

## ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ NLP ДЛЯ ОЦЕНКИ ЭФФЕКТИВНОСТИ МЕРОПРИЯТИЙ ПО ФОРМИРОВАНИЮ КОМФОРТНОЙ ГОРОДСКОЙ СРЕДЫ

© 2023 К. В. Галаган✉

*Югорский государственный университет  
ул. Чехова, 16, 628012 Ханты-Мансийск,  
Ханты-Мансийский автономный округ – Югра, Российская Федерация*

**Аннотация.** Формирование комфортной городской среды является одной из актуальных задач современного градостроительства. Для реализации этой задачи в 2018 г. Президентом и правительством Российской Федерации дан старт национальному проекту «Жилье и городская среда», в рамках которого определен исчерпывающий перечень мероприятий по формированию комфортной городской среды. Цель работы – формирование рейтинга эффективности таких мероприятий на основе анализа текстов сообщений граждан, пользователей социальных сетей, текстов публикаций в СМИ и описаний мероприятий проводимых на данной территории. В основу предлагаемого подхода положена оценка семантической близости описания мероприятий из набора датасетов по определенной территории. На основе сформированных для конкретных муниципальных образований датасетов выявлено расхождение ожиданий граждан с реализованными мероприятиями на данной территории.

**Ключевые слова:** национальный проект, комфортная городская среда, оценка мероприятий, NLP.

### ВВЕДЕНИЕ

В 2018 году был дан старт национальному проекту «Жилье и городская среда» (Проект). Срок реализации Проекта предусмотрен до 2030 года. В паспорте Проекта, разработанном Министерством строительства и жилищно-коммунального хозяйства РФ (Минстрой России), одной из целей указывается кардинальное повышение комфортности городской среды (КГС) [1]. Для достижения этой цели предполагается реализация мероприятий направленных на формирование комфортной городской среды (Мероприятия). Для оценки темпов и эффективности меро-

приятий, Минстроем России была предложена и утверждена методика формирования индекса качества городской среды (Методика) [2]. Подробное описание формирования индексов и их применение представлены в работе [3].

Основным критерием эффективности Мероприятий можно считать: своевременность и соответствие Мероприятий ожиданиям граждан. Для такой оценки Методика предписывает использовать индекс вовлеченности граждан в вопросы принятия решений связанных с формированием КГС (ФКГС), который рассчитывается на основе показателя доли граждан принявших участие в голосованиях на электронных площадках посвященных ФКГС. Как показывают исследования [4], такой способ коммуникации между

---

✉ Галаган Константин Владимирович  
e-mail: [galagankv@gmail.com](mailto:galagankv@gmail.com)



Контент доступен под лицензией Creative Commons Attribution 4.0 License.

The content is available under Creative Commons Attribution 4.0 License.

горожанами, как заказчиками, и властями муниципалитетов, как исполнителями Мероприятий, не достаточно эффективен. Низкая вовлеченность граждан в вопросы ФКГС, является одной из основных проблем реализации Проекта [5].

Для решения этой проблемы предлагается создание более качественных платформ для вовлечения граждан в вопросы управления ФКГС [6], а так же расширение каналов получения обратной связи от граждан, через проведение социологических исследований для выявления ожиданий граждан и привлечение экспертов [4].

В работе [7] авторами исследуется мотивация и факторы, влияющие на желание граждан участвовать в планировании городской среды. Авторы исследования выделяют четыре основных фактора влияющих на мотивацию граждан: «гражданское общество», «личный интерес», «социальное влияние» и «ограничения». Первые три можно рассматривать, как положительные факторы, «ограничения» являются отрицательным фактором. Фактор «ограничения» авторы напрямую связывают с низкой информированностью граждан.

Еще одним шагом в решении задачи повышения вовлеченности граждан в вопросы управления городской средой стала подготовка и утверждение в 2020 году Минстроем России Методических рекомендаций по вовлечению граждан, их объединений и иных лиц в решении вопросов развития городской среды [8]. В рамках этого решения в 2020 году Минстроем России была создана единая платформа для проведения голосований граждан по вопросам КГС [9].

В данном исследовании автором предлагается еще один способ оценки КГС через оценку эффективности Мероприятий. Способ основывается на семантическом сходстве наборов датасетов сообщений граждан с тематиками близкими к КГС и набора датасетов описания реализованных или планируемых к реализации Мероприятий для конкретной территории.

Семантическое сходство рассчитывается с применением нейросетевых моделей NLP.

## 1. ОБЗОР ЛИТЕРАТУРЫ

Анализ публикаций посвященных методам и подходам к оценке КГС показал, что существует несколько основных методик [10]. Широкое применение получил индексный метод [11], который позволяет оценивать, как отдельные пространства городской среды, так и состояние КГС в целом. Оценка, полученная индексным методом, позволяет проводить сравнительный анализ между различными городами, а так же рассматривать эту оценку, как интегральную составляющую в других методиках оценки [12]. К недостаткам этого метода можно отнести отсутствие развращенного мнения горожан.

Социальные опросы или анкетирование граждан [13–16] позволяют, прежде всего, получить обратную связь от граждан, а так же расширить оценку КГС, полученную индексным методом. Это важно для дальнейшего развития и улучшения городской среды. Оценка КГС в этих методах основывается на субъективном отношении граждан к городской среде. К недостаткам этих методов можно отнести ограниченность выборки, что не позволяет учесть мнение большинства горожан. Одним из способов преодоления данного недостатка, является получение обратной связи от неограниченного числа граждан. Так, например, для получения обратной связи предлагается анализировать сведения об использовании горожанами определенных мобильных приложений [17].

