

Цифровая экономика: теория и практика

Digital economy: theory and practice

Научная статья

УДК 338.46

DOI: <https://doi.org/10.18721/JE.15301>



ЦИФРОВИЗАЦИЯ УСЛУГ ДЛЯ УЛУЧШЕНИЯ КАЧЕСТВА СОЦИАЛЬНОЙ ИНФРАСТРУКТУРЫ

Г.И. Курчеева  , В.Б. Копылов

Новосибирский государственный технический университет,
г. Новосибирск, Российская Федерация

 kurcheeva@yandex.ru

Аннотация. В статье обосновывается возможность привлечения населения на основе процессов цифровизации к участию в улучшении качества городской среды. Работа выполнена в рамках концепции «Цифровой город» или «Умный город», которая направлена на решение многочисленных проблем быстро растущих городов и городов с большим количеством населения. В статье рассмотрены возможности извлечения и оценки мнения населения из неофициальных источников информации, а именно социальных сетей. В настоящее время данный подход применяется крайне ограниченно из-за больших затрат на привлечение групп аналитиков к оценке. В современном мире важнейшую роль играет социальное пространство в сети интернет. Но самыми популярными онлайн ресурсами являются социальные сети. Сентимент-анализ или анализ тональности текста, предназначен для автоматизированного выявления в текстах эмоционально окрашенной лексики и эмоциональной оценки авторов (мнений) по отношению к объектам, речь о которых идёт в тексте. Метод имеет большой потенциал применения для мониторинговых, аналитических и сигнальных систем, для систем документооборота и рекламных платформ, таргетированных по тематике веб-страниц. Процесс мониторинга сбора и обработки мнений населения о качестве жизни начинается с формирования мнения о качестве жизни и городских услугах у гражданина при взаимодействии с городской инфраструктурой будь то общественный транспорт, сфера здравоохранения, образование, жилищно-коммунальное хозяйство, экология и другие. Анализируя данную информацию, можно сформировать лучшую политику по модернизации. Разработка модели анализа эмоциональной тональности позволяет расширить возможности инструментов оценки, применяемых для принятия управленческих решений путем вовлечения населения в вопросы управления городской средой. В результате программной реализации модели сбора и анализа эмоциональной тональности были проанализированы активно обсуждаемые темы по улучшению качества городской среды. Получены варианты оценок. При этом, обработка этих данных, например, анализ настроений, может быть использована для уточнения оценок показателей качества жизни населения.

Ключевые слова: цифровой город, анализ эмоциональной тональности, управление, информационная модель, парсинг

Для цитирования: Курчеева Г.И., Копылов В.Б. Цифровизация услуг для улучшения качества социальной инфраструктуры // П-Economy. 2022. Т. 15, № 3. С. 7–21. DOI: <https://doi.org/10.18721/JE.15301>



DIGITALIZATION OF SERVICES TO IMPROVE THE QUALITY OF SOCIAL INFRASTRUCTURE

G.I. Kurcheeva  , V.B. Kopylov

Novosibirsk State Technical University,
Novosibirsk, Russian Federation

 kurcheeva@yandex.ru

Abstract. The article substantiates the possibility of involving the population to participate in improving the quality of the urban environment in the context of digital transformation. The study was carried out within the framework of the "Digital City" or "Smart City" concept, which is aimed at solving numerous problems of rapidly growing cities and cities with a large population. The article considers the possibilities of collecting and evaluating the opinion of the population from unofficial sources of information, namely social networks. Currently, this approach is used extremely limited due to the high costs of involving groups of analysts in the assessment. In the modern world, the most important role is played by the social space on the Internet. Notably, the most popular online resources are social networks. Sentiment analysis or text tonality analysis is designed for automated identification of emotionally colored vocabulary in texts and emotional evaluation of authors (opinions) in relation to the objects referred to in the text. The method has a great potential for application for monitoring, analytical and signaling systems, for document management systems and advertising platforms targeted by the subject of web pages. The process of monitoring the collection and processing of opinions of the population about the quality of life begins with forming an opinion about the quality of life and urban services of a citizen in interaction with urban infrastructure, be it public transport, healthcare, education, housing and communal services, ecology and others. Analyzing this information, it is possible to form a better modernization policy. A model for emotional tonality analysis model makes it possible to expand the capabilities of assessment tools used for managerial decision-making by involving the population in issues of urban environment management. As a result of the program implementation of the model for collecting and analyzing emotional tonality, actively discussed topics on improving the quality of the urban environment were analyzed. Variants of estimates are obtained. At the same time, processing of these data, for example, sentiment analysis, can yield refined estimates of indicators for the quality of life of the population.

Keywords: digital city, emotional tonality analysis, management, information model, parsing

Citation: G.I. Kurcheeva, V.B. Kopylov, Digitalization services to improve the quality of social infrastructure, *П-Economy*, 15 (3) (2022) 7–21. DOI: <https://doi.org/10.18721/JE.15301>

Введение

Концепция «Цифровой город» или «Умный город» активно применяется при развитии и модернизации городов и поселений по всему миру. Так, например, в Европе выделяют более 240 городов с населением свыше 100 тысяч человек, которые применяют технологии «Умных городов». Большая часть данных городов находится в западной части Европы. Более 30 таких городов имеют Испания, Италия, Англия, от 10 до 30 городов выделяют во Франции, Германии, Швеции и Дании, также наличие «Умных городов» отмечают в меньших количествах и в других странах Европы [1–3].

Концепция «Цифровой город» нашла своё развитие и в городах России. К 2021 году принято несколько программ развития городов, согласно которым должно происходить цифровое развитие городов и общества в целом. Например, проект цифровизации городского хозяйства «Умный город» для населённых пунктов, созданный Министерством строительства и жилищно-коммунального хозяйства России.



Проект направлен на повышение конкурентоспособности российских городов и поселений, формирование эффективной системы управления городским хозяйством, создание безопасных и комфортных условий для жизни граждан. Реализация городских проектов осуществляется в рамках выполнения национального проекта «Жильё и городская среда» и национальной программы «Цифровая экономика». Кроме того, для эффективного ведения проектов был создан Национальный Центр компетенций проекта «Умный город», который будет заниматься разработкой, внедрением и популяризацией технологий, оборудования, программ, направленных на повышение уровня цифровизации городского хозяйства, а также подготовкой и оказанием содействия проектам международного сотрудничества по вопросам жилищной политики, городского развития и управления природными ресурсами, прежде всего касающимся создания и функционирования таких городов [4, 5].

