмоциональную тональность этих терминов, что часто используется в рекламе недвижимости [2]. Напротив, деструктивные тексты разрушают и обесценивают ценности, а также снижают их позитивную тональность путем создания семантических связей с негативными словами. Негативизация — это одна из наиболее часто встречаемых деструктивных стратегий [3].

Семантические связи ценностей с позитивными и негативными словами влияют на эмоциональную тональность ценностей и поэтому являются факторами позитивного и негативного влияния. В научной литературе термин «фактор»

обозначает «элемент, условие или параметр, который влияет на исследуемый процесс или явление» [4]. Знание и понимание факторов помогает анализировать закономерности, предсказывать результаты, строить работающие модели.

В настоящее время широкое распространение получили методы тонального анализа текстов (sentiment analysis) [5], с помощью которых можно автоматически определять эмоциональную окраску/тональность текста и его фрагментов, чтобы понять, отражают ли они позитивное, негативное или нейтральное отношение автора к соответствующим объектам, явлениям и ценностям. В некоторых случаях для тонального анализа текстов используются тональные словари [6], [7]. Тональный словарь (sentiment dictionary, sentiment lexicon) — это список слов и словосочетаний с указанной тональностью (положительной или отрицательной). В тональном словаре словам и словосочетаниям сопоставляются соответствующие им оценки настроения или эмоциональные коннотации. Тональный словарь может создаваться экспертами [6] или с помощью опроса/краудсорсинга [7], в котором о каждом слове опрашивается определенное количество участников. Тональные словари используются для различных задач, связанных с тональным анализом текстов, например, для анализа отзывов о товарах и исследования социальных трендов. Особенно часто тональные словари применяются для тонального анализа редких языков, которые не имеют достаточного количества размеченных корпусов, необходимых для качественной работы программ тонального анализа на основе машинного обучения.

В ряде научных работ [8], [9], [10] тональный анализ текстов применяется для выявления факторов, влияющих на тональность интересующего объекта/явления. В этих работах в анализируемых текстах сначала вместе с определением тональности выявляется множество элементов таких как ключевые слова, темы, параметры, переменные, а затем среди выявленных элементов с помощью статистического анализа (например, корреляции) определяются элементы/факторы, статистически связанные с тональностью интересующего объекта/явления.

В настоящем исследовании факторы, влияющие на тональность слова «патриотизм», выявляются в корпусе текстов среди семантических связей, полученных с помощью нейросети Word2Vec и тональных словарей.

Методы и принципы исследования

Для проведения исследования использовался корпус текстов объемом 10 ГБ из сообщений телеграм-каналов, таких как Радио Говорит Москва, Владимир Соловьев и RT на русском, а также из связанных по ссылкам телеграм-каналов. В корпусе содержались данные за несколько последних лет от начала каждого телеграм-канала до января 2025 года. С помощью программы gensim.models.Word2Vec на языке Python по корпусу текстов были построены 3 лингвистические модели для слов с частотами >1000, >100 и >5. Затем для каждого слова в каждой модели с помощью функции model.most_similar были рассчитаны ассоциативные группы, содержащие наиболее близкие по семантике слова. Приведем три примера ассоциативных групп со словом «патриотизм».

- Группа 1: любовь, нежность, доброта, сострадание, вера, теплота, патриотизм, святость, милосердие, радость, чувство.
- Группа 2: патриотизм, духовность, любовь, нравственность, нравственный, гуманизм, чувство, воспитание, гордость, идеал, эгоизм.
- Группа 3: патриотизм, духовность, мужественность, нравственность, нравственный, гуманизм, любовь, гражданственность, эгоизм, коллективизм, чувство.

Вторая группа состоит из наиболее семантически близких слов с количеством вхождений в корпусе >1000, а третья — из слов с количеством вхождений в корпусе >100. Всего слово «патриотизм» вошло в 32 различные ассоциативные группы.