Задачи управления городской средой решаемые с помощью нейро-нечетких сетевых моделей условно можно разделить на два типа. Первый тип — это прикладные задачи, связанные с планированием застройки и инфраструктурой города [18]. В них нейро-нечеткие сетевые модели применяются для различного рода моделирования. Второго типа — это задачи связанные с обработкой неструктурированных данных. В задачах второго типа широко применяются методы NLP [19]. В частности, в работе [20] решается задача получения обратной связи от жителей города через анализ тональности в сообщениях социальной сети Twitter.

## 2. ПОСТАНОВКА ЗАДАЧИ И ДАННЫЕ

В общем виде задачу можно сформулировать так: необходимо оценить насколько семантически близко описание проблем в сообщениях граждан к описанию Мероприятий, которые публикуют официальные органы власти. Другими словами, направлены ли Мероприятия на решение тех проблем, о которых пишут граждане. Дополнительно необходимо было определить, какие из проблем являются наиболее значимыми в сообщениях граждан.

В качестве исходных данных было взято три набора датасетов с сообщениями. Два набора относились к сообщениям пользователей, третий содержал в себе сообщения СМИ о Мероприятиях. Каждый датасет включает в себя сообщения, сгруппированные по месяцу публикации, тематике и территории. Краткие характеристики наборов датасетов представлены в табл. 1.

## 3. ОПИСАНИЕ РЕШЕНИЯ

Решение было реализовано с помощью языка Python 3.8. Ключевые этапы алгоритма решения:

- Этап 1. Извлечение значимых N-грамм.
- Этап 2. Перевод полученных N-грамм в векторную форму, с учетом контекста.
- Этап 3. Определение семантической близости полученных векторов N-грамм.

В общем виде алгоритм решения представлен на рис. 1.

### 3.1. Извлечение значимых N-грамм

Для извлечения наиболее значимых ключевых N-грамм из сообщений, предложено использовать алгоритм TF-IDF [21]. Значи-

мость N-граммы определялась присвоенным ей рангом TF-IDF (R). В дальнейшем это позволит сформировать рейтинг значимости N-грамм для всех наборов датасета.

Опишем этапы алгоритма для извлечения и ранжирования значимых N-грамм.

1. Очистить все сообщения с помощью регулярных выражений.
2. Привести к начальной форме (лемматизировать) все слова в сообщениях.
3. Извлечь из каждого датасета первые 100 N-грамм для интервала N от 1 до 4.
4. Записать в базу данных извлеченные N-граммы с указанием ранга и исходного сообщения.
5. Просуммировать полученные ранги N-грамм по всем сообщениям.
6. Сформировать рейтинг N-грамм по всем датасетам набора

Каждый датасет будет обрабатываться отдельно. Регулярными выражениями должны быть удалены все символы кроме кириллических и знаков препинания. Так же должны быть удалены стоп-слова, фразы общепринятых обращений, приветствий, прощаний и имена собственные. Для лемматизации сообщений была использована библиотека `rutmorphu2` [22]. При извлечении N-грамм учитываются слова являющиеся существительными, прилагательными или глаголами.

### 3.2. Перевод полученных N-грамм в векторную форму, с учетом контекста

Для расчета семантической близости между отобранными N-граммами необходимо провести преобразование каждого слова в N-грамме в векторное представление. Существует несколько основных нейросетевых моделей позволяющих преобразовывать отдель-

Таблица 1. Краткие характеристики исходных наборов датасетов сообщений  
[Table 1. Brief characteristics of the original message datasets]

| Набор датасетов            | Количество датасетов | Общее количество сообщений в наборе |
|----------------------------|----------------------|-------------------------------------|
| Сообщения за 2019 г.       | 2466                 | 25891                               |
| Сообщения за 2020 г.       | 3778                 | 48006                               |
| Сообщения СМИ 2020-21 г.г. | 108                  | 290                                 |

