

УДК 004.91

А.М. Кузнецов

Информационная технология автоматизации социологических исследований: методы, технологии и преимущества

Проанализирована разработка информационно-аналитической системы (ИАС) для социологических исследований. Рассмотрены существующие программные решения для сбора, обработки и анализа данных, их преимущества и ограничения. Предложена интеграция различных функций в одной системе для обеспечения комплексного подхода к исследованиям, включая проактивное прогнозирование и анализ. Разработанная ИАС предоставляет широкие возможности для автоматизации ключевых этапов социологических исследований, улучшения качества данных и повышения точности аналитики. Интеграция онтологий в структуру ИАС позволила стандартизировать термины и понятия, установить семантические связи между ними, что уменьшило вероятность ошибок и сократило время на подготовку информации к ретроспективному анализу (сокращение времени на 20%), повысило качество аналитических выводов за счет более глубокой и комплексной обработки данных.

Ключевые слова: информационная технология, архитектура системы, социологические исследования, онтология, прогнозирование, система обработки данных (СОД), система поддержки принятия решений (СППР).

DOI: 10.21293/1818-0442-2024-27-3-109-117

Глобальный рынок программного обеспечения для сбора данных (Data Collection Software Market) является одним из наиболее динамично развивающихся рынков в сфере информационных технологий. Согласно отчету компании GlobalData PLC (Великобритания), ожидается, что глобальный рынок программного обеспечения для сбора данных будет расти значительными темпами в течение прогнозируемого периода, с 2023 по 2031 г. [1]. По прогнозу компании Fact.MR (Дубай), объем мирового рынка сбора данных и их маркировки составит 2,57 млрд долл. США в 2024 г., а среднегодовой темп роста составит 18%. Таким образом, к концу 2034 г. объем этого рынка может увеличиться до 13,45 млрд долл. США [2]. Предполагается, что инструменты для сбора, обработки и анализа данных будут играть решающую роль в планировании и осуществлении цифровой

трансформации бизнес-процессов в ближайшее десятилетие. Поэтому разработка проактивной информационно-аналитической системы (ИАС) для социологических исследований является актуальной задачей.

Анализ известных ИАС для социологических исследований и их сравнительные характеристики

В эпоху больших данных информационно-аналитические системы (ИАС) становятся ключевым инструментом для социологических исследований. Правильный выбор ИАС определяет глубину анализа и то, как результаты будут представлены.

Разнообразии программ и инструментов требует четкой классификации, включающей цели исследования, тип данных, функциональность и уровень анализа (рис. 1).



Рис. 1. Классы ИАС социологических исследований по типам задач

На основе этой классификации разработана сравнительная таблица различных инструментов для сбора, анализа и визуализации результатов социологических исследований (табл. 1).

Анализ специализированных программных решений показал их ограниченную эффективность для

комплексного сопровождения социологических исследований. Методологические проблемы, такие как выбор подхода, качество и объем выборки, а также точность формулировок, часто приводят к значительным расхождениям в результатах исследований с аналогичными условиями. Это становится особенно важным при

сопровождении избирательных процессов, где социологические данные могут существенно влиять на стратегии избирательных кампаний и выбор избирателей. Ограниченная функциональная специализация про-

граммного обеспечения, сосредоточенная либо на аналитических задачах, либо на предварительной обработке данных, представляет собой значительную проблему, требующую интегрированного решения.

Таблица 1

Анализ программного обеспечения для социологических исследований

Категория	ПО (примеры)	Основные недостатки	Источники
Статистические пакеты	IBM SPSS Statistics, Stata	Высокая стоимость, ограниченная гибкость, недостаточная поддержка современных методов анализа	IBM SPSS, [3–5]. Stata, [6, 7]
Среды программирования для анализа данных	Python, R	Требуют глубоких знаний программирования, сложны для новичков, недостаток интегрированных инструментов для специфических задач социологических исследований	Python, [8–10]. R Project, [6, 11]
ПО для качественного анализа	NVivo, MaxQDA, TextAnalyst, PolyAnalyst	Высокая стоимость, сложность в освоении, ограниченные возможности интеграции с количественными данными, отсутствие комплексности	NVivo, [12, 13]. MaxQDA, [12, 14, 15]. TextAnalyst, [12, 16]. PolyAnalyst [12, 17]
Платформы Big Data	Apache Hadoop, Apache Spark, RapidMiner, Elasticsearch	Требуют значительных технических компетенций для настройки и использования. Ограниченно применимы для социологических исследований. Возникают сложности в управлении и практическом применении исследователями-социологами	Hadoop, [18, 19]. Spark, [16, 19, 20]. RapidMiner, [21, 22]. Elasticsearch, [19, 23]
Системы Business Intelligence	Microsoft Power BI, SAS Business Intelligence, Tableau, Форсайт, Polymatica	Высокая стоимость, ограниченные возможности настройки под специфические требования социологических исследований, недостаточная поддержка неструктурированных данных и интеграции с другими системами	Power BI, [24, 25]. SAS, [26, 27]. Tableau, [6]. Форсайт, [25, 28]. Polymatica, [25]
ПО для сбора данных	Google Forms, SurveyMonkey, Qualtrics, REDCap, Анкетолог	Ограничения возможности настройки аналитических функций, недостаток интеграции с инструментами для анализа больших объемов данных, недостаточная поддержка пользовательских интерфейсов	Google Forms, [29, 30]. SurveyMonkey, [31]. Qualtrics, [32]. REDCap, [33]. Анкетолог, [21]
Проактивные и предиктивные системы	TensorFlow, IBM Watson, Scikit-learn, Apache Mahout, RecommenderLab	Трудности в изучении, высокие требования к ресурсам, недостаточная поддержка узкоспециализированных задач, высокая стоимость, необходимость в постоянном интернет-соединении для работы с облачными системами	TensorFlow, IBM Watson, Scikit-learn, [6, 12, 27]

Архитектура ИАС

Этапы анализа данных в рамках использования информационно-аналитической системы (ИАС) охватывают следующие процедуры: определение исследовательских целей и задач, разработка структурированных анкет и отбор соответствующих респондентов, организация подготовки анкетных материалов, реализация процесса анкетирования с применением разнообразных методик, первоначальная обработка и анализ собранных данных. На заключительном этапе происходит выработка прогнозов относительно тенденций развития событий или изменений в общественных настроениях.

