

*Научная статья*

УДК 338.434

DOI 10.18101/2304-4446-2026-1-155-163

**Анализ цифровых следов населения как инструмент диагностики  
и прогнозирования социально-экономических проблем  
сельских территорий региона (на примере Республики Бурятия)**

© **Тимофеев Александр Николаевич**

генеральный директор,

ООО «СибДиджитал»

Россия, 670046, г. Улан-Удэ, ул. Красной Звезды, 23, оф. 18

tan@sibdigital.net

© **Ванчикова Елена Николаевна**

доктор экономических наук, профессор,

Бурятская государственная сельскохозяйственная академия имени В. Р. Филиппова

Россия, 670010, г. Улан-Удэ, ул. Пушкина, 8

evanch@mail.ru

© **Садуев Нима Батодоржиевич**

кандидат физико-математических наук, доцент,

Бурятская государственная сельскохозяйственная академия имени В. Р. Филиппова

Россия, 670010, г. Улан-Удэ, ул. Пушкина, 8

saduev@yandex.ru

© **Ванзатова Елена Очировна**

кандидат экономических наук, доцент,

Бурятская государственная сельскохозяйственная академия имени В. Р. Филиппова

Россия, 670010, г. Улан-Удэ, ул. Пушкина, 8

vanzatova.elena@yandex.ru

© **Нечаев Дмитрий Витальевич**

аспирант,

Бурятская государственная сельскохозяйственная академия имени В. Р. Филиппова

Россия, 670010, г. Улан-Удэ, ул. Пушкина, 8

castuargreen@mail.ru

**Аннотация.** В условиях цифровой трансформации публичной сферы анализ неструктурированных текстовых обращений граждан становится критически важным инструментом для повышения эффективности регионального управления. В статье представлены результаты разработки и апробации комплексной аналитической системы, построенной на интеграции методов машинного обучения и больших языковых моделей (LLM) для исследования общественных настроений жителей Республики Бурятия. На основе массивов данных из трех ключевых каналов коммуникации за 2023–2025 гг. система обеспечивает автоматическое выявление латентных проблемных кластеров, их семантическую интерпретацию, генерацию аналитических резюме и прогнозирование динамики обращений. Практические результаты демонстрируют способность системы диагностировать пространственно-отраслевые диспропорции, проактивно обнаруживать новые социально-экономические вызовы и формировать доказательную базу для обоснования управленческих и инвестиционных решений. Вне-

дрение подобных решений позволяет осуществить переход от реактивной к проактивной модели управления, оптимизировать бюджетные расходы и повысить качество государственных услуг, что вносит вклад в развитие концепции управления, основанного на данных (data-driven governance) на региональном уровне.

**Ключевые слова:** региональная экономика, социально-экономическое развитие, цифровые следы, машинное обучение, общественное настроение, Республика Бурятия, прогнозирование, диагностика проблем.

### **Благодарности**

Работа выполнена при финансовой поддержке республиканского бюджета Республики Бурятия, проект №25-28-20352 «Применение методов машинного обучения для оценки эффективности реализации проектов по программе “Комплексное развитие сельских территорий на основе мониторинга изменений актуальных проблем сельских жителей”» и НИР «Построение математической модели для исследования общественных настроений их текстовых сообщений, опубликованных в социальных сетях на основе цифровых следов жителей Республики Бурятия».

### **Для цитирования**

Анализ цифровых следов населения как инструмент диагностики и прогнозирования социально-экономических проблем сельских территорий региона (на примере Республики Бурятия) / А. Н. Тимофеев, Е. Н. Ванчикова, Н. Б. Садуев [и др.] // Вестник Бурятского государственного университета. Экономика и менеджмент. 2026. № 1. С. 155–163.

### **Введение**

Современные вызовы регионального развития, усугубляемые высокой динамикой социально-экономических процессов и растущими ожиданиями граждан, требуют от органов власти принципиально новых, более чувствительных инструментов диагностики и мониторинга. Традиционные инструменты — статистические отчеты, плановые проверки и выборочные социологические опросы — обладают существенным системным лагом и часто фиксируют уже устоявшиеся, а не зарождающиеся тренды. Это формирует «слепые зоны» в управленческом восприятии, снижая способность системы к своевременному и адекватному реагированию.