Исследуя проблемы привлечения граждан в городское управление, хотелось бы отметить существующие на данный момент показатели, которые способны оценить уровень вовлечения населения в процессы управления. В 2019 году на основе паспортов национальных проектов была произведена оценка таких целевых показателей.

Рассматривая вопрос вовлечения граждан в городское управление, были выделены такие показатели, как, например, показатели, отражающие непосредственное участие людей в формировании проектов и решений для развития городской инфраструктуры. Этот показатель отражает долю граждан, принявших участие в решении вопросов городской среды от общего количества граждан в возрасте от 14 лет, проживающих в муниципальных образованиях, на территории которых реализуются проекты по созданию комфортной городской среды. На рис. 1 представлены основные значения показателя с прогнозом на будущее.

Базовое значение показателя составляет 5%, кроме того, отмечают, что в 2018 году показатель стал равен в среднем 6% на территории всей Российской Федерации, в 2019 году увеличился до 10,5% [6]. При дальнейшем развитии предполагается повышение показателя в 2024 году до 30%.

Исходя из представленных данных, можно сделать вывод, что показатель привлечения жителей в управление городским хозяйством является низким, и планируется его увеличение в рамках выполнения национальных проектов.

В современном мире важнейшую роль играет социальное пространство в сети интернет. Но самыми популярными онлайн ресурсами являются социальные сети. На сегодняшний день практически любая компания, независимо от предоставляемых ею услуг и выпускаемой продукции, имеет свои страницы в социальных сетях. Для получения таких оценок мнений предпринимаются попытки применить так называемый, сентимент-анализ или анализ тональности текста, предназначен для автоматизированного выявления в текстах эмоционально окрашенной лексики и эмоциональной оценки авторов (мнений) по отношению к объектам, речь о которых идёт в тексте различных сообщений [7, 8].

Цель исследования заключалась в адаптации метода сентимент-анализа к решению проблемы автоматизации процесса оценки вовлечения населения в процессы управления качеством городской среды.

Для выполнения этой цели в работе предлагалось выполнить ряд задач, направленных на изучение и разработку информационной модели, позволяющей оценить вовлечение населения на основе этого метода эмоциональной тональности.

Были поставлены следующие **задачи**:

- 1) обоснование выбора направления вовлечения населения в улучшение качества городской среды в рамках концепции «цифровой город»;
- 2) разработка информационной модели;
- 3) адаптация метода эмоциональной тональности для сбора мнения населения о качестве городской среды;

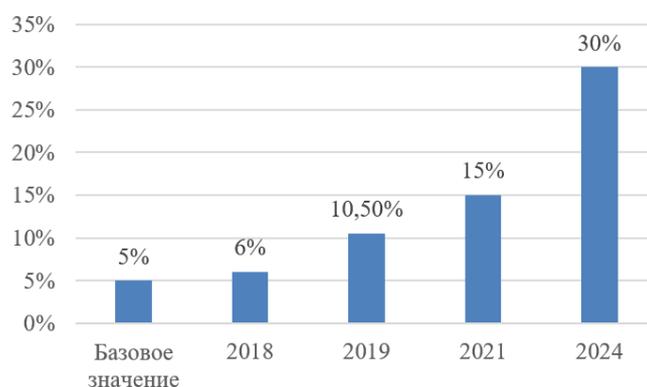


Рис. 1. Доля граждан, принявших участие в решении вопросов городской среды (*Разработано авторами*)
 Fig. 1. The percentage of citizens participating in issues of the urban environment (*Developed by the authors*)

- 4) разработка технологии участия населения в улучшении качества городской среды;
- 5) разработка программной реализации технологии участия населения в улучшении качества городской среды и анализ мнений населения на примере оценки качества транспортного обеспечения.

Литературный обзор

Специалисты (НИИТС – это научно-исследовательский институт технологий и связи, а также консалтинговый центр в области инфокоммуникационных технологий, г. Москва) рассматривают этот метод сентимент-анализа, как инструмент, имеющий большой потенциал применения для мониторинговых, аналитических и сигнальных систем, для систем документооборота и рекламных платформ, таргетированных по тематике веб-страниц. Они определяют понятие «Умный город» как системный подход к использованию информационных технологий на основе анализа данных для предоставления услуг по управлению природными, энергетическими и городскими ресурсами, способствующих устойчивому экономическому развитию и обеспечению высоких стандартов жизни [9]. В другом исследовании рассмотрено понятие «Умный город» как взаимосвязь информационных и коммуникационных технологий для управления городским имуществом, которые объединяют все сферы жизни города: здравоохранение, транспорт, энергетика, инфраструктура, экология [10].

Используя данную концепцию развития города, достигаются некоторые преимущества, по сравнению с городами активно не использующими высокие информационные технологии для управления.

Американский экономист и политолог Энтони Даунс в своей работе отмечает, что внедрение информационных технологий в городское управление и сам «умный» рост является естественной реакцией, которая следует за горизонтальным расширением города. того, чтобы повысить её эффективность функционирования [11].

Рассматривая концепцию «умного города» с точки зрения технологической структуры, можно выделить четыре базовых элемента «умного города», а именно интернет вещей, инфраструктуру передачи данных, системы анализа данных и систему агрегации и унификации данных [12].

На сегодняшний день существует множество компаний, которые целенаправленно занимаются развитиями технологий для умного города. [13].

Статьи новостных сайтов, посты в социальных сетях и сообщения на государственных форумах – это неофициальные источники информации, которые, как правило, содержат в себе мнение населения об уровне развития городских услуг и качества жизни в целом.

Для автоматического сбора информации данных ресурсов предполагается использование парсеров, которые позволят организовать автоматический сбор записей социальных сетей, форумов и новостных сайтов в одном месте, для их дальнейшей обработки [14, 15].

Как правило, автоматический сбор неструктурированной информации из различных интернет-источников представляет собой некий процесс обработки информации, состоящий из трёх этапов: консолидация данных, трансформация данных, очистка данных [16, 17].

Вся информация в сети Интернет представлена в виде большого количества интернет-страниц, которые, как правило, имеют индивидуальную архитектуру представления данных. Разрабатываемый веб-парсер, представляет собой программу или скрипт, который позволяет производить загрузку различного контента из веб-страниц, которые представлены в формате HTML [18].