Для анализа тональности слов в ассоциативных группах использовался объединенный тональный словарь на базе словарей RuSentiLex [6] и KartaSlovSent [7]. Словарь RuSentiLex содержит 2368 позитивных и 7177 негативных слов. Словарь KartaSlovSent содержит 6215 позитивных и 11863 негативных слов. В объединенный словарь вошли 6892 позитивных и 14156 негативных слов, которые имели одинаковую или непротиворечащую тональность в базовых словарях. Эти позитивные и негативные слова мы в дальнейшем называем тональными словами.

Предлагаемая методика использует статистические закономерности между тональными словами, входящими в ассоциативные группы. Тональности большинства слов в каждой ассоциативной группе часто совпадают. Поэтому большинство тональных слов, окружающих любое слово в различных ассоциативных группах, имеют одинаковую тональность. Подсчет подобной статистики совместной встречаемости тональных слов в ассоциативных группах для 21048 слов из объединенного словаря дал следующие результаты. Если в ассоциативных группах со словом X разность числа окружающих позитивных и негативных тональных слов равна +1 или -1, то с вероятностью 67% это слово X имеет такую же тональность как у большинства окружающих тональных слов. Данный факт был установлен для 964 слов из объединенного тонального словаря. Если разность равна +2 или -2, то вероятность повышается до 71.8% из 666 случаев. Если разность равна +3 или -3, то вероятность повышается до 74.4% из 523 случаев. Эта закономерность представлена на графике рис.1.

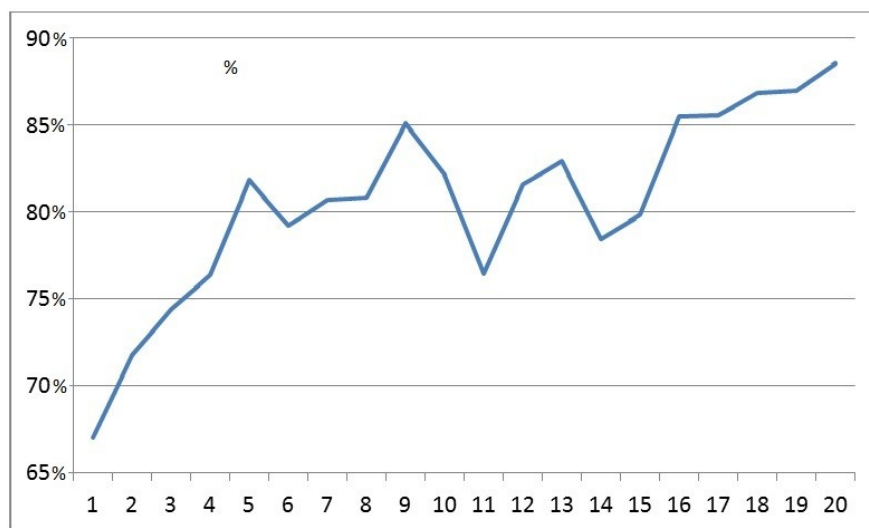


Рисунок 1 - График вероятности совпадения тональности слова X со знаком разности (Y-Z) в зависимости от абсолютного значения $|Y-Z|$:

Y – это количество вхождений позитивных тональных слов в ассоциативные группы со словом X; Z - количество вхождений негативных тональных слов в эти группы

DOI: <https://doi.org/10.60797/IRJ.2025.162.15.1>

Из графика видно, что тональность слова X зависит от разности числа семантически похожих позитивных и негативных тональных слов. Поэтому семантически похожие позитивные слова являются позитивными факторами, влияющими на тональность, причем их влияние тем сильнее, чем больше число Y. Аналогично, негативные слова являются негативными факторами, влияющими на тональность, причем их влияние тем сильнее, чем больше число Z.

В соответствии с описанной статистической закономерностью для определения факторов влияющих на тональность слова «патриотизм» в 32 группах со словом «патриотизм» были подсчитаны количества вхождений для каждого позитивного и негативного тонального слова. На основе полученной статистики были построены рейтинги позитивных и негативных слов по количеству их вхождений. Полученные рейтинги определяют наиболее сильные факторы влияния на тональность слова «патриотизм», и поэтому их можно назвать рейтингами позитивных и негативных факторов.