В этом контексте неструктурированные цифровые следы — спонтанные тексты обращений, жалоб и предложений граждан в официальные системы (такие как «Инцидент-менеджмент» или портал обращений) и публичные социальные сети — превращаются в уникальный источник оперативной, нефильтованной и массовой информации «с мест». Их анализ позволяет в режиме, близком к реальному времени, улавливать актуальные болевые точки, измерять общественную температуру и выявлять скрытые причинно-следственные связи проблем.

Дискуссия о роли цифровых данных в управлении сформировалась в рамках концепций «электронного правительства» (e-government) и «Government 2.0», предполагающих переход к партнерской модели взаимодействия власти и общества, основанной на диалоге и совместном создании решений [1–3]. Цифровые платформы обратной связи стали ключевым инструментом этой трансформации, обеспечивая канал для прямого вовлечения граждан. Исследования показывают, что в российских регионах наиболее распространены платформы, функциони-

рующие по принципу инцидент-менеджмента, которые позволяют фиксировать и адресно решать конкретные проблемы [1]. Однако их потенциал часто ограничивается реактивным реагированием, в то время как более сложные, интегрированные платформы, способные к анализу данных и проактивному управлению, внедрены пока лишь точечно [1]. Таким образом, существует разрыв между сбором цифровых обращений граждан и их системным использованием для стратегического планирования.

Параллельно в научной литературе активно развивается направление, связанное с применением больших данных (Big Data) и технологий искусственного интеллекта (ИИ) для повышения эффективности управления. Анализ больших объемов данных рассматривается как основа для перехода к более «умным», эффективным и прозрачным городам и регионам [4]. Исследования демонстрируют устойчивый рост интереса к использованию ИИ и машинного обучения не только в корпоративном, но и в публичном управлении, связывая с этими технологиями ожидания роста результативности и прозрачности [5–7]. Особое внимание уделяется анализу, ориентированному на гражданина (citizen-centered), который позволяет учитывать индивидуальные модели поведения и мнения для персонализации услуг и повышения качества решений [4].

В контексте данного исследования центральными являются две взаимосвязанные концепции:

1. «Цифровые следы» граждан — это большие массивы неструктурированных текстовых данных (обращения, жалобы, обсуждения в соцсетях), которые образуются в результате повседневного взаимодействия населения с цифровыми платформами. Их правовой режим и ценность как объекта анализа продолжают уточняться в научной литературе [8; 9].

2. Семантический анализ текста — это ключевая технология обработки естественного языка (NLP), которая позволяет перейти от простого учета ключевых слов к пониманию смысла, контекста и выявлению скрытых тематических паттернов в текстовых массивах [10]. Именно этот метод открывает возможность трансформации сырых данных в структурированное знание.

Несмотря на обилие теоретических и концептуальных работ, отмечается недостаток операциональных исследований, предлагающих конкретные и проверяемые пути (frameworks) для перехода от сбора данных к обоснованным управленческим действиям [4]. Большинство существующих работ носят общий характер и не предлагают детального методологического инструментария, адаптированного к специфике регионального управления и работе с разнородными источниками текстовой информации [4].

Особую актуальность такой анализ приобретает для регионов со сложной пространственно-территориальной организацией. Республика Бурятия, характеризующаяся значительной площадью, низкой плотностью населения, отраслевой спецификой и суровыми климатическими условиями, служит релевантной площадкой для апробации интеллектуальных систем анализа данных. Исходя из обозначенного научного пробела, **целью настоящего исследования** является разработка и апробация не просто математической модели, а целостной **аналитической платформы**, которая преодолевает разрыв между сбором цифровых следов и их использованием в экономической политике региона. Новизна работы заключается в **комплексном сочетании** нескольких элементов:

- **интеграция разнородных источников** данных (официальные порталы, системы инцидент-менеджмента, региональные СМИ) в единое проблемное поле;
- **применение каскадной методологии**, последовательно объединяющей семантическую векторизацию (например, на основе BERT), машинное обучение для кластеризации и генеративные языковые модели (LLM) для интерпретации результатов;
- **фокус на проактивное управление**, что выражается в модулях для выявления латентных проблем и прогнозирования динамики обращаемости.