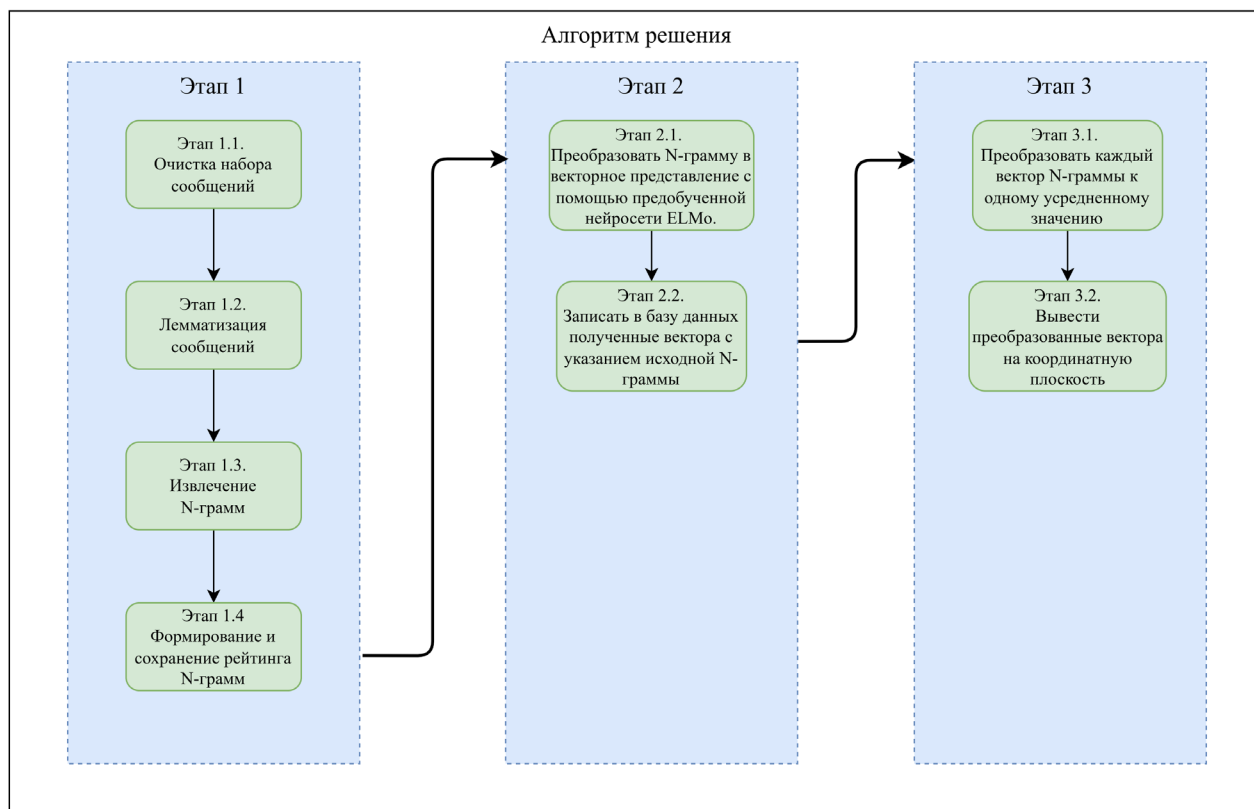


Рис. 1. Общий алгоритм решения  
[Fig 1. General solution algorithm]

ные слова или весь текст целиком в числовой вектор. Модели Word2Vec [23, 24] и GloVe [25] при определении вектора слова не учитывают контекст, в котором слово употребляется. Другими словами, у этих моделей слово «камера» в N-граммах «покрасочная камера» и «резиновая камера» дадут одинаковое значение вектора для слова «камера», что критично для нашей задачи, так как нам важен контекст слов в N-грамме. Модели BERT [26] и ELMo [27] лишены этого недостатка и позволяют получить контекстно-зависимые вектора для N-грамм.

Для получения контекстуального представления слов в N-граммах применяется следующий алгоритм:

1. Преобразовать N-грамму в векторное представление с помощью предобученной нейросети ELMo.
2. Записать в базу данных полученные вектора с указанием исходной N-граммы.

### 3.3. Определение семантической близости, полученных векторов N-грамм.

Задачу определения семантической близости можно представить, как вывод значений полученных на предыдущем этапе векторов на координатную плоскость. Для упрощения понимания полученных результатов, необходимо использовать двумерную координатную плоскость. Таким образом, алгоритм для определения семантической близости выглядит так:

1. Преобразовать каждый вектор N-граммы к одному усредненному значению.
2. Вывести преобразованные вектора на координатную плоскость.

## 4. ЭКСПЕРИМЕНТ

На основе предложенного выше алгоритма решения была произведена оценка эффективности Мероприятий. Для целей данной работы в расчет брались N-граммы, при N=2 (би-граммы), занимающие первые 20 мест

по сумме R. Характеристики значений R для ТОП-20 извлеченных N-грамм приведены в табл. 2. R для N-грамм при N=3 и N=4 имеют значения <1, в связи с чем было принято решение исключить их из дальнейшей обработки и анализа. N-граммы при N=1 малоинформативны и так же исключены из анализа.

Результаты первого этапа решения представлены на рис. 2, 3, 4.

Для получения числового вектора контекстуального представления слов в би-грамме была выбрана одна из предобученных моделей ELMo [28]. Реализация этой модели

позволяет получать усредненное значение вектора для каждого слова в би-грамме, что существенно облегчает последующую визуализацию результатов.

Визуализация преобразованных векторов для ТОП-20 би-грамм представлена на рис. 5. Как видно, значения векторов для би-грамм из сообщений граждан, слабо совпадают со значениями векторов для би-грамм из описания Мероприятий. Из чего автор делает вывод, что Мероприятия могут иметь низкую эффективность и не полностью удовлетворять потребности граждан в КГС.

Таблица 2. Краткие характеристики для ТОП-20 извлеченных N-грамм  
[Table 2 Brief characteristics for the TOP-20 extracted N-grams]

| Значение N | Сообщения граждан за 2019 г. |         | Сообщения граждан за 2020 г. |         | Сообщения СМИ за 2020-21 гг. |         |
|------------|------------------------------|---------|------------------------------|---------|------------------------------|---------|
|            | Max R                        | Min R   | Max R                        | Min R   | Max R                        | Min R   |
| N=2        | 11,3822                      | 1,33856 | 16,8417                      | 3,59251 | 1,41499                      | 0,41887 |
| N=3        | 0,44074                      | 0,31622 | 0,31622                      | 0,26726 | 0,42529                      | 0,18619 |
| N=4        | 0,33225                      | 0,28867 | 0,31622                      | 0,26726 | 0,27389                      | 0,13575 |