На первом этапе происходит разработка анкет, подготовка интервью и других инструментов сбора информации. Второй этап – анкетирование респондентов – проводится разнообразными методами (по месту проживания, на предприятиях, на улице, через почту, мобильные устройства и др.). Третий этап включает подготовительную работу с собранной информацией и её первичную обработку, что ранее осуществлялось вручную, а в рамках ИАС автоматизировано для обеспечения точного отображения процентного распределения изучаемых групп. На четвертом этапе проводится анализ данных, который может варьироваться от создания базовых графиков и диа-

грамм до комплексного изучения распределений ответов респондентов. В ИАС реализован компонент для оптимизации выбора аналитических методов разведочного анализа данных. На пятом этапе социолог-исследователь осуществляет теоретическую интерпретацию полученных данных, формулирует заключение и прогнозирует тенденции на основе проведённого анализа [34–36].

Разработанная архитектура ИАС предоставляет комплексное решение для проведения опросов с использованием различных методов сбора и обработки данных, обеспечивая эффективность и надежность процесса последующего исследования и прогнозирования [37, 38]. ИАС состоит из нескольких подсистем, каждая из которых решает свой подкласс задач и является логически изолированной и самостоятельной компонентой. Архитектура ИАС показана на рис. 2.

Основная цель создания информационно-аналитической системы (ИАС) социологических исследований заключается в обеспечении комплексного процесса сбора, обработки, анализа и толкования социологической информации, что, в свою очередь, помогает в выработке и принятии обоснованных управленческих решений.

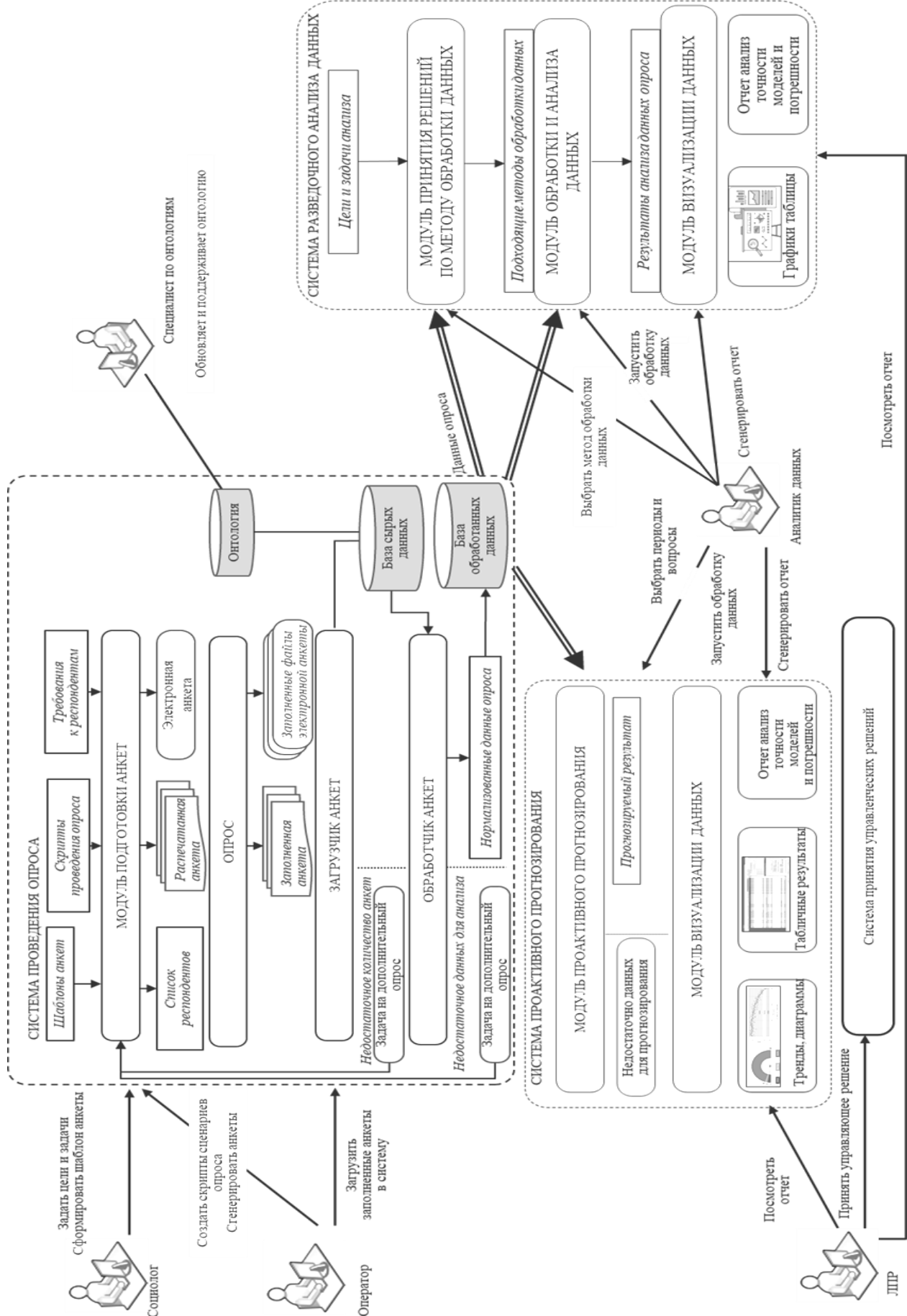


Рис. 2. Архитектура ИАС социологических исследований

ИАС позволяет автоматизировать и оптимизировать большинство этапов социологических исследований, увеличить эффективность и скорость обработки данных, а также улучшить точность предсказаний и рекомендаций для разработки стратегий развития и решения социальных вопросов.

ИАС социологических исследований состоит из следующих подсистем:

- подсистема проведения опроса (модуль подготовки анкет, загрузчик анкет, обработчик анкет);
- подсистема хранения (онтология, реляционная БД и база сырых данных);
- подсистема разведочного анализа данных (модуль принятия решений по методу обработки данных, модуль обработки и анализа данных, модуль визуализации данных);
- подсистема проактивного прогнозирования (модуль проактивного прогнозирования и модель визуализации данных);
- подсистема принятия управленческих решений.

В процессе функционирования системы основное внимание уделяется не только результативности процессов сбора и анализа данных, но и обеспечению высокого качества, надежного хранения и защиты приватных данных.