Таким образом, настоящее исследование вносит вклад как в теорию data-driven управления (предлагая конкретный операциональный фреймворк), так и в практику регионального развития, демонстрируя работоспособность подхода на примере Республики Бурятия.

#### **Материалы и методы исследования**

В основе разработанной методологии лежит идея создания сквозного аналитического конвейера, преобразующего поток разрозненных текстовых сообщений в иерархическую систему управленчески значимых индикаторов. Архитектура платформы построена по модульному принципу и включает несколько взаимосвязанных этапов обработки информации.

1. **Интеграция и нормализация разнородных источников данных.** Первоначальным этапом является формирование единого корпуса текстов из трех ключевых каналов коммуникации: системы «Инцидент-менеджмент», регионального информационного портала «БурятИнфо» и федерального портала обращений субъектов РФ. Для обеспечения сопоставимости данных проведена сложная работа по унификации категориальных схем и пространственных привязок, что позволило создать целостную картину проблемного поля региона.

2. **Семантическое моделирование и выделение смысловых паттернов.** Для перевода текстовой информации в формат, пригодный для машинного анализа, применялись современные методы обработки естественного языка. После процедур очистки и лемматизации тексты преобразовывались в векторные представления с использованием глубоких нейросетевых архитектур семейства BERT. Это позволило перейти от анализа ключевых слов к работе со смысловыми конструкциями и контекстными связями, что принципиально повысило качество последующей кластеризации.

3. **Алгоритмическое выявление тематических кластеров.** На этапе кластерного анализа векторные представления обращений подвергались автоматической группировке с применением алгоритма k-means. Ключевой особенностью данного этапа стала способность системы обнаруживать устойчивые тематические образования, не укладывающиеся в рамки заранее заданного административного классификатора. Эти «латентные кластеры» представляют особую ценность, так как часто указывают на новые, формирующиеся проблемы, еще не попавшие в фокус внимания органов власти.

4. **Содержательная интерпретация результатов с применением генеративных моделей.** Для перевода результатов машинного анализа на язык управленческой аналитики были задействованы большие языковые модели (Large Language Models). На этом этапе решались две взаимосвязанные задачи: автоматическая генерация содержательных названий для выявленных кластеров и создание развернутых аналитических резюме по каждому проблемному блоку. Ис-

пользование LLM позволило получать структурированные отчеты, включающие не только констатацию проблемы, но и анализ причин, географической привязки и возможных направлений решения.

**5. Прогнозно-аналитический модуль для стратегического планирования.** Отдельный блок платформы посвящен анализу временных рядов по обращениям граждан. С применением методов статистического моделирования (включая ARIMA и регрессионный анализ) система способна строить среднесрочные прогнозы динамики обращаемости по различным тематическим направлениям. Эта функциональность позволяет перейти от реактивного к упреждающему управлению, своевременно планируя ресурсы для решения ожидаемых проблем.

#### **Результаты и их обсуждение**

Внедрение разработанной платформы на данных Республики Бурятия за 2023–2025 гг. позволило получить ряд принципиально важных результатов, имеющих непосредственное значение для формирования региональной экономической политики.

1. Многоуровневая диагностика приоритетов населения и выявление структурных дисбалансов. Системный анализ массива обращений выявил четкую иерархию проблем, волнующих жителей региона. Так, вопросы жилищно-коммунального хозяйства и городского благоустройства составляют порядка 32% всего объема обращений, а проблемы транспортной инфраструктуры — около 28%. При этом традиционно значимые в публичной дискуссии сферы образования и здравоохранения формируют значительно меньший поток жалоб (8 и 12% соответственно). Эти данные указывают на существующий в общественном восприятии приоритет инфраструктурных проблем над социальными, что требует соответствующей корректировки акцентов в региональных программах развития.