Методы и материалы

На основании систематизации возникающих проблем при работе с такими порталами, а также на основании изучения отдельных зарубежных практик была предложена информационная модель, осуществляющая возможность привлечения населения к управлению, благодаря мониторингу социальных сетей и СМИ с целью получения мнений о различных вопросах городского хозяйства, а также привлечению открытых данных для повышения показателя оценки уровня открытости государственных данных.

В результате предполагается использовать информационную модель, схема работы которой представлена на рис. 2.

Основным отличием от имеющейся модели работы можно выделить мониторинг социальных сетей и СМИ, который позволит производить сбор и обработку информации о мнении населения касательно тех или иных вопросов городского управления и удовлетворённости качеством жизни и городскими услугами. Собираемые данные формируются в виде структу-

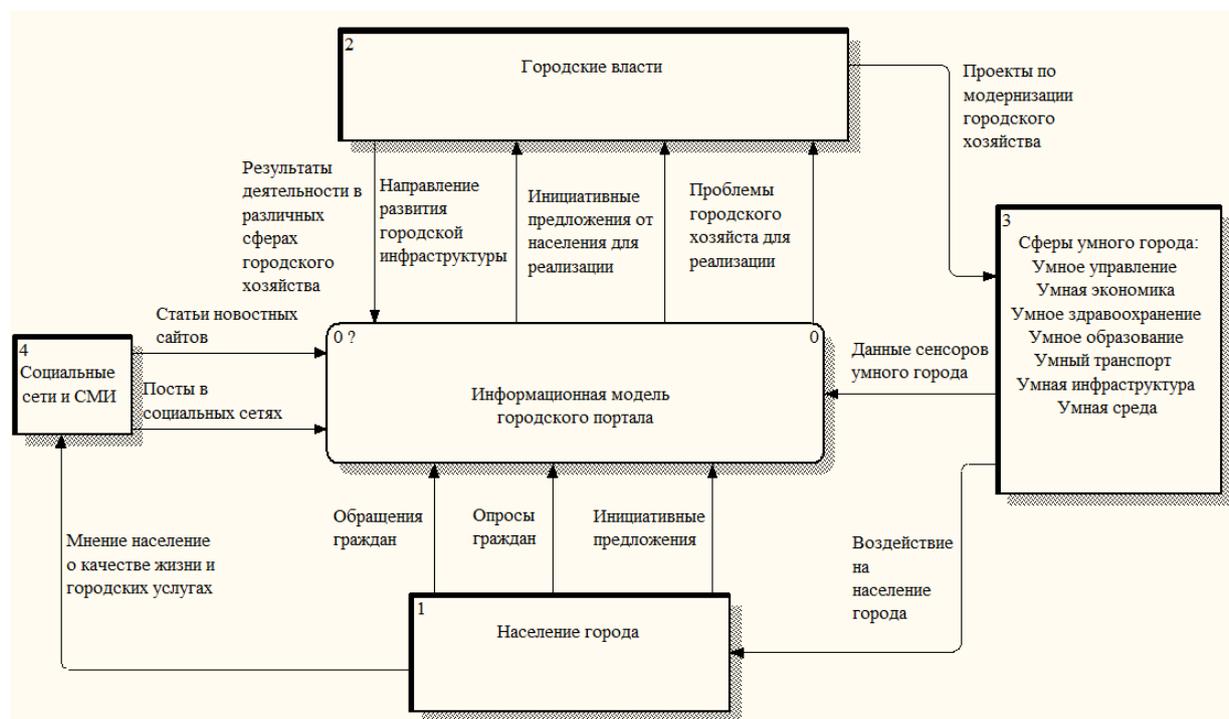


Рис. 2. Модель работы информационного портала TO-BE

Fig. 2. The model of the TO-BE information portal

рированных показателей, которые используются при формировании проектов модернизации городского хозяйства.

Тем временем администрация города предоставляет полную информацию, касаемо показателей города и документов, содержащих в себе планы по развитию городского хозяйства.

В ходе работы информационного портала происходит сбор показателей удовлетворённости населения городскими услугами и качеством жизни, показателей от сенсоров умного города, а также показателей деятельности в различных сферах городского хозяйства, которые формируют единую систему показателей качества жизни и городских услуг. На основании системы показателей качества жизни и городских услуг городскими властями формируется стратегия по модернизации тех или иных сфер городского хозяйства.

Благодаря предлагаемому функционалу, предоставляется возможность повысить степень вовлечения жителей города в управление городским хозяйством, а также позволить администрации повысить открытость и доступность информации. Наибольшую сложность для обработки и сбора составляют неофициальные источники. Ввиду того, что эти источники информации представляют собой массивы текстов на естественном языке, то требуется интеллектуальный анализ текстов, для сбора мнений. На сегодняшний момент было определено, что лучше всего для выполнения поставленной задачи будет справляться анализ эмоциональной тональности, который набирает свою популярность в сфере маркетинга для оценки продукта, путём оценки отзывов клиентов.

Результаты исследования. Анализ и оценка методов автоматизированного выявления в текстах эмоционально окрашенной лексики

Совокупность операций рассматриваемого процесса позволяет выполнить консолидацию данных, которая представляет собой комплекс методов и процедур, направленных на получение данных из различных источников, а также сохранение их в хранилище данных или аналитическую систему.

Главным фактором для проведения этого этапа является выбор источника информации, который содержит в себе требуемую для реализации определённого проекта информацию.

Документ формата HTML можно сравнить с обычным файлом любого текстового формата. Различием между ними является то, что первый имеет в своём составе строки, которые представляют собой управляющие команды, которые в последующем использовании формируют структуру документа. При автоматическом сборе информации из сети происходит анализ и обработка именно таких документов HTML¹ [13].

Работа парсера начинается с загрузки объектов содержимого HTML файла передачи его для дальнейшего извлечения. Полученный HTML файл обрабатывается с помощью синтаксического анализатора, который позволяет получить необходимые данные из файла.

Существует различные способы обработки HTML файлов:

– регулярные выражения – совокупность созданных регулярных выражений позволяют производить поиск по какому-либо шаблону во время обработки текста HTML файла. Способ зачастую используется для выполнения несложных задач, например, извлечение электронных адресов на веб-странице. Но в тоже время не подходит для извлечения уже более сложной информации, которая содержится в различных полях на странице;

– анализ HTML – способ наиболее часто используется при обработке и сбору данных с веб-страниц. Большая часть веб-сайтов опирается на некую базу данных из которой происходит представление контента на странице по определённому шаблону. Анализатор превращает код HTML в иерархическую структуру, по которой можно перемещаться программно, используя различные языки запросов или CSS-селекторы для извлечения данных;

¹ Автоматическая обработка текстов на естественном языке и анализ данных: учеб. пособие / Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. — М.: Изд-во НИУ ВШЭ, 2017. — 269 с.



