

14. СОВЕРШЕНСТВОВАНИЕ СИСТЕМЫ АНАЛИЗА ДАННЫХ НАЦИОНАЛЬНОГО СТАТИСТИЧЕСКОГО КОМИТЕТА РЕСПУБЛИКИ БЕЛАРУСЬ НА ОСНОВЕ ТЕХНОЛОГИИ BIG DATA

Парфёнов А. В., студент гр.272301, Кудин И. С., студент гр.272301

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Лазаревич И.М. – канд. экон. наук., доцент

Аннотация: В статье представлены результаты исследования сущности технологии Big Data и возможности ее внедрения при анализе социально-экономических данных Белстат. Описаны основные элементы, методы и инструменты системы анализа данных на основе Big Data, преимущества и недостатки ее использования. В работе предложено решение для внедрения такой системы в Белстат, описан возможный эффект.

Ключевые слова. Технология Big Data, Белстат, аналитика, социально-экономические процессы.

На современном этапе развития, информация является еще одним важным национальным стратегическим ресурсом. В условиях цифровой трансформации общества особую актуальность приобретает оперативное получение информации, способствующей повышению конкурентоспособности экономики страны в целом, субъектов хозяйствования в частности. Технология Big Data, ставшая в наше время неотъемлемой частью информационной инфраструктуры, открыла огромные возможности в сборе, хранении, обработке и анализе данных.

Big Data – это структурированные или неструктурированные массивы данных большого объема, которые обрабатывают при помощи специальных автоматизированных инструментов для статистики, анализа, прогнозов и принятия решений. Big Data характеризуется следующими признаками: объём данных, скорость передачи данных, разнообразие данных, достоверность, ценность, изменчивость. При обработке и анализе данных используются такие инструменты и технологии как: специальное программное обеспечение (NoSQL, MapReduce, Hadoop, R); Data mining; ИИ и нейросети и др. К основным методам анализа Big Data относят: описательную, прогнозную, предписательную, диагностическую аналитику [1]. Эффективное использование данных способствует разработке более точных и адаптивных стратегий поддержки инноваций. Предоставление информации о текущих тенденциях и потребностях рынка позволяет государству и бизнес-сектору выстраивать свои действия в соответствии с динамикой спроса и технологическими изменениями.

Одной из важнейших областей применения технологии Big Data является анализ социально-экономических процессов, которые играют ключевую роль в принятии решений на уровне государства, бизнеса и общества в целом. В различных странах и организациях аналитические системы на основе больших данных демонстрируют свою эффективность. Например, в Соединенных Штатах аналитические системы используются для прогнозирования и предотвращения преступлений, оптимизации городского транспорта и предоставления персонализированных медицинских услуг. В Европейском союзе такие системы помогают в борьбе с финансовым мошенничеством, оптимизации поставок и снабжения, а также в разработке политики общественного здравоохранения.

Big Data открывает новые возможности для Национального статистического комитета Республики Беларусь (Белстат) и его пользователей. Белстат представляет собой независимый при осуществлении государственной статистической деятельности республиканский орган

государственного управления в области государственной статистики, осуществляющий регулирование и управление, а также координацию деятельности других государственных органов в этой области. Подчиняется Президенту Республики Беларусь. Цель работы – удовлетворение потребности в статистической информации различных категорий пользователей (общества, государства, предприятий и международного сообщества) [2].

Белстат сталкивается с различными проблемами в работе с данными, включая их обработку, анализ и интерпретацию. Некоторые из этих проблем включают в себя ограниченные ресурсы, несогласованность данных, отсутствие оперативности и сложности в анализе больших объемов информации. Внедрение новой системы анализа данных на основе Big Data позволит решить эти проблемы, обеспечивая оперативный доступ к качественной и достоверной информации, а также автоматизируя процессы анализа и принятия решений.

Внедрение системы анализа данных на основе Big Data имеет как положительные, так и отрицательные эффекты. Преимущества и недостатки системы для различных категорий пользователей систематизированы в таблице 1.

Внедрение системы анализа данных на основе Big Data для использования статистиками может предоставить государству возможность самостоятельного управления и использования этой информации. Она предоставит ученым и государственным служащим возможность беспрепятственно и наилучшим образом использовать данные для проведения исследований, разработки политик и мониторинга социально-экономического развития.

Таблица 1 – Преимущества и недостатки системы анализа данных на основе Big Data

Преимущества	Недостатки
Увеличение оперативности принятия решений	Зависимость от качества данных, предоставляемых Белстат
Повышение качества предоставляемой информации	Возможные ошибки алгоритмов машинного обучения
Оптимизация процессов управления	Необходимость постоянного обновления и адаптации алгоритмов к изменяющейся ситуации
Повышение конкурентоспособности	Риски приватности данных и конфиденциальности
Автоматизация процессов анализа	Сложности в интерпретации результатов
Рост эффективности деятельности	

Белстат для эффективного применения системы анализа данных на основе Big Data будут целесообразными следующие меры

1. Провести организационное преобразование посредством создания отдельного структурного подразделения, специализирующегося на проведении описательной, прогнозной, предписательной, диагностической аналитической работы.