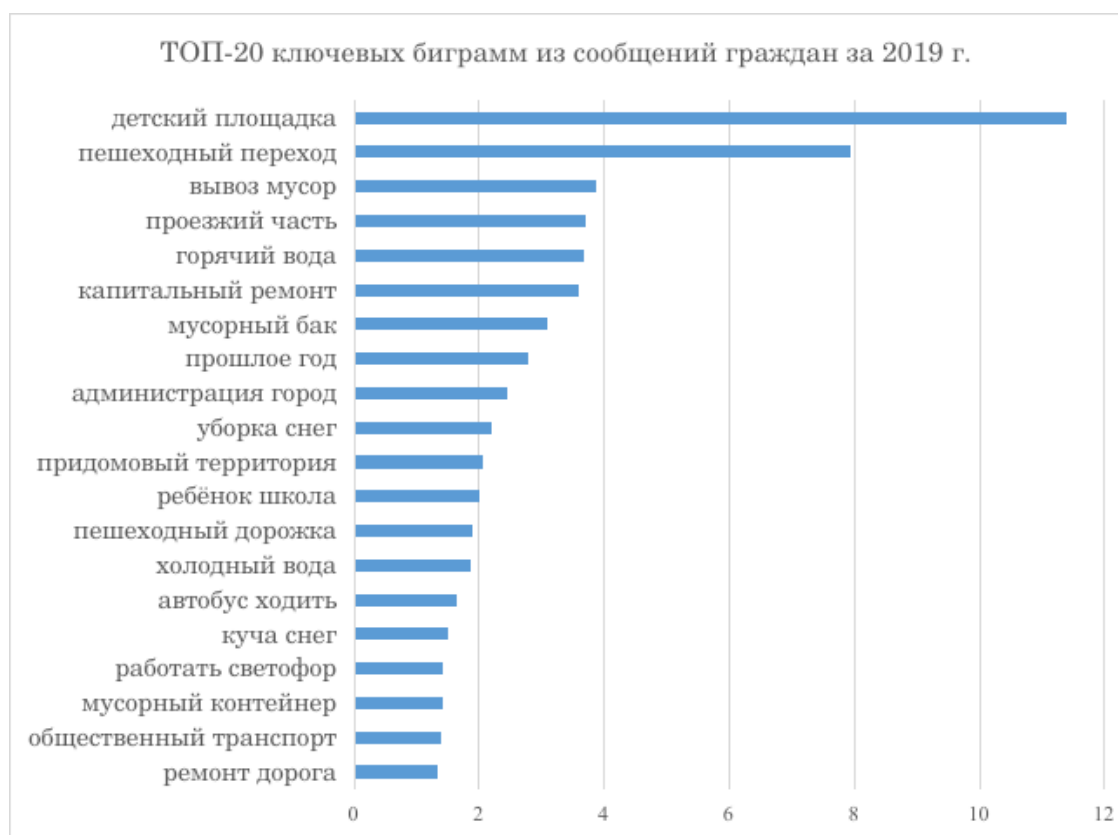


Рис. 2. ТОП-20 би-грамм из сообщений граждан за 2019 г.  
[Fig. 2. TOP 20 bi-grams from citizens' messages in 2019]



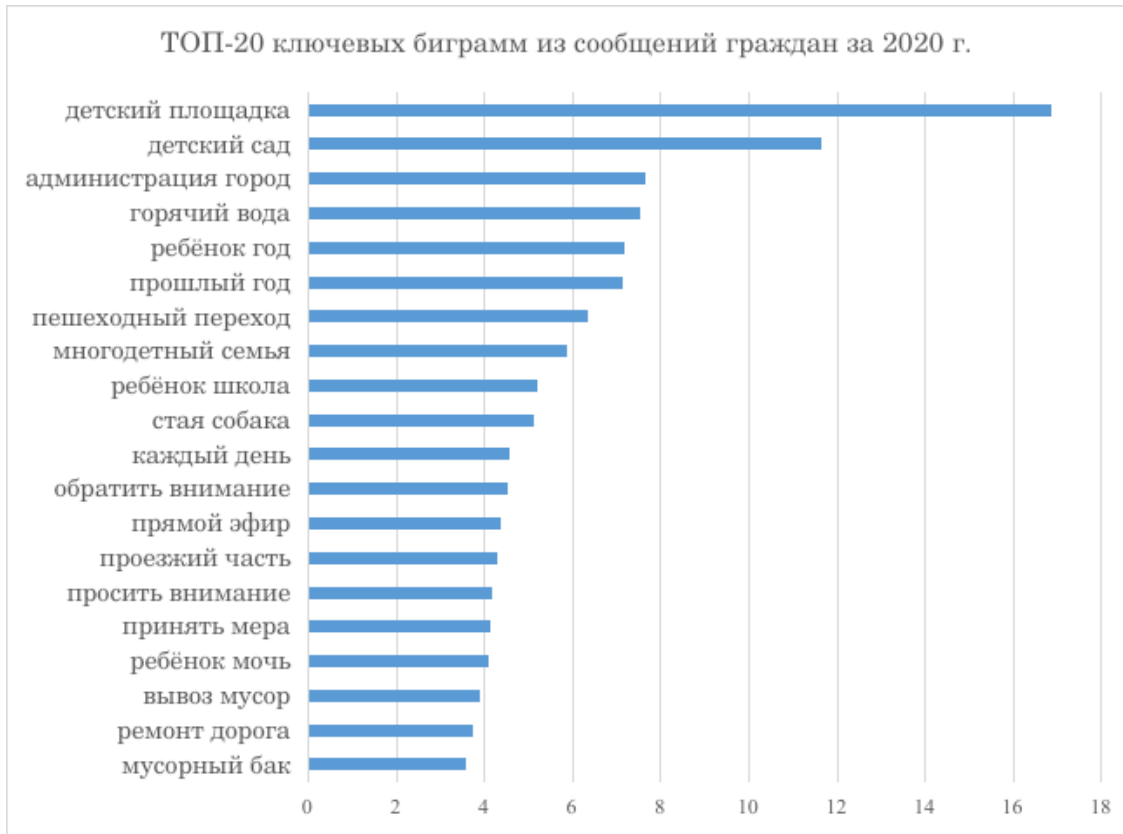


Рис. 3. ТОП-20 би-грамм из сообщений граждан за 2020 г.  
[Fig. 3. TOP 20 bi-grams from citizens' messages in 2020]