Основные результаты

В соответствии с описанной методикой был построен следующий рейтинг позитивных факторов: нравственный (16), духовность (15), нравственность (13), любовь (10), гуманизм (8), мораль (7), воспитание (7), вера (6), эмпатия (5), гордость (5), духовный (5), святость (4), идеал (4), моральный (4), коллективизм (4), сострадание (3), радость (3), мудрость (3), задор (3), восхищение (3), благодарность (3), мотивация (3), стимул (3), добродетель (3), мужественность (3), обожание (3), нежность (2), доброта (2), благородство (2), стойкость (2), мощь (2), человеческий (2), социализация (2), наставничество (2), воспитательный (2), единение (2), гибкость (2), ценностный (2), гуманистический (2), ценность (2), подрастать (2), самоуважение (2), мотивированность (2), юношество (2), трудолюбие (2), образованность (2), осознанность (2), всепрощение (2), целомудрие (2).

Рейтинг негативных факторов: эгоизм (4), стыд (3), презрение (2), шовинизм (2), совковый (2). В скобках указано количество вхождений слова в 32 ассоциативные группы, что соответствует силе данного фактора.

Полученные результаты находят подтверждение в официальной и научной литературе. Приведем примеры работ, подтверждающие некоторые выявленные факторы и их тональность. В приведенных ниже примерах выявленные факторы выделены жирным шрифтом. Так, в работе [11] утверждается, что «противоположностью или врагом патриотизма является эгоизм и его разновидности (эгоцентризм, нарциссизм и пр.)».

«Открыто приняв в 90-х годах западные ценности и западную культуру, мы как-то растеряли, а молодое поколение так и не сформировало, чувство гордости за свою землю, свой народ, свое государство» [12]. Западная система образования вместо гордости за свою Родину и ее исторические достижения насаждала эгоизм, жажду наживы, презрение к патриотизму, стыд за свою страну. «Слово «патриотизм» в устах некоторых идеологов и политиков стало ругательным, уничижительным, отождествляемым с такими понятиями, как «совки», «комуняки» [13].

Председатель Совета Федерации Валентина Матвиенко уверена, что патриотизм «должен стать неотъемлемой частью характера человека. Но при этом не перерастать в шовинизм. Подлинный патриот — тот, кто любит свою Родину и в то же время с уважением относится к другим народам» [14].

Таким образом, защите и укреплению патриотизма способствует уменьшение в обществе проявлений негативных факторов, таких как эгоизм, шовинизм, стыд за свою страну и презрение к патриотизму, а также увеличение проявлений позитивных факторов, включая нравственность, духовность, любовь к Родине, гордость за свою страну и ее исторические достижения.

Заключение

В работе представлен эффективный метод выявления позитивных и негативных факторов, влияющих на эмоциональную тональность слова «патриотизм», в анализируемом корпусе текстов. Факторы, влияющие на тональность, выявляются среди семантических связей, полученных с помощью нейросети Word2Vec и тональных словарей.

Предлагаемый метод позволяет осуществлять мониторинг факторов, влияющих на тональность, в большом динамическом корпусе текстов, что может повысить эффективность работы по патриотическому воспитанию, защите и укреплению традиционных ценностей, а также по продвижению патриотизма как традиционной ценности в информационной среде.

Конфликт интересов

Не указан.

Рецензия

Сообщество рецензентов Международного научно-исследовательского журнала

DOI: <https://doi.org/10.60797/IRJ.2025.162.15.2>

Conflict of Interest

None declared.