– анализ DOM – так как развитие Интернета привело к созданию большого количество сложных веб-приложений, которые зависят от многих программных компонентов, то простой загрузки веб-страницы и кода HTML становится недостаточно. Такие страницы динамически обновляют данные внутри браузера, не отправляя пользователя на новую страницу. При этом загружая HTML файл страницы, пользователь получает только внешнюю HTML оболочку веб-приложения. Данный файл предоставляет только часть ссылок приложения и неполный контент, и данные. Для анализа подобных веб-страниц необходимо воспользоваться с помощью селектора DOM:

– автоматическое извлечение с использованием искусственного интеллекта – этот способ извлечения информации является наиболее сложным среди всех, в основном он используется для анализа нескольких сайтов, которые имеют схожую структуру. При этом, веб-парсеры обучаются с помощью машинного обучения по извлечению данных [14, 15].

Кроме описанных выше способов парсинга различных веб-страниц существует так называемый API. Данный термин представляет собой установку функций и правил, позволяющих взаимодействовать между различным программным обеспечением, которое предоставляет API и программными компонентами, которые используют API для получения данных первых. В веб-разработке API обычно представляет собой набор элементов кода (например, методов, свойств, событий и URL), которые разработчик может использовать в своих приложениях для взаимодействия с компонентами веб-браузера пользователя или другим программным/аппаратным обеспечением на компьютере пользователя, а также сторонними веб-сайтами и сервисами [16].

Способ хорошо подходит для сбора информации из социальной сети «ВКонтакте», так как данный источник имеет большой перечень методов взаимодействия с помощью API.

Сбор информации включает необходимость трансформация данных для представления ее в таком виде, чтобы она могла быть использована наиболее эффективно. Трансформация – комплекс методов и алгоритмов, направленных на оптимизацию представления и форматов данных с точки зрения решаемых задач и целей анализа. Трансформация данных не ставит целью изменить информационное содержание данных [17].

В заключение процесса сбора информации обязательным этапом является очистка информации для получения данных в формате, позволяющем производить дальнейший анализ с целью получения необходимых показателей и данных. Сам процесс очистки принято определять, как – процесс выявления и исключения различных факторов, мешающих корректному анализу данных: аномальных и фиктивных значений, пропусков, дубликатов и противоречий, шумов [18].

Ввиду того, что полученные в ходе автоматического сбора данные представляют собой неструктурированные записи, то возникают проблемы с их обработкой. Ручная обработка большого массива данных требует больших трудовых ресурсов, что делает работу сайта нецелесообразной. Поэтому для получения необходимой информации необходимо перейти к интеллектуальному анализу текстов. Для выполнения, задачи обычно принято использовать какие-либо регулярные выражения, операции со строками и методы поиска.

Автоматический анализ тональности текстов, а также выявление мнения авторов текста касательно предмета описания, является одной из самых активно развивающихся технологий в сфере компьютерной лингвистики. Во многом это связано с развитием социальных сетей, онлайн сервисов с отзывами многих людей и других инструментах.

Зачастую анализ тональности используется при определении популярности или репутации той или иной компании или же продукта. В данном случае происходит мониторинг позитивных и негативных отзывов на различных информационных ресурсах.

В сфере финансовых рынков анализ отчетов и новостей позволяет определить тренды на фондовом и валютном рынке. Тональность упоминания терминов в научных статьях позволяет определить наиболее важные понятия и научные тренды.

Также растёт популярность автоматических методов определения эмоциональной тональности сообщений в социальных сетях для политических и социальных исследований, позволяя определить предпочтения и результаты будущих выборов.

Автоматический анализ тональности текстов, написанных на естественном языке, как правило, осуществляется с помощью нескольких типов подходов:

- лингвистический – подход, основанный на словарях и правилах;
- статистический – подход на основе машинного обучения;
- смешанный – совместное использование лингвистического и статистического подхода [19].

На данный момент создаются смешанные методы, которые позволяют интегрировать тональные словари оценочной лексики в методы машинного обучения. Разработанная информационная модель позволяет собрать комментарии по конкретному вопросу городского хозяйства и провести анализ эмоциональной тональности, в результате которого, могут быть сформированы оперативные данные для короткого периода, используемые для выделения актуальных проблем, а также стратегические данные для длительного периода для формирования показателей качества жизни, используемых для оценки городского развития.

Для реализации данной технологии был выбран язык python, благодаря обработке HTML документов, а также методов API ВКонтакте были собраны комментарии к тем или иным вопросам. С помощью библиотеки Dostoevsky, которая предназначена для анализа эмоциональной тональности и была обучена на массиве комментариев из социальных сетей был произведён анализ тональности с выделением отношения пользователей к тому или иному вопросу [20, 21].

Для возможной демонстрации полученных результатов с помощью фреймворка Django был разработан раздел обработки неофициальных источников, который может быть использован на различных порталах. В итоге, с помощью технологии привлечения населения на основе анализа эмоциональной тональности были получены различные результаты для сферы транспорта в городе Новосибирск, которые позволяют оценить удовлетворение населением услугами сферы транспорта в целом за различные периоды, а также оценить отношение людей с различными характеристиками, в том числе была произведена оценка отношения людей различных возрастных групп. Также для определения актуальных задач одна из тем сферы транспорта была разделена на конкретные вопросы дорожного хозяйства, в результате чего было определена различная актуальность при различных временных периодах.

Результаты исследования. Анализ тональности рассматриваемых примеров

Рассматривая города России и анализируя показатели привлечения населения к управлению, был сделан вывод о низком уровне вовлечения, которых находится на начальных этапах развития. Для возможности развития этого направления в стране предлагается использовать информационные модели, позволяющие собирать воедино различные источники информации для принятия управленческого решения, при условии, что будут учитываться оценки, получаемые из источников, отражающих мнение населения.

В ходе работы были изучены существующие информационные модели, которые обычно применяются в управлении для привлечения населения. Как правило, модели представлены в виде информационных порталов, которые позволяют населению формировать обращения о проблемах, а также проходить опросы о качестве тех или иных вопросах городского хозяйства.