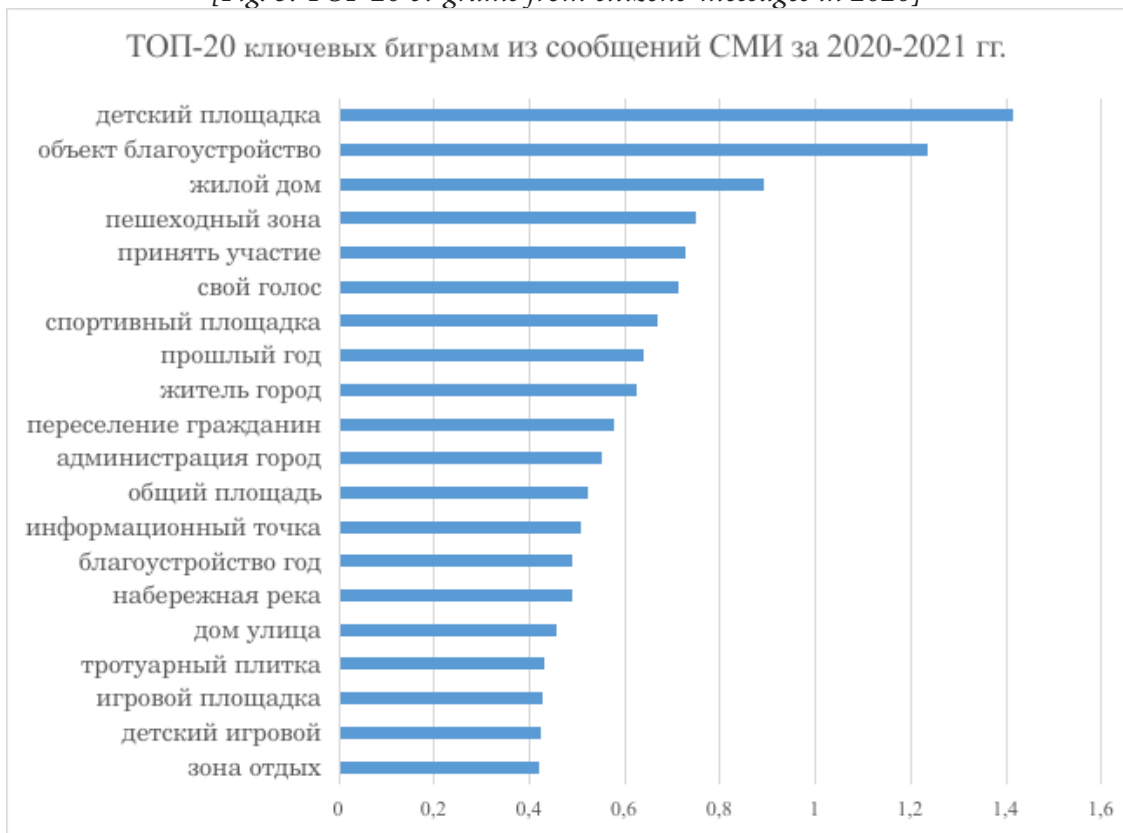


Рис. 4. ТОП-20 би-грамм из сообщений СМИ за 2020-2021 г.г.  
[Fig. 4. TOP 20 bi-grams from media reports for 2020-2021]

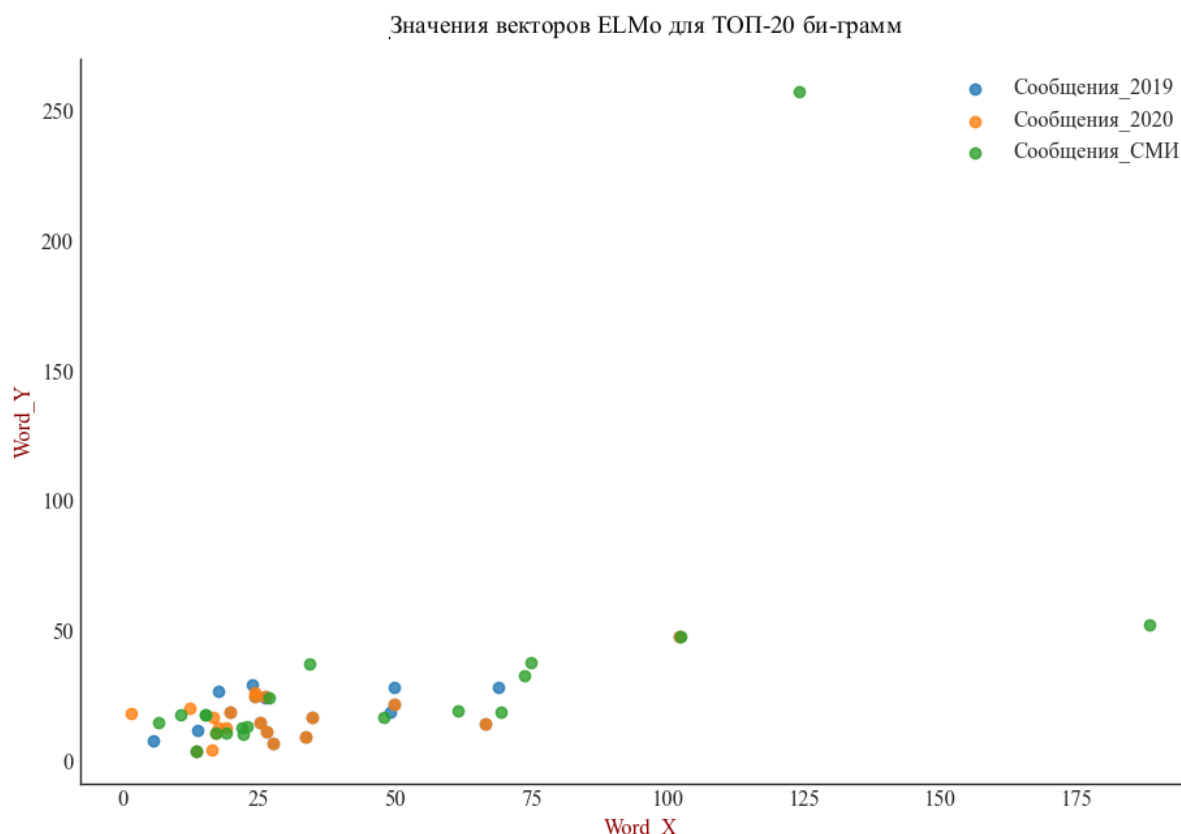


Рис. 5. Значения векторов ELMo для ТОП-20 би-грамм  
 [Fig. 5. ELMo vector values for TOP-20 bi-grams]