На стадии обработки и анализа данных ключевую роль играет подсистема выбора метода обработки, позволяющая аналитику данных выбрать наиболее подходящие под заданные исследовательские вопросы методы статистического анализа. Адаптивность системы к разнообразию задач позволяет использовать широкий спектр статистических и аналитических моделей, что делает возможным проведение как разведочного, так и предиктивного анализа. Разведочный и предиктивный анализы данных позволяют выявить основные тенденции и закономерности, а также сформировать прогнозы на будущее, на основании которых принимаются управленческие решения.

Важную роль в повышении эффективности работы ИАС социологических исследований играет интеграция онтологий, обеспечивающая более точный поиск и анализ данных за счет структурирования и классификации информации. Специалист по онтологиям регулярно обновляет и расширяет онтологическую базу, адаптируя её под текущие задачи исследования. Онтологии помогают обеспечить семантическую согласованность данных, что критически важно для точности аналитических выводов. Организация данных с помощью онтологий значительно улучшает возможности поиска необходимой информации и взаимосвязи между различными данными, что существенно повышает эффективность аналитического процесса. Разработана базовая онтология социологических исследований применительно к Тамбовской области и постоянно пополняется в течение последних пяти лет.

Онтология в ИАС обеспечивает семантическую согласованность данных, что особенно важно для поиска и выборки данных в файловом хранилище результатов опросов, где сохранялись необработанные

анкеты начиная с 2014 г. В этих анкетах использовались разные термины для обозначения одного и того же понятия (например, «дистанционная работа», «удаленная работа», «работа из дома»), а также нестандартизированные значения (например, возраст мог быть записан как числовое значение «25–30 лет» или категориальные данные «молодые специалисты»). Онтология позволила стандартизировать термины и понятия и установить семантические связи между ними, что уменьшило вероятность ошибок и сократило время на обработку. Благодаря этому исследователи могут объединять и анализировать данные из различных источников и периодов, даже если первоначально они были структурированы по-разному. В результате время, затрачиваемое на подготовку информации к ретроспективному анализу, сократилось на 20%, что ускоряет общий процесс исследования.

В ИАС применяется гибридная структура хранения данных социологических исследований, сочетающая реляционную базу данных и файловое хранилище, которая создает условия для эффективного сбора, хранения и анализа различных типов данных, что позволяет обеспечивать необходимую гибкость при их обработке и анализе.

Применение автоматизированных инструментов обработки данных позволяет значительно снизить процент ошибок (до 0,5%) за счёт исключения человеческого фактора на этапе первичной обработки анкет. Затраты на обслуживание снизились на 8,5%, подтверждая достижение целей по комплексному обеспечению автоматизации всех этапов процесса проведения и обработки социологических опросов.

Использование специализированного модуля для автоматического выбора методов обработки данных повышает качество анализа.

Информационная технология сопровождения социологического исследования

Как показал анализ программного обеспечения для социологических исследований (см. табл.1), на протяжении многих лет в социологических исследованиях для обработки данных широко использовались зарубежные программные продукты, такие как SPSS, Excel, Stata, NVivo, MaxQDA, Tableau и IBM Watson. Несмотря на их широкие аналитические возможности, они имеют ряд ограничений, таких как высокая стоимость, сложность настройки под специфические задачи и зависимость от зарубежных серверов и обновлений. В условиях экономических и политических санкций возникает необходимость разработки отечественных решений, способных не только предложить аналогичный функционал, но и обеспечивать безопасность данных и независимость от внешних поставщиков [39, 40].

Предложенная в данной работе информационно-аналитическая система (ИАС) позволяет решить эти проблемы за счёт интеграции всех этапов социологических исследований в единую структуру, начиная с этапа проектирования анкет и заканчивая анализом и прогнозированием. ИАС разработана с целью полного сопровождения социологических исследований

с минимальными затратами и высокой надёжностью, что делает её актуальным инструментом в современных условиях.

Ранее специализированные программные продукты применялись преимущественно на первом, третьем и частично на четвертом этапах социологических исследований. Их функциональность включала возможности распечатки анкет, ввода данных с опросных листов и простейший подсчет распределенных ответов на вопросы. Разработанная ИАС значительно расширяет сферу применения методов цифро-

визации при проведении социологических опросов. Применение этой системы позволяет социологам на подготовительном этапе анализа предмета исследования получать доступ к существующим данным, их распределениям и выявленным зависимостям, осуществляя вторичный анализ данных с использованием онтологий.

Сопоставление функциональности компонентов разработанной ИАС социологических исследований и этапов проведения социологических исследований представлено на рис. 3.