Важным функционалом для привлечения населения к управлению городским хозяйством являются анкетированные опросы населения. Данные, получаемые из опросов, аналогично неофициальным источникам информации, позволяют оценить удовлетворённость и отношение населения к тем или иным городским процессам и проблемам.

При дистанционном заполнении формы на сайте, во время которой необходимо указать, где человек столкнулся с проблемой, к какой сфере городского хозяйства относится проблема



и суть проблемы с возможностью добавления изображений. После заполнения данной формы обращения распределяются по ответственным инстанциям для проверки и исправления, согласно определённой в форме сферы городского хозяйства, будь то транспорт, жилищно-коммунальное хозяйство и другие.

Данный функционал играет важную роль в работе городского управления в виде актуального перечня проблем для исправления с целью повышения качества жизни населения в виде дополнительного способа привлечения населения к управлению городским хозяйством, кроме того, анализируя полученные обращения можно сформировать дополнительные показатели системы показателей качества жизни и городских услуг, отражающие наиболее часто встречающиеся проблемы для формирования будущей политики улучшения городского хозяйства.

Также способом привлечения населения к управлению городским хозяйством является предоставление возможности формирования инициативных предложений населением по улучшению городской среды (рис. 3).

Процесс мониторинга сбора и обработки мнений населения о качестве жизни начинается с формирования мнения о качестве жизни и городских услугах у гражданина при взаимодействии с городской инфраструктурой будь то общественный транспорт, сфера здравоохранения, образование, жилищно-коммунальное хозяйство, экология и другие. Такое мнение является наиболее релевантным для муниципальных органов, так как оно отражает фактический уровень состояния городской среды. Анализируя полученную информацию, можно сформировать лучшую политику по модернизации. Сформированное мнение граждан может оставить в виде записи или комментария в любом источнике неофициальной информации, например, социальные сети, СМИ, форумы и другие.

Затем, информационная система городского информационного портала автоматизированным способом, с помощью различных парсеров, позволяет собрать мнение гражданина по поводу какого-либо вопроса, которое он оставил в социальных сетях и других источников, в виде массива текстовых комментариев и записей.

Обработка массива записей предполагается с помощью средств интеллектуального анализа данных, а именно с помощью анализа эмоциональной тональности для получения какого-либо показателя для его использования в системе показателей качества жизни и городских услуг для формирования управленческого решения по модернизации городского хозяйства.

Взяв за основу описание методов автоматизированного сбора и обработки информации, из социальных сетей выделены основные процессы работы модели. На основании параметров, которые определены для решения поставленной задачи, происходит сбор постов из определённого сообщества социальной сети «ВКонтакте».

Далее происходит поиск ключевых слов в тексте собранных постов для соотнесения их к определённой сфере городского хозяйства для дальнейшего сбора комментариев. Если в посте нет ключевых слов, то его обработка заканчивается. В случае если пост содержит ключевые слова, то происходит сбор комментариев для дальнейшей работы. Так как при автоматизированном сборе текстовой информации могут присутствовать различные шумы, то проверяется необходимость обработки комментариев для избегания искажения результатов анализа эмоциональной тональности.

Если такой необходимости нет, то комментарий анализируется и для него выставляется показатель эмоциональной тональности, в случае если необходима обработка, то, в первую очередь, удаляются шумы, а затем происходит анализ эмоциональной тональности.

Для выбора темы обсуждения проанализировали по 1000 записей интернет-сообществ «АСТ-54» и «Новосибирск» из которых выбраны записи тематики дорожного хозяйства. Результат представлен в табл. 1.

В табл. 1 представлены количество постов, содержащих данную тему обсуждения, количество комментариев для этих постов, а также количество уникальных пользователей, оставляю-

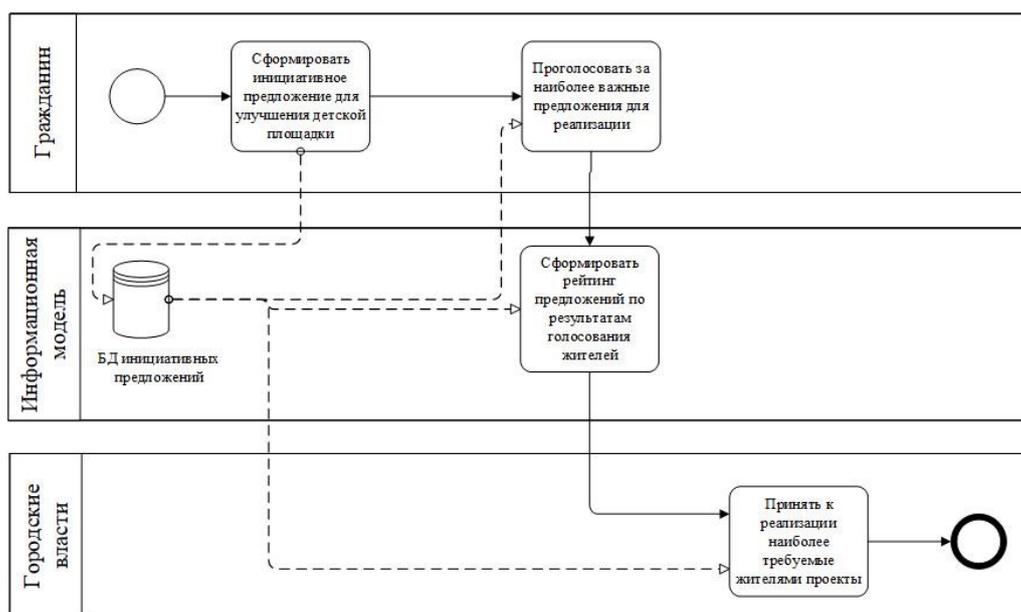


Рис. 3. Процесс сбора инициативных предложений
 Fig. 3. The process of collecting initiative proposals

Таблица 1. Анализ сообществ социальных сетей в разрезе тем обсуждения
Table 1. Analysis of social network communities by discussion topics

	Кол-во постов	Кол-во комментариев	Кол-во людей	Положительное мнение		Негативное мнение		Нейтральное мнение	
				Кол-во	Доля %	Кол-во	Доля %	Кол-во	Доля %
Общественный транспорт	102	10770	4511	581	5	2107	19	7497	70
Дорожное хозяйство	172	18634	6928	1044	6	624	20	12947	70
Пешеходы	86	9504	4030	513	5	1792	18	67	72

щих своё мнение по вопросам, отражённым в тексте поста. Кроме, того в таблице представлена информация, полученная непосредственно в ходе анализа эмоциональной тональности.