## ЗАКЛЮЧЕНИЕ

Результатом предложенного решения явился набор N-грамм, извлеченных из текстовых сообщений граждан и СМИ, на основании которого была дана оценка эффективности Мероприятий. Оценка основывалась на сравнении семантического сходства N-грамм, извлеченных из текстовых сообщений граждан и N-грамм, извлеченных из текстовых описаний Мероприятий. Данная оценка выявила расхождение между ожиданиями граждан и Мероприятиями проводимыми властями муниципалитетов.

Основываясь на полученных результатах была разработана модель оценки Мероприятий на базе нейросетевых моделей NLP. Данный подход позволяет полнее учитывать ожидания граждан в отношении КГС при планировании Мероприятий, дополняя и расширяя результаты традиционных методов получения обратной связи, таких как: социологические опросы, анкетирование и общественные обсуждения на электронных площадках.

Решение так же может быть интегрировано в систему поддержки принятия решений при проведении оценки значимости Мероприятий. Своевременная реализация Мероприятий, направленных на решение реальных потребностей жителей города положительно скажется на оценке гражданами деятельности местных властей и повысит вовлеченность граждан в вопросы управления городскими пространствами.

## КОНФЛИКТ ИНТЕРЕСОВ

Автор декларирует отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

## СПИСОК ЛИТЕРАТУРЫ

1. Паспорт национального проекта «Жилье и городская среда», утверждённый на заседании президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам 24 декабря 2018 г.

2. Распоряжение Правительства РФ от 23.03.2019 № 510-р (ред. от 30.12.2020) «Об утверждении Методики формирования индекса качества городской среды».
3. Королева П. В. Индекс качества городской среды как способ оценки государственной политики по формированию комфортной городской среды / П. В. Королева // Сборник научных статей VI Международной научной конференции студентов и молодых ученых: в 3 т. – Т. 1: Экономика. Гуманитарные науки. Курск, 2019. – С. 306–312.
4. Опыт общественного участия в планировании комфортной городской среды на примере Архангельской области / А. Г. Деменев [и др.] // Арктика и Север. – 2018. – № 33. – С. 91–117. DOI: 10.17238/issn2221-2698.2018.33.91
5. Дмитриева Н. Н. Формирование комфортной городской среды – как стратегическое направление развития проекта «ЖКХ и городская среда» / Н. Н. Дмитриева, Т. М. Ипатова // Социально-экономическое управление: Теория и практика. – 2018. – № 1(32). – С. 95–98.
6. Мухаметов Д. Р. Модели платформ вовлечения граждан для создания в России умных городов нового поколения / Д. Р. Мухаметов // Вопросы инновационной экономики. – 2020. – Т. 10, № 3. – С. 1605–1622.
7. Analysis of citizens' motivation and participation intention in urban planning / W. Li [et al.] // Cities. 2020, – Vol. 106. – P. 102921. DOI: 10.1016/j.cities.2020.102921
8. Приказ Минстроя России № 913/пр от 30.12.2020 «Об утверждении методических рекомендаций по вовлечению граждан, их объединений и иных лиц в решение вопросов развития городской среды».
9. Делаем город комфортнее. – Режим доступа: <https://za.gorodsreda.ru> – (Дата обращения: 31.03.2022).
10. Подходы к оценке качества городской среды / Е. А. Белякова [и др.] // Вестник ПГУ-АС: Строительство, наука и образование. – 2019. – № 2(9). – С. 3–9.
11. Нотман О. В. Индексный метод оценки качества городской среды: международный и российский опыт / О. В. Нотман // Вестник Адыгейского государственного университета. Серия 1: Регионоведение: философия, история, социология, юриспруденция, политология, культурология. – 2021. – № 2 (279). – С. 89–99.
12. Стригунов В. В. Оценка состояния города Хабаровска по федеральным индексам качества городской среды и IQ городов / В. В. Стригунов, Т. А. Бочарова // Ученые заметки ТОГУ. – 2021. – Т. 12, № 1. – С. 5–12.
13. Аксенова В. В. Комфортная городская среда: общественное мнение москвичей / В. В. Аксенова // Социальная политика и социология. – 2020. – Т. 19, 4 (137). – С. 76–84. DOI: 10.17922/2071-3665-2020-19-4-76-84
14. Богданова Л. П. Оценка качества городской среды населением города Твери / Л. П. Богданова, М. А. Глушкова // Вестник тверского государственного университета. Серия: География и геоэкология. – 2021. – № 2(34). – С. 14–24. DOI: 10.26456/2226-7719-2021-2-14-24
15. Дунаева Д. О. Дискурсивные практики горожан как коммуникативный механизм формирования образа комфортного города (опыт полевого исследования) / Д. О. Дунаева // Вестник томского государственного университета. Философия. Социология. Политология. – 2021. – № 60. – С. 137–150. DOI: 10.17223/1998863X/60/13
16. Russova O. N. Assessment of the Comfort of the Urban Environment as a Factor in the Social Well-Being of Citizens of the Arkhangelsk Oblast / O. N. Russova, T. S. Smak, I. A. Tarasov // Arctic and North. – 2020. – № 41. – P. 236–247. DOI: 10.37482/issn2221-2698.2020.41.236.
17. Алексеев С. А. Коммуникативно-информационные технологии в повышении качества городской среды / С. А. Алексеев // Управление устойчивым развитием. – 2019. – № 6(25). – С. 36–40.
18. Мыльников В. А. Моделирование процессов оценки качества инженерной инфраструктуры городского района с использованием нейро-нечетких сетей / В. А. Мыльников, В. В. Бут // Вестник Ивановского государственного энергетического университета. – 2017. – № 3. – С. 75–82. DOI: 10.17588/2072-2672.2017.3.075-082