Review

Community of Reviewers of the International Research Journal

DOI: <https://doi.org/10.60797/IRJ.2025.162.15.2>

Список литературы / References

1. Указ Президента Российской Федерации от 9 ноября 2022 г. № 809 // Сайт Президента Российской Федерации. — 2022. — URL: <http://www.kremlin.ru/acts/bank/48502?erid=LjN8K8S> (дата обращения: 07.11.25)
2. Тихомирова Л.С. Специфика реализации коммуникативных стратегий и тактик в рекламе недвижимости (на материале интернет-сайтов г. Перми) / Л.С. Тихомирова, А.С. Черноусова // Медиалингвистика. — 2024. — № 11 (3). — С. 325–340. — DOI: 10.21638/spbu22.2024.303
3. Карабулатова И.С. Специфика лингвистической параметризации деструктивного массмедийного текста с обесцениванием исторической памяти / И.С. Карабулатова, Г.А. Копнина // Медиалингвистика. — 2023. — № 10 (3). — С. 319–335. — DOI: 10.21638/spbu22.2023.303
4. Атаманова Г.И. Термин «Фактор» как общенаучный феномен / Г.И. Атаманова, А.С. Аверьянская // Международный научно-исследовательский журнал. — 2024. — № 12 (150). — URL: <https://research-journal.org/archive/12-150-2024-december/10.60797/IRJ.2024.150.150> (дата обращения: 07.11.25). — DOI: 10.60797/IRJ.2024.150.150
5. Wankhade M. A survey on sentiment analysis methods, applications, and challenges / M. Wankhade, A.C.S. Rao, C. Kulkarni // Artificial Intelligence Review. — 2022. — № 55 (7). — P. 5731–5780.
6. Loukachevitch N. Creating a general Russian sentiment lexicon / N. Loukachevitch, A. Levchik // Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16). — 2016. — P. 1171–1176.
7. Kulagin D. Publicly available sentiment dictionary for the Russian language KartaSlovSent / D. Kulagin // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”. — 2021. — № 20. — P. 1106–1119.
8. Liu R. Factors affecting users' satisfaction with urban parks through online comments data: Evidence from Shenzhen, China / R. Liu, J. Xiao // International Journal of Environmental Research and Public Health. — 2021. — № 18 (1). — P. 253.
9. Yan X. Analyzing Antecedent Configurations of Group Emotion Generation in Public Emergencies: A Multi-Factor Coupling Approach / X. Yan, Y. Liu, Y. Chen et al. // Behavioral Sciences. — 2025. — № 15 (1). — P. 41. — DOI: 10.3390/bs15010041
10. Li M. Sentiment analysis and prediction model based on Chinese government affairs microblogs / M. Li, Y. Shi // Heliyon. — 2023. — № 9 (8).
11. Воеводин А.П. Нравственные смыслы патриотизма / А.П. Воеводин // Вестник Московского государственного университета культуры и искусств. — 2024. — № 4 (120). — С. 39–46. — DOI: 10.24412/1997-0803-2024-4120-39-46
12. Шинкевич В.Е. Опыт преподавания курса «Основы Российской государственности» в контексте формирования патриотизма у обучающихся высшей школы / В.Е. Шинкевич // Вестник Восточно-Сибирской Открытой Академии. — 2024. — № 53 (53).
13. Стрелкова Н.В. Семья — основа воспитания патриотизма и духовности / Н.В. Стрелкова // Проблемы патриотического воспитания сотрудников правоохранительных органов и противодействия идеологии неонацизма и экстремизма. Сборник научных трудов. Московский университет Министерства внутренних дел Российской Федерации. — Москва, 2023. — С. 151–156.
14. Матвиенко В.И. Здесь раскрывается завеса будущего нашей Родины / В.И. Матвиенко // Блог. — 2023. — URL: <http://council.gov.ru/services/discussions/blogs/143112/> (дата обращения: 07.11.25)

Список литературы на английском языке / References in English

1. Ukaz Prezidenta Rossiiskoi Federatsii ot 9 noyabrya 2022 g. № 809 [Decree of the President of the Russian Federation of November 9, 2022 No. 809] // Website of the President of the Russian Federation. — 2022. — URL: <http://www.kremlin.ru/acts/bank/48502?erid=LjN8K8S> (accessed: 07.11.25) [in Russian]
2. Tixomirova L.S. Specifika realizacii kommunikativny'x strategij i taktik v reklame nedvizhimosti (na materiale internet-sajtov g. Permi) [Specifics of implementing communication strategies and tactics in real estate advertising (based on materials