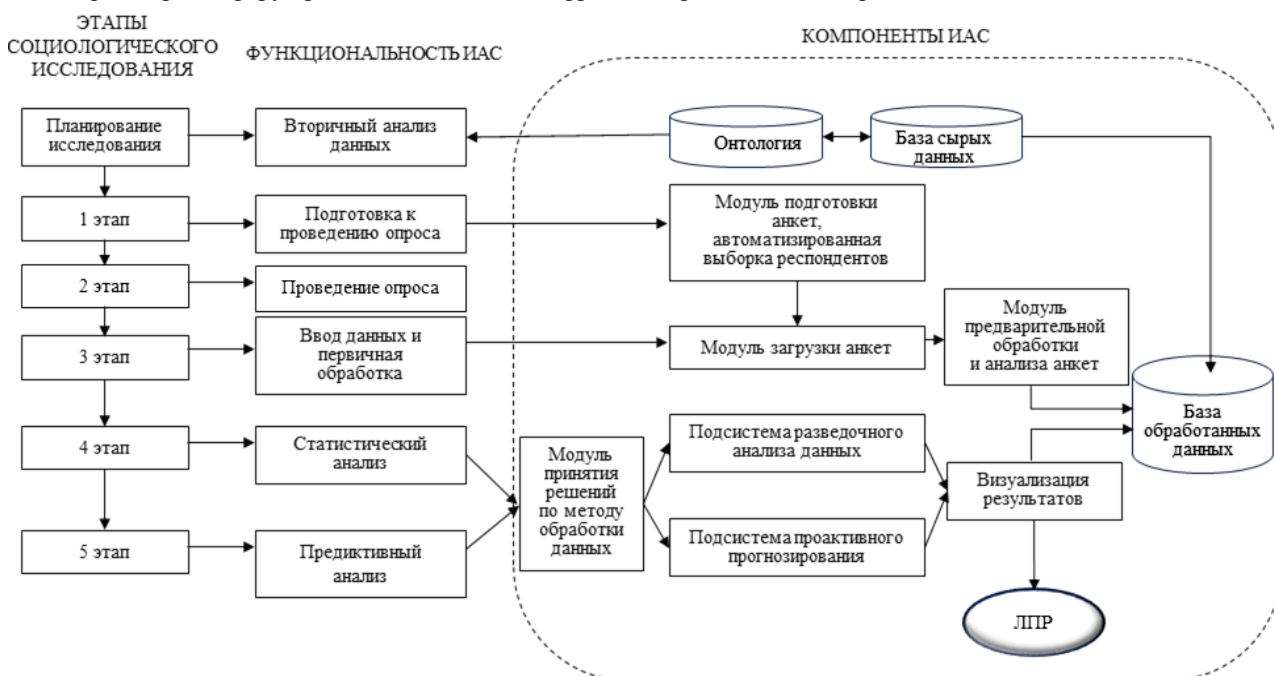


Рис. 3. Информационная технология сопровождения ИАС на этапах социологического исследования

Предлагаемая информационная технология проведения социологического исследования включает применение модулей ИАС практически на всех этапах социологического исследования, не связанных с непосредственным опросом респондентов.

Необходимо отметить, что ИАС для проведения социологических опросов должны обладать свойствами как систем поддержки принятия решений (СППР), так и систем обработки данных (СОД). В связи с этим информационная технология использования ИАС имеет двоякий характер.

Для систем обработки данных (СОД) характерна начальная стадия ввода и обработки результатов социологических опросов. Эффективность этих процессов в ИАС значительно возросла благодаря внедрению дополнительных функций, которые повышают точность и скорость обработки социологической информации:

- поиск информации на основе онтологических запросов в хранилищах необработанных данных социологических исследований;
- автоматизированный ввод первичной информации с анкетных листов в ИАС (компонент «Загрузчик анкет»);
- предварительная обработка и загрузка данных в реляционную базу данных (компонент «Обработчик анкет»);

- применение инструментов подсистемы разведочного анализа данных для контроля критических параметров выборки, выбора метода обработки данных и построения распределений.

В процессе аналитических исследований реализуются важные функциональные возможности, которые необходимы специалистам в области социологии для анализа и толкования полученных результатов и прогнозирования будущих тенденций. Эти возможности являются характеристиками систем поддержки принятия решений (СППР) и включают следующее:

- поиск необходимой информации социологом-исследователем осуществляется через онтологические запросы к базам данных, содержащим как исходные данные предыдущих опросов, так и обработанную информацию;
- онтологии формализуют описание исследуемых объектов и процессов, обеспечивая единый формат данных, что улучшает взаимодействие между системами и качество аналитики, позволяя проводить более точные и комплексные исследования;
- фреймворк для модуля прогнозирования в ИАС позволяет интегрировать различные методологические подходы и аналитические инструменты для обработки и анализа социологических данных.

Можно утверждать, что разработанная ИАС социологических исследований обеспечивает ком-

плексный подход, интегрируя функции ввода, обработки и анализа данных. Это не только способствует повышению точности и надежности исследований, но и расширяет возможности для глубокой аналитики и прогнозирования. Благодаря использованию онтологий становится возможным проведение более точных и комплексных исследований, так как они позволяют учитывать и анализировать широкий спектр данных, структурированных по единому принципу.

Результаты опытной эксплуатации ИАС

В рамках опытной эксплуатации ИАС для социологических исследований была проведена апробация на базе нескольких реальных социологических опросов. Цель – оценить эффективность системы по ряду эксплуатационных показателей (скорость обработки запросов, точность прогнозов, уровень автоматизации и экономия затрат на обслуживание). В ходе апробации были собраны данные, подтверждающие, что предложенная ИАС значительно улучшает исследовательский процесс за счёт интеграции всех этапов исследования – от сбора данных до прогностического анализа.

Для прогнозирования результатов выборов различных уровней в Тамбовской области использовались данные ежеквартальных опросов с выборкой в 1 000 человек за период с 2017 по 2023 г. Часть этих данных была использована в качестве обучающей выборки для построения модели, а оставшиеся данные были выделены как тестовая выборка для проверки точности прогнозов. Модель SARIMAX была выбрана из-за её способности учитывать сезонные колебания и тренды. Средняя абсолютная процентная ошибка (MAPE) составила 3,61%, что является приемлемым результатом для моделей временных рядов и подтверждает пригодность модели для прогнозирования социологических показателей данного типа.

Экспериментальная апробация ИАС социологических исследований показала высокую производительность (обработка 650–700 запросов в минуту со средним временем ответа 1–3 с).

Основные результаты опытной эксплуатации представлены в табл. 2.

Заключение

Разработка информационно-аналитической системы (ИАС) для социологических исследований в условиях санкционного давления предоставляет отечественным социологам инструмент, который не только полностью удовлетворяет потребности в анализе данных, но и обеспечивает независимость и безопасность информационных процессов. В отличие от существующих программных продуктов, сосредоточенных на отдельных аспектах исследования, предложенная ИАС интегрирует все этапы – от создания анкет до прогностического анализа. Это позволяет не только повысить точность и надежность данных, но и обеспечить более глубокий и всесторонний анализ. Использование онтологий в ИАС обеспечивает семантическую согласованность и структурированность данных, что важно для точности анализа и прогноза. Онтологии упрощают поиск и интерпретацию информации, а также способствуют улучшению взаимодействия между различными подсистемами.

Это позволяет не только ускорить процесс анализа, но и повысить качество выводов за счёт более точной интерпретации данных.