19. *Cai M.* Natural language processing for urban research: A systematic review / M. Cai // *Heliyon*. – 2021. – Vol. 7, № 3. – e06322. DOI: 10.1016/j.heliyon.2021.e06322
20. *Estévez-Ortiz F.-J.* An application of people's sentiment from social media to smart cities / F.-J. Estévez-Ortiz, A. García-Jiménez, P. Glösekötter // *El Profesional de la Información*. – 2016. – Vol. 25, № 6. – P. 851. DOI: 10.1007/978-3-319-26123-2\_31
21. *Jones K. S.* A statistical interpretation of term specificity and its application in retrieval / K. S. Jones // *Journal of Documentation*. – 2004. – Vol. 60, № 5. – P. 493–502. DOI: 10.1108/00220410410560573
22. *Korobov M.* Morphological Analyzer and Generator for Russian and Ukrainian Languages / M. Korobov // *Communications in Computer and Information Science*. – 2015. – P. 320–32. DOI: 10.1007/978-3-319-26123-2\_31.
23. Distributed Representations of Words and Phrases and their Compositionality / T. Mikolov [et al.] // *Advances in Neural Information Processing Systems*. – 2013. – №26. DOI: 10.48550/arXiv.1310.4546
24. *Mikolov T.* Linguistic Regularities in Continuous Space Word Representations / T. Mikolov, W.-t. Yih, G. Zweig // *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, – 2013. – P. 746–751.
25. *Pennington J.* Glove: Global Vectors for Word Representation / J. Pennington, R. Socher, C. D. Manning // *EMNLP*. – 2014. – Vol. 14. – P. 1532–1543. DOI: 10.3115/v1/D14-1162.
26. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin [et al.] // *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. P. 4171–4186. DOI: 10.18653/v1/N19-1423
27. Deep contextualized word representations / M. E. Peters [et al.] // *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. P. 2227–2237. DOI: 10.18653/v1/N18-1202
28. Multilingual ELMo and the Effects of Corpus Sampling / V. Ravishankar [et al.] // *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*, P. 378–384.

**Галаган Константин Владимирович** — аспирант кафедры цифровых технологий, института цифровой экономики Югорского государственного университета.

E-mail: galagankv@gmail.com

ORCID iD: <https://orcid.org/0000-0003-3587-8334>

## APPLICATION OF NLP NEURAL NETWORK MODELS TO EVALUATE THE EFFECTIVENESS OF MEASURES TO CREATE A COMFORTABLE URBAN ENVIRONMENT

© 2023 K. V. Galagan✉

*Yugra State University  
16, Chekhov Street, 628012 Khanty-Mansiysk,  
Khanty-Mansiysk Autonomous Okrug – Yugra, Russian Federation*

**Annotation.** The formation of a comfortable urban environment is one of the urgent tasks of modern urban planning. To achieve this goal, in 2018 the President and the Government of the Russian Federation launched the national project “Housing and the Urban Environment”, within the framework of which an exhaustive list of measures to create a comfortable urban environment was determined. The purpose of the work is to form a rating of the effectiveness of such events based on the analysis of the texts of messages from citizens, users of social networks, texts of publications in the media and descriptions of events held in a given territory. The proposed approach is based on an assessment of the semantic similarity of the description of events from a set of datasets for a certain territory. On the basis of the datasets formed for specific municipalities, a discrepancy between the expectations of citizens and the implemented activities in this territory was revealed.

**Keywords:** national project, comfortable urban environment, evaluation of events, NLP.

### CONFLICT OF INTEREST

The author declares the absence of obvious and potential conflicts of interest related to the publication of this article.

### REFERENCES

1. Passport of the national project “Housing and Urban Environment”, approved at a meeting of the Presidium of the Council under the President of the Russian Federation for Strategic Development and National Projects on December 24, 2018 (In Russian)

2. Decree of the Government of the Russian Federation of March 23, 2019 No. 510-r (as amended on December 30, 2020) “On Approval of the Methodology for Forming the Urban Environment Quality Index”. (In Russian)

3. Koroleva P. and Bochanov M. (2019) The urban environment quality index as a way to assess the state policy for the formation of a com-

fortable urban environment. *Sbornik nauchnyh statej VI Mezhdunarodnoj nauchnojkonferencii studentov i molodyh uchenyh: v 3 t. T. 1: Ekonomika. Gumanitarnye nauki* [Collection of scientific articles of the VI International Scientific Conference of Students and Young Scientists: in 3 vols. Vol. 1: Economics. Humanities]. Kursk; P. 306–312. (In Russian)

4. Demenev A. G., Shubina T. F., Shubina P. V., Nenashva M. V., Makulin A. V. and Tarasov I. A. (2018) Public participation in planning a comfortable urban environment: case of the Arkhangelsk region *Arctic and North*. (33). P. 91–117. (In Russian) DOI: 10.17238/issn2221-2698.2018.33.91