Пространственный анализ дополнил эту картину важными территориальными различиями. Для столичного округа Улан-Удэ характерна концентрация проблем, связанных с качеством услуг в многоквартирном жилом фонде и организацией городской среды. В то же время для периферийных и сельских районов (таких как Баунтовский или Муйский) на первый план выходят вопросы транспортной доступности, качества связи и физической удаленности от центров предоставления государственных и муниципальных услуг. Выявленная дихотомия городской и сельской проблематики свидетельствует о необходимости разработки дифференцированных подходов к управлению развитием разных типов территорий.

2. Проактивное выявление формирующихся проблемных зон. Одним из наиболее значимых результатов работы платформы стала ее способность идентифицировать новые, ранее не фиксировавшиеся тематические кластеры. Среди автоматически обнаруженных и содержательно описанных проблемных областей можно выделить:

- комплекс вопросов, связанных с логистикой, безопасностью и социально-экономической адаптацией приграничных территорий в условиях меняющегося геополитического контекста.
- проблематику административных барьеров и сложности процедурного соответствия (*procedural compliance*) при получении мер государственной поддержки субъектами малого и среднего предпринимательства.

– вопросы сохранения и содержания объектов историко-культурного наследия в малых населенных пунктах, где эти проблемы часто остаются вне фокуса внимания из-за ограниченности ресурсов.

– возникающие экологические риски, связанные с изменением режима использования природных ресурсов и антропогенной нагрузкой на локальные экосистемы.

Обнаружение таких латентных трендов на ранней стадии их формирования предоставляет региональным властям возможность не просто реагировать на уже оформившиеся кризисы, а упреждающе формировать политики и программы, направленные на смягчение потенциальных негативных последствий.

3. Создание целевой доказательной базы для обоснования бюджетных и инвестиционных решений. Сгенерированные языковыми моделями аналитические резюме по кластерам представляют собой готовые элементы управленческой документации. Например, по кластеру дорожных проблем система может сформировать отчет следующего содержания: «Систематическое разрушение дорожного полотна на участке федеральной трассы Р-258 "Байкал" (километры 120–145) в весенне-осенний период. Основные упоминаемые причины: превышение допустимой нагрузки грузовым транспортом, неудовлетворительное состояние системы водоотведения. Ключевые последствия: рост аварийности, увеличение транспортных издержек для регионального бизнеса, ухудшение транспортной связности населенных пунктов». Такая детализированная и привязанная к конкретным объектам информация служит прямым обоснованием для включения соответствующих проектов в государственные и муниципальные программы, позволяет точно оценить необходимые объемы финансирования и спрогнозировать социально-экономический эффект от их реализации.

4. Оптимизация операционного управления и бюджетирования на основе прогнозной аналитики. Прогнозный модуль платформы продемонстрировал высокую точность в предсказании сезонных и событийных всплесков обращаемости по различным тематикам. Возможность заблаговременного прогнозирования увеличения жалоб, связанных с подготовкой к отопительному сезону, последствиями паводков или сезонной транспортной нагрузкой, позволяет соответствующим службам перейти от работы в авральном режиме к плановой деятельности. На практике это означает возможность:

- заблаговременного формирования материально-технических резервов;
- оптимизации графиков работы и загрузки персонала;
- точно рассчитанного и обоснованного закрепления необходимых финансовых ресурсов в проектах бюджетов различных уровней;
- снижения непродуктивных расходов, связанных с ликвидацией последствий, а не с их предотвращением.

Разработанная методика и ее практическая реализация имеют существенное значение для повышения эффективности регионального управления и экономики публичного сектора. Можно выделить несколько ключевых направлений потенциального положительного воздействия:

1. Повышение адресности и результативности бюджетных расходов. Использование объективных данных об актуальных проблемах населения позволяет перейти от инерционного, основанного на исторических паттернах распределения бюджетных средств, к целевому, проблемно-ориентированному финансированию

нию. Это создает прямые предпосылки для роста отдачи от каждого бюджетного рубля и повышения общей эффективности публичных финансов.

2. Формирование системы мониторинга результативности государственных программ и политик. Динамика обращений граждан по тематическим кластерам может служить независимым и оперативным индикатором эффективности реализуемых мер. Снижение количества и остроты жалоб по определенной проблеме после реализации соответствующей программы является объективным свидетельством ее успешности, в то время как отсутствие изменений или рост негативной обратной связи сигнализирует о необходимости корректировки подхода.