Таблица 2

Результаты опытной эксплуатации ИАС

Критерий	Традиционные методы	ИАС	Преимущество ИАС
Время обработки данных	10–12 ч	1–2 ч	Сокращение времени обработки на 80%
Процент ошибок при обработке анкет	2–7%	Менее 0,5%	Снижение ошибок за счет автоматизации
Точность прогнозов	75–85%	90–95%	Увеличение точности прогнозов (применение моделей SARIMAX и LSTM)
Интеграция этапов исследования	Отсутствует	Полная интеграция	Полный цикл исследований в одной системе
Использование онтологий	Нет	Да	Повышение согласованности и точности данных, сокращение времени на поиск релевантной информации на 20%
Стоимость внедрения	Высокая (различные решения на каждом этапе)	Средняя (единое решение)	Снижение затрат на 8,5% за счет комплексного решения на всех этапах
Время подготовки отчета	5–7 дней	1 день	Ускорение подготовки отчета

Опытная эксплуатация ИАС показала, что её применение позволяет сократить время обработки данных на 80%, снизить процент ошибок на менее 0,5%, увеличить точность прогнозирования до 10%, а также уменьшить затраты на обслуживание на 8,5%. Таким образом, информационная технология социологических исследований на основе ИАС обеспечивает высокую эффективность и скорость обработки социологической информации, а также позволяет усовершенствовать исследовательский процесс.

Литература

1. Data Collection and Labelling Market Size, Share, Trends and Analysis by Region, Type, Vertical, and Segment Forecast to 2030 [Электронный ресурс]. – Режим доступа: <https://www.globaldata.com/store/report/data-collection-and-labelling-market-analysis>, свободный (дата обращения: 30.07.2024).
2. Data Collection and Labeling Market [Электронный ресурс]. – Режим доступа: <https://www.factmr.com/report/4726/data-collection-and-labelling-market>, свободный (дата обращения: 31.07.2024).
3. Обзор перечня и особенностей программного обеспечения, применяемого для статистической обработки данных при выполнении экспертной работы / К.А. Кошечкин, А.В. Козлов, В.Н. Котиков, А.Н. Миронов // Ведомости научного центра экспертизы средств медицинского применения. – 2014. – № 3. – С. 46–51.

4. Андреева А.А. Возможности использования профессионального статистического пакета для социальных наук SPSS Statistics для анализа и обработки данных исследования общественного мнения // Теория и практика современной науки. – 2020. – № 1(67). – С. 30–43.
5. Орестова В.Р., Бастрон А.А. Применение статистического пакета анализа данных SPSS Statistics в психологических исследованиях на примере факторного анализа // История и архивы. – 2017. – № 2(8). – С. 38–51.
6. Демаков В.И. Сравнительный обзор статистических пакетов для анализа данных / В.И. Демаков, Е.Ю. Ларионова, Ю.Э. Голодков, В.И. Рерке // Вестник БГУ. Математика, информатика. – 2023. – № 3. – С. 78–89.
7. Унгуряну Т.Н. Однофакторный дисперсионный анализ с использованием пакета статистических программ Stata / Т.Н. Унгуряну, А.М. Гржибовский // Экология человека. – 2014. – № 5. – С. 60–64.
8. Смирнов В.А. Новые компетенции социолога в эпоху «Больших данных» // Мониторинг. – 2015. – № 2 (125). – С. 44–54. DOI: <https://doi.org/10.14515/monitoring.2015.2.04>.
9. Андрюк Д.С. Семантический анализ структуры ценностей группы с помощью тезауруса Роже: автоматизированный алгоритм / Д.С. Андрюк, А.С. Ливитина, Н.С. Сушко // Государственное управление. Электронный вестник. – 2022. – № 91. – С. 148–161.
10. Использование машинного обучения для изучения качества жизни населения: методологические аспекты / Е.В. Щекотин, В.Л. Гойко, П.А. Басина, В.В. Бакулин // Цифровая социология. – 2022. – № 1. – С. 87–97.
11. Рагимова Н.А. Применение технологий больших данных для мониторинга социальных процессов в университете / Н.А. Рагимова, В.Г. Абдуллаев, М.Э. Халилов // Science Rise. – 2019. – № 8 (61). – С. 17–23.
12. Гегер А.Э. Компьютерные программы для анализа качественных и смешанных данных / А.Э. Гегер, Ю.А. Чупахина, С.А. Гегер // Петербургская социология сегодня. – 2015. – № 6. – С. 374–383.
13. Каныгин Г.В. Аналитическое кодирование / Г.В. Каныгин, В.С. Корецкая // Петербургская социология сегодня. – 2018. – № 10. – С. 99–121.
14. Хуажев А.А. Технологические средства анализа имиджформирующих политических медиатекстов: возможности и ограничения // Изв. Саратов. ун-та. Сер.: Социология. Политология. – 2022. – № 2. – С. 232–236.
15. Чудновская И.Н. Модель коммуникации в цифровом образовательном пространстве в рефлексии студентов / И.Н. Чудновская, И.Н. Бухтиярова, М.Е. Липатова // Социология. – 2022. – № 1. – С. 173–183.
16. Харламов А.А. Инструментарий для информационно-аналитической экспертной оценки научной продукции с целью выявления приоритетных научных направлений и коллективов / А.А. Харламов, Б.И. Васин // Вестник Моск. гос. лингвистического ун-та. Гуманитарные науки. – 2018. – № 6 (797). – С. 164–174.
17. Ученые ТГУ с помощью платформы PolyAnalyst измерили настроения россиян [Электронный ресурс]. – Режим доступа: <https://www.megaputer.ru/uchenye-izmerilischaste-gossijan/>, свободный (дата обращения: 30.07.2024).
18. Василенко Л.А. Социология цифрового общества: монография / Л.А. Василенко, Н.Н. Мещерякова. – Томск: Изд-во Том. политехн. ун-та, 2021. – 226 с.
19. Утепбергенов А.О. Большие данные в управлении человеческими ресурсами // Экономика и социум. – 2022. – № 4–3(95). – С. 513–519.
20. Анализ социальных сетей: методы и приложения / А. Коршунов, И. Белобородов, Н. Бузун, В. Аванесов, Р. Пастухов, К. Чихрадзе, И. Козлов, А. Гомзин, И. Андрианов, А. Сысоев, С. Ипатов, И. Филоненко, К. Чуприна, Д. Турдаков, С. Кузнецов // Труды ИСП РАН. – 2014. – № 26 (1). – С. 439–456.
21. Груданова А.А. Обзор инструментов онлайн-опросов для информационного обеспечения маркетинговых исследований / А.А. Груданова, Н.А. Груданов // StudNet. – 2021. – № 7. – С. 1814–1829.
22. Никонова М.Л. Компьютерная модель решения задач классификации в программной среде Rapid Miner // Медицинское образование и профессиональное развитие. – 2017. – № 2–3. – С. 28–29.
23. Василькова В.В. Боты на публичных аренах социальных сетей / В.В. Василькова, Н.И. Легостаева // Социологический журнал. – 2021. – № 4. – С. 99–117.
24. Гончарова И.В. Технологии больших данных как инструмент решения задач устойчивого развития общества / И.В. Гончарова, Г.Б. Прончев // Право и управление. – 2023. – № 5. – С. 215–226.
25. Рогожа А.И. Сравнительный анализ российских систем Business Intelligence // Скиф. – 2022. – № 8 (72). – С. 282–287.
26. Хасанов А.Р. Влияние предиктивной аналитики на деятельность компаний // Стратегические решения и риск-менеджмент. – 2018. – № 3. – С. 108–113. <https://doi.org/10.17747/2078-8886-2018-3-108-113>.
27. Цыпин А.П. Статистические пакеты программ в социально-экономических исследованиях / А.П. Цыпин, А.С. Сорокин // АНИ: экономика и управление. – 2016. – № 4 (17). – С. 379–384.
28. Шимановский К.В. Импортзамещение в области программного обеспечения бизнес-аналитики // Международный журнал прикладных наук и технологий «Integral». – 2022. – № 2. – С. 527–537.
29. Борханова А.Д. Применение анкеты Google Forms для изучения педагогических условий реализации инновационных проектов в школе // КПО. – 2021. – № 3 (27). – С. 32–36.
30. The Google Forms Online Service is Used to Assess the Socio-psychological Qualities of a Person and the Level of Professional Competence of Average Medical Professionals / A. Syzdykova, L. Orakbay, A. Mansharipova, B. Shuzheev, M. Kushniyazova // Актуальные проблемы теоретической и клинической медицины. – 2021. – № 2. – С. 35–40.
31. Кед А.П. Интернет-опрос как метод социологического исследования / А.П. Кед, П.М. Агаева // Проблемы современной экономики (Новосибирск). – 2015. – № 27. – С. 112–116.
32. Пантелеев Д.Н. Исследование и характеристика маркетингового атрибута инноваций // Вестник Академии знаний. – 2022. – № 6 (53). – С. 453–456.
33. Информационная система REDCap для сбора и хранения данных популяционных биомедицинских исследований / А.В. Аталян, Л.И. Колесникова, С.И. Колесников, А.М. Гржибовский, Л.В. Сутурина // Экология человека. – 2019. – № 2. – С. 52–59.
34. Кузнецов А.М. Структура модели информационно-аналитической системы обработки социологических данных // Нанобиотехнологии. Теплоэнергетика. Математическое моделирование: сборник статей междунар. науч.-практ. конф., Липецк, 27–28 февраля 2023 г. – Липецк: Липецкий гос. техн. ун-т, 2023. – С. 258–262.
35. Саклаков В.М. Методология цифрового социологического исследования: общественная система как базовый инструмент моделирования // Доклады ТУСУР. – 2023. – Т. 26, № 4. – С. 61–77. DOI: [10.21293/1818-0442-2023-26-4-61-77](https://doi.org/10.21293/1818-0442-2023-26-4-61-77).
36. Св-во о гос. регистрации программы для ЭВМ № 2024617985 / А.М. Кузнецов. – Заявка № 202461581. Дата поступ.: 22 марта 2024 г. Зарег. в Реестре программ для ЭВМ: 08 апреля 2024 г.