Можно заметить, что большая часть постов, а именно 172 затрагивают вопросы дорожного хозяйства, в итоге происходит анализ 18634 комментариев, что позволяет оценить мнение 6928 людей. Значительно меньше обсуждений происходит в вопросах общественного транспорта, а именно 102 записи, к которым были собраны 10770 комментариев от 4511 человек. Наименьшую популярность имеет тема, затрагивающая вопросы пешеходов. Эта тема рассматривается в 86 записях, и обсуждается в 9504 комментариях от 4030 человек.

Статистика позволяет сделать вывод, что наибольшее количество вопросов и проблем замечено в теме дорожного хозяйства, что подтверждается значительным перевесом значений собранных показателей. Рассматривая результаты проведённого анализа эмоциональной тональности, был сделан вывод, что подавляющее большинство комментариев имеет нейтральный тон, меньшее количество комментариев имеют негативный тон и меньше всего комментариев с позитивным эмоциональным тоном.

Рассматривая общие результаты анализа эмоциональной тональности, был сделан вывод, что наибольшая доля негативных комментариев у вопроса уборки дорог – 21%, наименьшее количество негативных комментариев у вопроса дорожной инфраструктуры 17%.

Для выборки, собранной за 8 месяцев, была произведена оценка поднимаемых вопросов для периода времени, что позволит определить для них актуальное время обсуждения. В табл. 2 представлен результат анализа актуальность поднимаемых вопросов в сообществах социальных сетей. В данной табл. 2 сокращения «К.П» – количество постов, в которых рассмотрены данные вопросы, «К.К» – количество комментариев для выбранного периода и темы, «К.Л» – количество людей, участвующих в обсуждении.

Таблица 2. Результат анализа актуальности рассматриваемых вопросов
Table 2. The result of the analysis of the relevance of the issues under consideration

Период (за месяц)	Состояние дорог			Заторы на дорогах			Дорожная инфраструктура			Уборка дорог		
	К.П.	К.К.	К.Л.	К.П.	К.К.	К.Л.	К.П.	К.К.	К.Л.	К.П.	К.К.	К.Л.
05.2021	21	2004	1125	13	1315	758	18	2566	1141	5	438	309
04.2021	36	4380	2365	11	1093	616	30	2937	1560	14	1614	994
03.2021	19	1984	1145	26	3268	1720	37	5458	2442	67	8901	3944
02.2021	18	1655	908	24	3463	1677	28	4316	1913	52	6449	2698
01.2021	14	1568	766	23	3094	1377	28	2711	1408	34	4699	2196
12.2020	11	1074	567	21	2706	1203	29	2811	1441	37	4409	2047
11.2020	19	1926	1095	9	339	243	25	2540	1227	23	2540	1310
10.2020	32	3154	1617	15	1486	717	32	2863	1316	15	1201	775

Вопросы состояния дорог имели большую популярность в октябре 2020 года, а также в апреле 2021 года. В табл. 3 представлены результаты анализа эмоциональной тональности по вопросам дорожного хозяйства с распределением по месяцам с октября 2020 года по май 2021 года.

Таблица 3. Результат анализа тональности рассматриваемых вопросов
Table 3. The result of the analysis of the tonality of the issues under consideration

Период (за месяц)	Состояние дорог			Заторы на дорогах			Дорожная инфраструктура			Уборка дорог		
	П. %	Нег. %	Ней. %	П. %	Нег. %	Ней. %	П. %	Нег. %	Ней. %	П. %	Нег. %	Ней. %
05.2021	7	20	68	7	24	64	5	16	75	7	19	68
04.2021	5	21	70	5	20	70	5	18	71	6	20	66
03.2021	5	20	70	7	18	72	5	18	72	5	20	70
02.2021	5	18	72	4	20	70	6	17	71	5	21	69
01.2021	7	15	72	6	18	72	5	17	75	6	21	70
12.2020	6	14	76	6	16	72	6	17	70	6	21	67
11.2020	7	17	70	8	17	66	6	16	73	8	22	65
10.2020	7	17	70	4	20	70	4	16	75	9	15	70

В таблице представлена информация о доли каждого вида мнений для рассматриваемых вопросов, «П.» – позитивные комментарии, «Нег.» – негативные комментарии, «Ней» – нейтральные комментарии. По результатам можно сделать вывод, что наибольшая доля негативных комментариев для вопроса состояния дорог, который включает в себя качество дорожного покрытия, а именно наличие ям, выбоин и прочего, составляет 21% в апреле 2021 года, как раз в тот период, в котором прослеживается максимальная актуальность. Для вопроса заторы на дорогах максималь-

ная доля негативных комментариев прослеживается в мае 24%, апреле, феврале и октябре – 20%. Для вопроса дорожной инфраструктуры самое большое значение доли негативных комментариев в апреле и марте 2021 года. Рассматривая вопрос уборки дорог, большая доля негативных комментариев заметно прослеживается в период наибольшей актуальности данного вопроса с ноября 2020 года по февраль 2021 года 22–21%.

Заключение

В соответствии с поставленной целью выполнены поставленные задачи:

- 1) обоснован выбор направления вовлечения населения в улучшение качества городской среды в рамках концепции «цифровой город»;
- 2) разработана информационная модель;
- 3) выполнена адаптация метода эмоциональной тональности для сбора мнения населения о качестве городской среды;
- 4) разработана технология участия населения в улучшении качества городской среды;
- 5) разработана программная реализация технологии участия населения в улучшении качества городской среды и выполнен анализ мнений населения на примере оценки качества транспортного обеспечения.

Роль населения в улучшении качества жизни растет, особенно с развитием концепций «цифрового города», поэтому адаптация и развитие самого метода оценки эмоциональной тональности, может быть востребовано во всех направлениях, а именно, «цифрового управления», «умного образования», «умной экологии» и других.

В ходе анализа социальных сетей были выделены основные темы, которые обсуждаются в анализируемых сообществах социальных сетей для сферы транспорта. К таким темам были отнесены: вопросы общественного транспорта, дорожное хозяйство, проблемы пешеходов. Результаты анализа показали, что появляется возможность оперативного реагирования на возникающие проблемы.