3. Снижение информационной асимметрии и рост общественного доверия. Систематический анализ и учет мнения граждан, выраженного в цифровых обращениях, сокращают разрыв между ожиданиями населения и действиями органов власти. Прозрачность и доказательность принимаемых на основе данных решений способствуют укреплению социального капитала и повышению легитимности региональных и местных органов управления, что является критически важным нематериальным активом для устойчивого развития территории.

4. Стимулирование развития региональной инновационной экосистемы. Разработка и внедрение сложных аналитических систем создают устойчивый внутренний спрос на высокотехнологичные продукты и услуги в сфере Data Science, искусственного интеллекта и разработки программного обеспечения. Это способствует закреплению в регионе квалифицированных кадров, созданию высокопроизводительных рабочих мест и формированию компетенций, востребованных в цифровой экономике.

### **Заключение**

Проведенное исследование подтверждает высокий потенциал анализа цифровых следов населения как инструмента повышения эффективности регионального управления и экономической политики. Разработанная и апробированная на материалах Республики Бурятия методологическая платформа демонстрирует возможность трансформации массива неструктурированных текстовых обращений в систему управленчески значимых индикаторов, прогнозов и аналитических продуктов.

Реализация подобных решений на практике открывает путь к переходу от реактивной, основанной на интуиции и инерции модели управления, к проактивной data-driven парадигме, где каждое решение подкреплено доказательствами и направлено на решение реально существующих, а не гипотетических проблем. Для регионов Российской Федерации, сталкивающихся со сложными вызовами пространственного развития и ограниченности ресурсов, внедрение таких систем может стать одним из ключевых факторов повышения качества жизни населения, оптимизации бюджетных расходов и достижения целей устойчивого развития.

### **Литература**

1. Ковальчук А. Г., Цыбаков Д. Л. Анализ современных цифровых платформ вовлечения граждан в региональное управление: опыт российских субъектов // Общество: политика, экономика, право. 2025. № 9(146). С. 33–41. DOI 10.24158/per.2025.9.3. EDN MCCWAT. Текст: непосредственный.
2. Васильева Е. И., Орфониций А. В. Платформы гражданского участия как направление цифровизации публичного управления // Муниципалитет: экономика и управление.

2022. № 3. С. 49–60. DOI: 10.22394/2304-3385-2022-3-49-60 EDN: VCBSLS. Текст: непосредственный.

3. Порошина Л. А., Зелева А. Д. Реализация цифровой трансформации через платформу «Единая централизованная цифровая платформа в социальной сфере» // Финансовая экономика. 2025. № 5. С. 222–226. EDN HPSATB. Текст: непосредственный.

4. Ju J., Liu L., Feng Y. Citizen-centered big data analysis-driven governance intelligence framework for smart cities. *Telecommunications Policy*. 2018; 42(10): 881–896. DOI: 10.1016/j.telpol.2018.01.003.

5. Мейрияни М., Варганегара Д. Л., Андини В. Большие данные, машинное обучение, искусственный интеллект и блокчейн в корпоративном управлении // Форсайт. 2023. Т. 17, № 4. С. 69–78. DOI 10.17323/2500-2597.2023.4.69.78. EDN SZPLVY. Текст: непосредственный.

6. Завьялова Е. А., Погадаева Н. Ю. Цифровая трансформация государственных и муниципальных услуг // Вестник Кемеровского государственного университета. Сер. Политические, социологические и экономические науки. 2021. Т. 6, № 2(20). С. 219–228. DOI: 10.21603/2500-3372-2021-6-2-219-228 EDN: TNHVLQ. Текст: непосредственный.

7. Мухаметов Д. Р. Модели платформ вовлечения граждан для создания в России умных городов нового поколения // Вопросы инновационной экономики. 2020. № 3. С. 1605–1622. EDN: GUXXWD. Текст: непосредственный.