5. Dmitrieva N. N. and Ipatova T. M. (2018) Formation of comfort city environment – as a strategic direction of development of the project «Housing and city environment». *Social and Economic Management: Theory and Practice*. 1 (32). P. 95–98. (In Russian)

6. Muhametov D. R. (2020) Models of citizen engagement platforms for creating a new generation of smart cities in Russia. *Russian Journal of Innovation Economics*. 10(3). P. 1605–1622. (In Russian)

---

✉ Galagan Konstantin V.  
e-mail: galagankv@gmail.com

7. Wenshu Li, T. Feng, H. Timmermans, Zhi-gang Li, M. Zhang and Bowen Li (2020) Analysis of citizens' motivation and participation intention in urban planning. *Cities*. 106. P. 102921. DOI: 10.1016/j.cities.2020.102921
8. Order of the Ministry of Construction of Russia No. 913/pr dated December 30, 2020 "On approval of guidelines for the involvement of citizens, their associations and other persons in resolving issues of the development of the urban environment". (In Russian)
9. Making the city more comfortable. (2022) <https://za.gorodsreda.ru> (accessed 31.03.2022). (In Russian)
10. Belyakova E. A., Moskvina R. N., Yurova V. S. and Utyugova E. S. (2019) Approaches to Quality Assessment of City Environment. *PGUAS Bulletin: construction, science and education*. 2(9). P. 3–9. (In Russian)
11. Notman O. V. (2021) Index Method for Assessing the Quality of the Urban Environment: International and Russian Experience. *Vestnik Adygejskogo gosudarstvennogo universiteta. Seriya 1: Regionovedenie: filosofiya, istoriya, sociologiya, yurisprudenciya, politologiya, kul'turologiya*. 2(279). P. 89–99. (In Russian)
12. Strigunov V. V. and Bocharova T. A. (2021) Assessment of the State of the City of Khabarovsk by Federal Indices of the Urban Environment and IQ of Cities. *Uchenye zametki TOGU*. 12(1). P. 5–12. (In Russian)
13. Aksenova V. V. (2020) Comfortable Urban Environment: Public Opinion of Muscovites. *Social policy and sociology*. 19(4). P. 76–84. (In Russian) DOI: 10.17922/2071-3665-2020-19-4-76-84
14. Bogdanova L. P. and Glushkova M. A. (2021) Assessment of Quality of the City Environment by Populations of the City Tver. *Herald of Tver State University. Series: Geography and Geoecology*. 2(34). P. 14–24. (In Russian) DOI: 10.26456/2226-7719-2021-2-14-24
15. Dunaeva D. O. (2021) Discursive Practices of Citizens as a Communicative Mechanism for Forming the Image of a Comfortable City (Practical Research Experience). *Tomsk State University Journal Of Philosophy Sociology And Political Science*. (60). P. 137–150. (In Russian) DOI: 10.17223/1998863X/60/13
16. Russova O.N., Smak T. S. and Tarasov I. A. (2020) Assessment of the Comfort of the Urban Environment as a Factor in the Social Well-Being of Citizens of the Arkhangelsk Oblast. *Arctic and North*. (41). P. 236–247. DOI: 10.37482/issn2221-2698.2020.41.236
17. Alekseev S. A. (2019) Communicative Information Technologies to Increase the Quality of the City Environment. *Managing Sustainable Development*. 6(25). P. 36–40. (In Russian)
18. Myl'nikov V. A. and But V. V. (2017) Modeling of urban engineering infrastructure quality assessment processes using neural-fuzzy networks. *Vestnik Ivanovskogo gosudarstvennogo energeticheskogo universiteta*. (3). P. 75–82. (In Russian) DOI: 10.17588/2072-2672.2017.3.075-082
19. Cai M. (2021) Natural language processing for urban research: A systematic review. *Heliyon*. 7(3). e06322. DOI: 10.1016/j.heliyon.2021.e06322
20. Estévez-Ortiz F.-J., Garcí a-Jiménez A. and Glösekötter P. (2016) An application of people's sentiment from social media to smart cities. *El Profesional de la Información*. 25(6). P. 851. DOI: 10.3145/epi.2016.nov.02
21. Spärck Jones K. (2004). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*. 60(5). P. 493–502. DOI: 10.1108/00220410410560573
22. Korobov M. (2015). Morphological Analyzer and Generator for Russian and Ukrainian Languages. *Communications in Computer and Information Science*. Vol 542. DOI: 10.1007/978-3-319-26123-2\_31
23. Mikolov T., Sutskever I., Chen K., Corrado G. S. and Dean J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*. P. 3111–3119.
24. Mikolov T., Yih W.-t. and Zweig G. (2013) Linguistic Regularities in Continuous Space Word Representations. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. P. 746–751.
25. Pennington J., Socher R. and Manning C. D. (2014) Glove: Global Vectors for Word Representation. *EMNLP*. 14. P. 1532–1543. DOI: 10.3115/v1/D14-1162

26. Devlin J., Chang M-W., Lee K. and Toutanova K. (2019) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. P. 4171–4186. DOI: 10.18653/v1/N19-1423
27. Peters M. E., Neumann M., Iyyer M., Gardner M., Clark C., Lee K. and Zettlemoyer L. (2018) Deep contextualized word representations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. P. 2227–2237. DOI: 10.18653/v1/N18-1202
28. Ravishankar V., Kutuzov A., Ovrelid L. and Velldal E. (2021) Multilingual ELMo and the Effects of Corpus Sampling. *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*. P. 378–384.

**Galagan Konstantin V.** — PhD student, Department of Digital Technologies, Institute of Digital Economy, Yugra State University.  
E-mail: galagankv@gmail.com  
ORCID iD: <https://orcid.org/0000-0003-3587-8334>