Разработанная модель может стать базисом для развития такого инструмента, который учитывает специфику русского языка, и настроен на работу с определенным брендом и социальными сетями. В результате работы программной реализации модели сбора и анализа эмоциональной тональности были проанализированы выделенные темы.

Выполненная декомпозиция на конкретные задачи для сбора и анализа эмоциональной тональности позволяет органам управления получать оперативные данные по актуальным вопросам для их оперативного решения, что является важной задачей улучшения качества жизни населения.

Направление развития концепции

Выполненное исследование в целом направлено на улучшение качества жизни населения. С позиции цифровизации услуг, интерес представляет не только привлечение населения к оценке государственных услуг, но и дальнейшее расширение области применения метода в сфере развития конкурентоспособного бизнеса.

СПИСОК ИСТОЧНИКОВ

1. **Aletdinova A.A., Bakaev M.** Cultural and environmental factors promoting innovative activities in digital economy: the comparative studies between South Korea and Russia // Korea and the World Economy. 2018. Vol. 19, № 2. Pp. 169–191. http://www.akes.or.kr/wp-content/uploads/2018/09/3_MS180526_Page_Proof.pdf



2. Курчеева Г.И., Хворостов В.А. Результативность бенчмаркинга веб-сайтов // Экономический анализ: теория и практика. 2014. № 10. С. 23–28.
3. Kurcheeva G.I., Klochkov G.A. Smart rural settlement: digitalizing the quality of life // IOP Conference Series: Earth and Environmental Science. – 2021. – Vol. 666 : All-Russian conference with international participation economic and phytosanitary rationale for the introduction of feed plants, Bol'shie Vyazemy, Moscow Oblast, 10–11 June 2020. – Art. 032040 (6 p.). – DOI: 10.1088/1755-1315/666/3/032040
4. Bakaev M.A., Khvorostov V.A. Component-based engineering of web user interface designs for evolutionary optimization // 19 IEEE/ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD): proc., Korea, Busan, 27–29 June 2018. – IEEE, 2018, – Pp. 335–340. – ISBN: 978-153865889-5. – DOI: 10.1109/SNPD.2018.8441135
5. ЕМИСС Государственная статистика [Электронный ресурс] Режим доступа: <https://www.fedstat.ru/indicator/59148> (Дата обращения: 10.09.21)
6. Национальные проекты: целевые показатели и основные результаты [Электронный ресурс] Режим доступа: <http://government.ru/news/35675/> (Дата обращения: 10.08.21)
7. Парсинг сайтов [Электронный ресурс] Режим доступа: <https://msiter.ru/articles/chto-takoe-parsing-saytov> (Дата обращения: 12.08.21)
8. API [Электронный ресурс] Режим доступа: <https://developer.mozilla.org/ru/docs/Glossary/API> (Дата обращения: 12.08.21)
9. Индикаторы умных городов НИИТС 2017 [Электронный ресурс] Режим доступа: <http://niitc.ru/publications/SmartCities.pdf>
10. «Умные города», или Smart Cities = Happy Citizens [Электронный ресурс] Режим доступа: <https://habr.com/ru/company/huawei/blog/323382/>
11. Downs A., Smart G. Why we Discuss It More than We Do It. Journal of the American Planning Association, Vol. 71, No. 4, Autumn 2005.
12. Smart City в России: быть ли «умным городам»? [Электронный ресурс] Режим доступа: <http://strategyjournal.ru/articles/smart-city-v-rossii-byt-li-umnym-gorodam/>
13. Концепция умных городов и ее реализация. Опыт компании Cisco [Электронный ресурс] Режим доступа: https://www.cisco.com/c/dam/m/ru_ru/internet-of-everything-ioe/iac/assets/pdfs/celc-iggirkin-smartcity-overview.pdf
14. Романов А.С., Васильева М.И., Куртукова А.В., Мещеряков Р.В. Анализ тональности текста с использованием методов машинного обучения / Proceedings 2nd International Conference "R. PI-OTROWSKI'S READINGS LE & AL'2017" (Saint-Petersburg, 2017). Санкт Петербург: М. Jeusfeld c/o Redaktion Sun SITE, Informatik V, 2018. С. 86–95.
15. Байнова М.С., Соколов А.М. Инструменты автоматизированного сбора и анализа социологической информации о территориальной идентичности жителей крупных городов // Прикладная информатика. 2021. Т. 16. № 2. С. 92–102. DOI: 10.37791/2687-0649-2021-16-2-92-102
16. Акжолов Р.К., Верига А.В. Анализ тональности текста // Международный научный журнал Вестник науки № 9 (18) том 4. ISSN: 2712-8849. С. 29–31. 2019 г. // Электронный ресурс: <https://www.вестник-науки.pf/article/2084> (дата обращения: 22.10.2021 г.)
17. Bakaev M.A., Khvorostov V.A., Heil S., Gaedke M. Web intelligence linked open data for web-site design reuse // Lecture Notes in Computer Science. – 2017. – Vol. 10360. – Pp. 370–377. – Web Engineering : proc., 17 international conference, ICWE 2017, Italy, Rome, 5-8 June 2017. – DOI: 10.1007/978-3-319-60131-1_22
18. Apishev M., Koltcov S., Koltsova O., Nikolenko S., Vorontsov K. Additiveregularization for topic modeling in sociological studies of user-generated text content // MICA 2016, 15th Mexican International Conference on Artificial Intelligence. – Vol. 10061 – Springer, Lecture Notes in Artificial Intelligence, 2016. – Pp. 166–181.
19. Nikolenko S.I., Koltcov S., Koltsova O. Topic modelling for qualitative studies // Journal of Information Science. – 2017. – Vol. 43, no. 1. – Pp. 88–102
20. Rubin T. N., Chambers A., Smyth P., Steyvers M. Statistical topic models for multi label document classification // Machine Learning. 2012. Vol. 88, no. 1-2. Pp. 157–208.
21. Семина Т.А. Анализ тональности текста: современные подходы и существующие проблемы // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 6, Языкознание: Реферативный журнал. – 2020. – № 4. С. 23–27.