37. Св-во о гос. регистрации программы для ЭВМ № 2024666569 / А.М. Кузнецов. – Заявка № 2024665224. Дата поступ.: 3 июля 2024 г. Зарег. в Реестре программ для ЭВМ: 15 июля 2024 г.

38. Кузнецов А.М. Выбор модели и алгоритма прогнозирования событий на основе результатов обработки данных социологических исследований // Нейрокомпьютеры и их применение: сборник тезисов XXI Всерос. науч. конф., Москва, 28 марта 2023 г. – М.: Моск. гос. психолого-педагогический ун-т, 2023. – С. 93–94.

39. Курбатова А.С. Перспективы импортозамещения в российском секторе информационных технологий в условиях санкционного давления // Современная мировая экономика. – 2023. – Т. 1, № 4 (4). – С. 108–128. DOI: 10.17323/2949-5776-2023-1-4-108-128.

40. Злобина О.В. Перспективы автоматизации и цифровизации производства в условиях введения санкций / О.В. Злобина, Г.Ю. Пешкова // Вестник Алтайской академии экономики и права. – 2022. – № 8-1. – С. 66–73.

Кузнецов Александр Михайлович

Аспирант каф. информационных систем

и защиты информации (ИСЗИ)

Тамбовского гос. технического ун-та (ГГТУ)

Советская ул., 106/5, г. Тамбов, Россия, 392000

ORCID: 0000-0003-2640-3417

Тел.: +7 (475-2) 63-00-54

Эл. почта: 068t@mail.ru

Kuznetsov A.M.

Information technology for automation of sociological research: methods, technologies and advantages

The article is devoted to the analysis and development of an information-analytical system (IAS) for sociological research. It reviews existing software solutions for data collection, processing, and analysis, outlining their advantages and limitations. The article proposes integrating various functions into a single system to provide a comprehensive approach to research, including proactive forecasting and analysis. The developed IAS offers extensive capabilities for automating key stages of sociological research, improving data quality, and enhancing analytical accuracy. The integration of ontologies into the IAS structure has standardized terms and concepts, established semantic relationships between them, which reduced the likelihood of errors and shortened the preparation time for retrospective analysis (by 20%), while improving the quality of analytical conclusions through more in-depth and comprehensive data processing.

Keywords: information technology, system architecture, sociological research, ontologies, forecasting, data processing system (DPS), decision support system (DSS).

DOI: 10.21293/1818-0442-2024-27-3-109-117

References

1. Data Collection and Labelling Market Size, Share, Trends and Analysis by Region, Type, Vertical, and Segment Forecast to 2030. Available at: <https://www.global-data.com/store/report/data-collection-and-labelling-market-analysis>, free (Accessed: July 30, 2024).

2. Data Collection and Labeling Market. Available at: <https://www.factmr.com/report/4726/data-collection-and-labelling-market>, free (Accessed: July 31, 2024).