- from Perm city websites)] / L.S. Tixomirova, A.S. Chernousova // *Medialinguistics*. — 2024. — № 11 (3). — P. 325–340. — DOI: 10.21638/spbu22.2024.303 [in Russian]
3. Karabulatova I.S. Specifika lingvističeskoj parametrizacii destruktivnogo massmedijnogo teksta s obescenivaniem istoričeskoj pamyati [The specificity of linguistic parameterization of destructive mass media text with devaluation of historical memory] / I.S. Karabulatova, G.A. Kopnina // *Medialinguistics*. — 2023. — № 10 (3). — P. 319–335. — DOI: 10.21638/spbu22.2023.303 [in Russian]
4. Atamanova G.I. Termin «Faktor» kak obshhenauchnyj fenomen [The term "factor" as a general scientific phenomenon] / G.I. Atamanova, A.S. Aver'yaskina // *International Research Journal*. — 2024. — № 12 (150). — URL: <https://research-journal.org/archive/12-150-2024-december/10.60797/IRJ.2024.150.150> (accessed: 07.11.25). — DOI: 10.60797/IRJ.2024.150.150 [in Russian]
5. Wankhade M. A survey on sentiment analysis methods, applications, and challenges / M. Wankhade, A.C.S. Rao, C. Kulkarni // *Artificial Intelligence Review*. — 2022. — № 55 (7). — P. 5731–5780.
6. Loukachevitch N. Creating a general Russian sentiment lexicon / N. Loukachevitch, A. Levchik // *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. — 2016. — P. 1171–1176.
7. Kulagin D. Publicly available sentiment dictionary for the Russian language KartaSlovSent / D. Kulagin // *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"*. — 2021. — № 20. — P. 1106–1119.
8. Liu R. Factors affecting users' satisfaction with urban parks through online comments data: Evidence from Shenzhen, China / R. Liu, J. Xiao // *International Journal of Environmental Research and Public Health*. — 2021. — № 18 (1). — P. 253.
9. Yan X. Analyzing Antecedent Configurations of Group Emotion Generation in Public Emergencies: A Multi-Factor Coupling Approach / X. Yan, Y. Liu, Y. Chen et al. // *Behavioral Sciences*. — 2025. — № 15 (1). — P. 41. — DOI: 10.3390/bs15010041
10. Li M. Sentiment analysis and prediction model based on Chinese government affairs microblogs / M. Li, Y. Shi // *Heliyon*. — 2023. — № 9 (8).
11. Voevodin A.P. Nравstvenny'e smysly' patriotizma [Moral meanings of patriotism] / A.P. Voevodin // *Bulletin of the Moscow State University of Culture and Arts*. — 2024. — № 4 (120). — P. 39–46. — DOI: 10.24412/1997-0803-2024-4120-39-46 [in Russian]
12. Shinkevich V.E. Opy't prepodavaniya kursa «Osnovy' Rossijskoj gosudarstvennosti» v kontekste formirovaniya patriotizma u obuchayushhixsya vy'sshej shkoly' [Experience of teaching the course "Fundamentals of Russian Statehood" in the context of developing patriotism among students of higher education] / V.E. Shinkevich // *Bulletin of the East Siberian Open Academy*. — 2024. — № 53 (53). [in Russian]
13. Strelkova N.V. Semya — osnova vospitaniya patriotizma i dukhovnosti [Family: The Foundation for Fostering Patriotism and Spirituality] / N.V. Strelkova // *Problems of Patriotic Education of Law Enforcement Officers and Countering the Ideology of Neo-Nazism and Extremism. A Collection of Scientific Papers. Kikot Moscow University of the Ministry of Internal Affairs of the Russian Federation*. — Moscow, 2023. — P. 151–156. [in Russian]
14. Matvienko V.I. Zdes raskrivaetsya zavesa budushchego nashei Rodini [Here the curtain on the future of our Motherland is lifted] / V.I. Matvienko // *Blog*. — 2023. — URL: <http://council.gov.ru/services/discussions/blogs/143112/> (accessed: 07.11.25) [in Russian]