8. Аюшеева И. З. Большие данные: проблемы определения гражданско-правового режима // *Lex Russica (Русский закон)*. 2023. Т. 76, № 10(203). С. 125–134. DOI 10.17803/1729-5920.2023.203.10.125-134. EDN KRLORF. Текст: непосредственный.

9. Василевская Л. Ю., Подузова Е. Б., Тасалов Ф. А. Цифровизация гражданского оборота: big data в механизме гражданско-правового регулирования (цивилистическое исследование): монография: в 5 томах. / ответственный редактор Л. Ю. Василевская. Москва: Проспект, 2023. Т. 5. 360 с. EDN: FKCCGN. Текст: непосредственный.

10. Захарова О. И. Разработка модели семантического анализа и синтеза текстовых данных // Приборы и системы. Управление, контроль, диагностика. 2024. № 4. С. 49–57. DOI 10.25791/pribor.4.2024.1493. EDN NGMXXF. Текст: непосредственный.

Статья поступила в редакцию 16.12.2025; одобрена после рецензирования 10.01.2026; принята к публикации 16.01.2026.

### Analysis of Population Digital Footprints as a Tool for Diagnosing and Forecasting Socioeconomic Problems in Rural Areas (case study of the Republic of Buryatia)

*Aleksandr N. Timofeev*

ООО “SibDigital”

18 office, 23 Krasnoy Zvezdy St., Ulan-Ude 670046, Russia

tan@sibdigital.net

*Elena N. Vanchikova*

Dr. Sci. (Econ.), Prof.,

Filippov Buryat State Agricultural Academy

8 Pushkina St., Ulan-Ude 670010, Russia

evanch@mail.ru

*Nima B. Saduev*

Cand. Sci. (Phys. and Math.), A/Prof.

Filippov Buryat State Agricultural Academy

8 Pushkina St., Ulan-Ude 670010, Russia  
saduev@yandex.ru

*Elena O. Vanzatova*  
Cand. Sci. (Econ.), A/Prof.  
Filippov Buryat State Agricultural Academy  
8 Pushkina St., Ulan-Ude 670010, Russia  
vanzatova.elena@yandex.ru

*Dmitry V. Nechaev*  
Research Assistant,  
Filippov Buryat State Agricultural Academy  
8 Pushkina St., Ulan-Ude 670010, Russia  
cactuargreen@mail.ru

*Abstract.* In the context of the digital transformation of the public sphere, an analysis of unstructured textual citizen communications is becoming a critical tool for improving the efficiency of regional governance. The article presents the results of developing and testing a comprehensive analytical system built on the integration of machine learning methods and large-scale language models (LLM) for studying public sentiment in the Republic of Buryatia. Using datasets from three key communication channels for 2023–2025, the system automatically identifies latent problem clusters, interprets them semantically, generates executive summaries, and forecasts communication trends. Practical results demonstrate that the system is capable of diagnosing spatial and sectoral imbalances, proactively identifying new socioeconomic challenges, and generating an evidence base to support management and investment decisions. The implementation of such solutions enables the transition from a reactive to a proactive management model, optimizes budget expenditures, and improves the quality of public services, thereby contributing to the development of data-driven governance at the regional level.

*Keywords:* regional economy, socioeconomic development, digital footprints, machine learning, public sentiment, the Republic of Buryatia, forecasting, problem diagnosis.

#### *Acknowledgments*

The study was supported by the Republic of Buryatia's budget, Project No. 25-28-20352, “Application of machine learning methods to assess the efficiency of projects under the program 'Integrated Development of Rural Areas Based on Monitoring Changes in Urgent Problems of Rural Residents' ” and the research project “Construction of a Mathematical Model for Studying Public Sentiment in Text Messages Published on Social Networks Based on the Digital Footprints of Residents of the Republic of Buryatia”.

#### *For citation*

Timofeev A. N., Vanchikova E. N., Saduev N. B. et al. Analysis of population digital footprints as a tool for diagnosing and forecasting socioeconomic problems in rural areas (case study of the Republic of Buryatia). *Bulletin of Buryat State University. Economy and Management.* 2026; 1: 155–163 (In Russ.).

*The article was submitted 16.12.2025; approved after reviewing 10.01.2026; accepted for publication 16.01.2026.*