3. Koshechkin K.A., Kozlovich A.V., Kotikov V.N., Mironov A.N. Obzor perechnya i osobennostey programmnovo obespecheniya, primenyaemogo dlya statisticheskoy obrabotki dannykh pri vypolnenii ekspertnoy raboty [Review of the list and the features of the software used for data processing in statistics when carrying out the expert evaluation], *The Journal Regulatory Research and Medicine Evaluation*, 2014, no. 3, pp. 46–51 (in Russ.).

4. Andreeva A.A. Vozmozhnosti ispolzovaniya professionalnogo statisticheskogo paketa dlya sotsialnykh nauk SPSS Statistics dlya analiza i obrabotki dannykh issledovaniya obshchestvennogo mneniya [The advantages of using a professional software package SPSS Statistics in social sciences to process the data collected from public views], *Theory and Practice of Modern Science*. 2020, no. 1(67), pp. 30–43 (in Russ.).

5. Orestova V.R., Bastron A.A. Primeneniye statisticheskogo paketa analiza dannykh SPSS Statistics v psikhologicheskikh issledovaniyakh na primere faktornogo analiza [The use of the SPSS Statistics for psychological research illustrated by factor analysis], *History and Archives*, 2017, no. 2(8), pp. 38–51 (in Russ.).

6. Demakov V.I., Larionova E.Yu., Golodkov Yu.E., Rerke V.I. Sravnitelnyy obzor statisticheskikh paketov dlya analiza dannykh [Comparative study of statistical packages for data processing], *Vestnik BGU. Mathematics, Informatics*, 2023, no. 3, pp. 78–89 (in Russ.).

7. Unguryanu T.N., Grzhibovskiy A.M. Odnofaktornyy dispersionnyy analiz s ispolzovaniem paketa statisticheskikh program [Single-factor analysis of variance using statistical package], *Stata // Human Ecology*, 2014, no. 5, pp. 60–64 (in Russ.).

8. Smirnov V.A. Novye kompetentsii sotsiologa v epokhu «Bolshikh dannykh» [Sociologist new competences in the Big Data Age], *Monitoring*, 2015, no. 2(125), pp.44–54 (in Russ.).

9. Andreyuk D.S., Livitina A.S., Sushko N.S. Semanticheskyy analiz struktury tsennostey gruppy s pomoshchyu tezavra Rozhe: avtomatizirovannyy algoritm [Semantic analysis of the group values structure using the Roget's Thesaurus: an automated algorithm], *Public Administration. E-journal (Russia)*, 2022, no. 91, pp. 148–161 (in Russ.).

10. Shchekotin E.V., Goyko V.L., Basina P.A., Bakulin V.V. Ispolzovanie mashinnogo obucheniya dlya izucheniya kachestva zhizni naseleniya: metodologicheskie aspekty [Using Machine Learning to evaluate the quality of life: methodological aspects], *Digital Sociology*, 2022, № 1, pp. 87–97 (in Russ.).

11. Ragimova N.A., Abdullaev V.G., Khalilov M.E. Primeneniye tekhnologiy bolshikh dannykh dlya monitoringa sotsialnykh protsessov v universitete [The use of Big Data technologies to monitor social processes at university], *ScienceRise*, 2019, no. 8(61), pp. 17–23 (in Russ.).

12. Geger A.E., Chupakhina Yu.A., Geger S.A. Kompyuternye programmy dlya analiza kachestvennykh i smeshannykh dannykh [Computer programs to analyze quality and mixed data], *St. Petersburg Sociology Today*, 2015, no. 6, pp. 374–383 (in Russ.).

13. Kanygin G.V., Koretskaya V.S. Analiticheskoe kodirovanie [Analytical coding], *St. Petersburg Sociology Today*, 2018, no. 10, pp. 99–121 (in Russ.).

14. Khuazhev A.A. Tekhnologicheskie sredstva analiza imidzhformiruyushchikh politicheskikh mediatekstav: vozmozhnosti i ogranicheniya [Tools to analyze the image making political media texts], *Izvestiya of Saratov University. Sociology. Politology*, 2022, no. 2, pp. 232–236 (in Russ.).

15. Chudnovskaya I.N., Bukhtiyarova I.N., Lipatova M.E. Model kommunikatsii v tsifrovom obrazovatelnom prostanstve v refleksii studentov [Communication model in digital education environment based on students opinions], *Sociology*, 2022, no. 1, pp. 173–183 (in Russ.).