REFERENCES

1. **A.A. Aletdinova, M. Bakaev**, Cultural and environmental factors promoting innovative activities in digital economy: the comparative studies between South Korea and Russia // *Korea and the World Economy*. 2018. Vol. 19, № 2. Pp. 169–191. http://www.akes.or.kr/wp-content/uploads/2018/09/3_MS180526_Page_Proof.pdf
2. **G.I. Kurcheeva, V.A. Hvorostov**, Rezul'tativnost' benchmarkinga veb-sajtov // *Ekonomicheskij analiz: teoriya i praktika*. 2014. № 10. S. 23–28.
3. **G.I. Kurcheeva, G.A. Klochkov**, Smart rural settlement: digitalizing the quality of life // *IOP Conference Series: Earth and Environmental Science*. – 2021. – Vol. 666: All-Russian conference with international participation economic and phytosanitary rationale for the introduction of feed plants, Bol'shie Vyazemy, Moscow Oblast, 10–11 June 2020. – Art. 032040 (6 p.). – DOI: 10.1088/1755-1315/666/3/032040
4. **M.A. Bakaev, V.A. Khvorostov**, Component-based engineering of web user interface designs for evolutionary optimization // 19 IEEE/ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD): proc., Korea, Busan, 27–29 June 2018. – IEEE, 2018, – Pp. 335–340. – ISBN: 978-153865889-5. – DOI: 10.1109/SNPD.2018.8441135
5. EMISS Gosudarstvennaya statistika [Elektronnyj resurs] Rezhim dostupa: <https://www.fedstat.ru/indicator/59148> (Data obrashcheniya: 10.09.21)
6. Nacional'nye proekty: celevye pokazateli i osnovnye rezul'taty [Elektronnyj resurs] Rezhim dostupa: <http://government.ru/news/35675/> (Data obrashcheniya: 10.08.21)
7. Parsing sajtov [Elektronnyj resurs] Rezhim dostupa: <https://msiter.ru/articles/chto-takoe-parsing-sajtov> (Data obrashcheniya: 12.08.21)
8. API [Elektronnyj resurs] Rezhim dostupa: <https://developer.mozilla.org/ru/docs/Glossary/API> (Data obrashcheniya: 12.08.21)
9. Indikatory umnyh gorodov NIITS 2017 [Elektronnyj resurs] Rezhim dostupa: <http://niitc.ru/publications/SmartCities.pdf>
10. «Umnye goroda», ili Smart Cities = Happy Citizens [Elektronnyj resurs] Rezhim dostupa: <https://habr.com/ru/company/huawei/blog/323382/>
11. **A. Downs, G. Smart**, Why we Discuss It More than We Do It. *Journal of the American Planning Association*, Vol. 71, No. 4, Autumn 2005.
12. Smart City v Rossii: byt' li «umnym gorodom»? [Elektronnyj resurs] Rezhim dostupa: <http://strategyjournal.ru/articles/smart-city-v-rossii-byt-li-umnym-gorodom/>
13. Koncepciya umnyh gorodov i ee realizaciya. Opyt kompanii Cisco [Elektronnyj resurs] Rezhim dostupa: https://www.cisco.com/c/dam/m/ru_ru/internet-of-everything-ioe/iac/assets/pdfs/celc-iggirkin-smartcity-overview.pdf
14. **A.S. Romanov, M.I. Vasil'eva, A.V. Kurtukova, R.V. Meshcheryakov**, Analiz tonal'nosti teksta s ispol'zovaniem metodov mashinnogo obucheniya / Proceedings 2nd International Conference "R. PI-OTROWSKI'S READINGS LE & AL'2017" (Saint-Petersburg, 2017). Sankt Peterburg: M. Jeusfeld c/o Redakction Sun SITE, Informatik V, 2018. S. 86–95.
15. **M.S. Bajnova, A.M. Sokolov**, Instrumenty avtomatizirovannogo sbora i analiza sociologicheskoy informacii o territorial'noj identichnosti zhitel'ev krupnyh gorodov // *Prikladnaya informatika*. 2021. T. 16. No. 2. S. 92–102. DOI: 10.37791/2687-0649-2021-16-2-92-102
16. **R.K. Akzholov, A.V. Veriga**, Analiz tonal'nosti teksta // *Mezhdunarodnyj nauchnyj zhurnal Vestnik nauki* № 9 (18) tom 4. ISSN: 2712-8849. S. 29–31. 2019 g. // Elektronnyj resurs: <https://www.vestnik-nauki.rf/article/2084> (data obrashcheniya: 22.10.2021 g.)
17. **M.A. Bakaev, V.A. Khvorostov, S. Heil, M. Gaedke**, Web intelligence linked open data for website design reuse // *Lecture Notes in Computer Science*. – 2017. – Vol. 10360. – Pp. 370–377. – *Web Engineering : proc.*, 17 international conference, ICWE 2017, Italy, Rome, 5-8 June 2017. – DOI: 10.1007/978-3-319-60131-1_22
18. **M. Apishev, S. Koltcov, O. Koltsova, S. Nikolenko, K. Vorontsov**, Additiveregularization for topic modeling in sociological studies of user-generated text content // *MICAI 2016, 15th Mexican International Conference on Artificial Intelligence*. – Vol. 10061 – Springer, *Lecture Notes in Artificial Intelligence*, 2016. – Pp. 166–181.
19. **S.I. Nikolenko, S. Koltcov, O. Koltsova**, Topic modelling for qualitative studies // *Journal of Information Science*. – 2017. – Vol. 43, no. 1. – Pp. 88–102.



20. **T.N. Rubin, A. Chambers, P. Smyth, M. Steyvers**, Statistical topic models for multi label document classification // Machine Learning. 2012. Vol. 88, no. 1-2. Pp. 157–208.

21. **T.A. Semina**, Analiz tonal'nosti teksta: sovremennyye podhody i sushchestvuyushchie problemy // Social'nye i gumanitarnyye nauki. Otechestvennaya i zarubezhnaya literatura. Ser. 6, YAzykoznanie: Referativnyj zhurnal. – 2020. – № 4. S. 23–27.

СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTHORS

КУРЧЕЕВА Галина Ивановна

E-mail: kurcheeva@yandex.ru

Galina I. KURCHEEVA

E-mail: kurcheeva@yandex.ru

КОПЫЛОВ Виктор Борисович

E-mail: viktorkopylov1997@mail.ru

Victor B. KOPYLOV

E-mail: viktorkopylov1997@mail.ru

Поступила: 07.05.2022; Одобрена: 27.06.2022; Принята: 28.06.2022.

Submitted: 07.05.2022; Approved: 27.06.2022; Accepted: 28.06.2022.