16. Kharlamov A.A., Vasin B.I. Instrumentariy dlya informatsionno-analiticheskoy ekspertnoi otsenki nauchnoy produktssii s tsel'yu vyavleniya prioritnykh nauchnykh napravleniy i kollektivov [Tools for analytical expert evaluation of a research products aimed at determining the priority research directions and teams], *Vestnik of Moscow State Linguistic University. Humanities*, 2018, no. 6(797), pp. 164–174 (in Russ.).
17. Uchenye TGU s pomoshchyu platformy PolyAnalyst izmerili nastroyeniya rossiyan [TSU researchers evaluated the mood of the Russian citizens using PolyAnalyst platform], Available at: <https://www.megaputer.ru/uchenye-izmerili-shchaste-rossijan/>, free (Accessed: July 30, 2024) (in Russ.).
18. Vasilenko L.A., Meshcheryakova N.N. Sotsiologiya tsifrovogo obshchestva: monografiya [Digital Society Sociology: Monograph]. Tomsk: Publ. office of Tomsk Polytechnic University, 2021, 226 p. (in Russ.).
19. Utepbergenov A. O. Bolshie dannye v upravlenii che-lovecheskimi resursami [Big Data in Human Resources Management], *Economics and Socium*, 2022, no. 4–3 (95), pp. 439–456 (in Russ.).
20. Korshunov A., Beloborodov I., Buzun N., Avanesov V., Pastukhov R., Chikhhradze K., Kozlov I., Gomzin A., Andrianov I., Sysoev A., Ipatov S., Filonenko I., Chuprina K., Turdakov D., Kuznetsov S. Analiz sotsialnykh setey: metody i prilozheniya [Social Networks analysis: techniques and applications], *Proceedings of the Institute for System Programming of the RAS*, 2014, no. 1, pp. 439–456 (in Russ.).
21. Grudanova A.A., Grudanov N.A. Obzor instrumentov onlayn-oprosov dlya informatsionnogo obespecheniya market-ingovykh issledovaniy [Overview of computer-assisted web interviewing tools used in market research], *StudNet*, 2021, no. 7, pp. 1814–1829 (in Russ.).
22. Nikonorova M.L. Kompyuternaya model resheniya zadach klassifikatsii v programmnoy srede Rapid Miner [Computer model to solve classification tasks in programming environment Rapid Miner], *Medical Education and Professional Development*, 2017, no. 2–3, pp. 28–29 (in Russ.).
23. Vasilkova V.V., Legostaeva N.I. Boty na publichnykh arenakh sotsialnykh setey [Bots in social networks], *Sociological Journal*, 2021, no. 4, pp. 99–117 (in Russ.).
24. Goncharova I.V., Pronchev G.B. Tekhnologii bolshikh dannykh kak instrument resheniya zadach ustoychivogo razvitiya obshchestva [Big data technologies as a tool for solving the problems of sustainable development of the society], *Law and Management*, 2023, no. 5, pp. 215–226 (in Russ.).
25. Rogoza A.I. Sravnitelnyy analiz rossiyskikh sistem Business Intelligence [Comparative analysis of a Russian systems «Business Intelligence»], *Skif*, 2022, no. 8(72), pp. 282–287 (in Russ.).
26. Khasanov A.R. Vliyanie prediktivnoy analitiki na deyatelnost kompaniy [Impact of predictive analytics to the company operation], *Strategic Decisions and Risk Management*, 2018, no. 3, pp. 108–113 (in Russ.). DOI: 17747/2078-8886-2018-3-108-113.
27. Tsypin A.P., Sorokin A.S. Statisticheskie pakety programm v sotsialno-ekonomicheskikh issledovaniyakh [Statistical packages in socio-economic studies], *ASR: Economics and Management*, 2016, no. 4 (17), pp. 379–384 (in Russ.).
28. Shimanovskiy K. V. Importozameshchenie v oblasti programmno obespecheniya biznes-analitiki [Import-substituting in software for business analytics], *International Journal of Applied Sciences and Technology «Integral»*, 2022, no. 2, pp. 52–537 (in Russ.).
29. Borkhanova A.D. Primeneniye ankety Google Forms dlya izucheniya pedagogicheskikh usloviy realizatsii innovatsionnykh projektov v shkole [Using Google Forms to study the pedagogical conditions of the implementation of innovative projects at schools], *Pedagogic Correction Education Journal*, 2021, no. 3(27), pp. 32–36 (in Russ.).
30. Syzdykova A., Orakbay L., Mansharipova A., Shuzheev B., Kushniyazova M. The Google Forms Online Service is Used to Assess the Socio-psychological Qualities of a Person and the Level of Professional Competence of Average Medical Professionals. *Actual Problems of Theoretical and Clinical Medicine*, 2021, no. 2, pp. 35–40
31. Ked A.P., Agaeva P.M. Internet-opros kak metod sotsiologicheskogo issledovaniya [Computer-aided interview as a tolls used for sociological study], *Problems of Modern Economy (Novossibirsk)*, 2015, no. 27, pp. 112–116 (in Russ.).
32. Pantelev D.N. Issledovanie i kharakteristika market-ingovogo atributa innovatsiy [Study and characteristics of marketing aspects in innovation], *Bulletin of the Academy of Knowledge*, 2022, no. 6(53), pp. 453–456 (in Russ.).
33. Atalyan A.V., Kolesnikova L.I., Kolesnikov S.I., Grzhibovskiy A.M., Suturina L.V. Informatsionnaya sistema REDCap dlya sbora i khraneniya dannykh populatsionnykh biomeditsinskikh issledovaniy [REDCap information system to collect and store the data collected from population and biomedical studies], *Human Ecology*, 2019, no. 2, p. 52–59 (in Russ.).
34. Kuznetsov A.M. Struktura modeli informatsionno-analiticheskoy sistemy obrabotki sotsiologicheskikh dannykh [Model structure for information analytical system used for sociological data processing], *Nano-biotechnologies, Heat and Power Engineering, Mathematical Modeling Proceedings of the International Conference*, Lipetsk, 2023, p. 258–262 (in Russ.).
35. Saklakov V.M. Metodologiya tsifrovogo sotsiologicheskogo issledovaniya: obshchestvennaya sistema kak bazovyy instrument modelirovaniya [Methodology of digital sociological research: The social system as a basic modeling tool], *Proceedings of TUSUR University*, 2023, vol. 26, no. 4, pp. 61–77. DOI: 10.21293/1818-0442-2023-26-4-61-77 (in Russ.).
36. Certificate of state registration of computer program No. 2024617985. / A.M. Kuznetsov. – Application No. 202461581. Date of receipt: March 22, 2024. Registered in the Register of computer programs: April 08, 2024. (in Russ.).
37. Certificate of state registration of computer program No. 2024666569. A.M. Kuznetsov. Application No. 2024665224. Date of receipt: July 3, 2024. Registered in the Register of computer programs: July 15, 2024 (in Russ.).
38. Kuznetsov A.M. Vybor modeli i algoritma prognozirovaniya sobytii na osnove rezul'tatov obrabotki dannykh sotsiologicheskikh issledovaniy [Choosing a model and an algorithm to predict an event based on the results of sociological data processing], *Neuro-Computers and their Application: proceedings of the XXI All-Russian Research Conference*, Moskva: MGPPU, 2023. p. 93–94 (in Russ.).
39. Kurbatova A.S. Perspektivy importozameshcheniya v rossiyskom sektore informatsionnykh tekhnologiy v usloviyakh sanktsionnogo davleniya [Import-substituting prospects in Russian information technologies in sanctions pressure conditions], *Modern World Economy*, 2023, vol. 1, no. 4(4). DOI: 10.17323/2949-5776-2023-1-4-108-128 (in Russ.).
40. Zlobina O. V., Peshkova G. Yu. Perspektivy avtomatizatsii i tsifrovizatsii proizvodstva v usloviyakh vvedeniya sanktsiy [Automation and Digitalization prospects under sanctions], *Bulletin of the Altai Academy of Economics and Law*, 2022, no. 8–1, pp. 66–73 (in Russ.).

Alexander M. Kuznetsov

Postgraduate student, Department of Information Systems and Information Security,
Tambov State Technical University (TSTU)
106, Sovetskaya st., bldg. 5 Tambov, Russia, 392000
ORCID: 0000-0003-2640-3417
Phone: +7 (475-2) 63-00-54
Email: 068t@mail.ru