2. Создать инфраструктуру сбора и хранения данных, обеспечивающую высокопроизводительные и масштабируемые базы данных для хранения больших объемов информации.

3. Совершенствовать механизм сбора данных, включая применение различных источников, таких как открытые источники, API, базы данных и др.

4. Разработать и применить алгоритмы машинного обучения, статистических методов и других аналитических инструментов для обработки и анализа данных.

5. Создать пользовательский интерфейс и инструменты для визуализации результатов анализа данных в понятной и наглядной форме.

6. Разработать системы уведомлений для оперативного оповещения пользователей о важных изменениях и аномалиях в данных.

7. Обеспечить безопасность и конфиденциальность данных во время их хранения, передачи и обработки.

8. Интегрировать новую систему с существующими информационными системами и базами данных для обеспечения совместимости и обмена данными.

9. Обеспечить регулярное обновление и поддержку системы для устранения ошибок, добавления новых функций и адаптации к изменяющимся потребностям и требованиям пользователей.

Становится заметно, что вышеперечисленные меры потребуют определения технических требований к системе, которые будут составлять ее основу. Технические требования должны состоять из необходимых функций, обосновывающих сущность всей системы и являющихся фундаментально важными для достижения поставленной цели.

Система анализа данных на основе Big Data будет способна выявлять как значимые, так и незначительные изменения в исследуемых процессах на временных интервалах различной длины, обеспечивая пользователям более оперативное и точное принятие решений. Данные возможности

должны быть реализованы в полной мере и первую очередь, определяя собой основные функции системы анализа данных. Данные функции являются фундаментальной особенностью самой системы, и их реализация будет важнейшей целью всей работы. Также для повышения эффективности системы и улучшения удобства использования могут быть реализованы дополнительные функции (таблица 2) [3; 4].

Таблица 2 – Основные и дополнительные функции системы анализа данных

Функции	Назначение
Основные	
Оперативный анализ в реальном времени	Осуществление анализа данных в реальном времени для оперативного реагирования на любые изменения факторов
Использование алгоритмов машинного обучения	Автоматическое обнаружение даже незначительных изменений в данных
Быстрый анализ больших объемов данных	Анализ данных в реальном времени для оперативного реагирования на любые изменения факторов
Использование статистических методов и тестов	Проверка статистической значимости обнаруженных изменений для подтверждения в их достоверности и значимости
Визуализация данных	Наглядная демонстрация изменений в статистиках на кратких периодах времени
Оповещения и уведомления	Предоставление пользователю возможности настройки оповещений и уведомлений о любых изменениях в статистиках для оперативного реагирования на них
Дополнительные	
Прогнозирование и моделирование	Прогнозирование будущих тенденций и моделирование различных сценариев на основе загруженных данных
Коллаборативный анализ	Создание возможности для коллективного анализа и обмена опытом между пользователями
Поддержка множества языков и форматов	Доступность для пользователей из разных стран и областей с разными предпочтениями и требованиями
Интеграция с внешними источниками данных	Получение более полной «картины» и улучшение качества анализа данных

Получателями информации системы анализа данных Big Data Белстат могут стать:

1. Правительство Республики Беларусь, министерства и ведомства, эксперты и ученые для проведения мониторинга социально-экономического развития, эффективного принятия решений и разработки программ развития страны.

2. Субъекты предпринимательской деятельности для формирования условий их устойчивого эффективного функционирования и развития.

3. Международные интеграционные объединения. На текущий момент – ЕАЭС при формировании интеграционной статистической системы для проведения сравнительного анализа различных процессов развития стран-участниц Союза.

Оценка затрат и ресурсов, необходимых для внедрения и поддержки новой системы, включая обучение персонала и обновление программного обеспечения, является важным этапом. На стадии внедрения проекта источниками финансирования могут стать бюджетные средства или средства белорусского инновационного фонда. В последующем, расходы по обслуживанию системы могут быть разделены на различные источники и внедрена дифференцированная система оплаты услуг. Например, для государственных институтов – за счет бюджетных средств, тарифы за услуги обеспечивают только окупаемость затрат, для субъектов предпринимательской деятельности – на платной основе по тарифам, окупающим затраты и обеспечивающим эффективную прибыльность.

Таким образом, создание специализированного подразделения Белстат, осуществляющего разработку инфраструктуры для обработки и хранения данных на основе Big Data, аналитическую работу помогут обеспечить высокую эффективность анализа и улучшить оперативность принятия решений национальными институтами различных уровней.

Список использованных источников:

1. Абдирахимов, И. Э. Проблемы и решение в Big Data. // Цифровые технологии в промышленности, 2023.– С.158-163.
2. Задачи и функции Белстат. [Электронный ресурс]. – Режим доступа: https://www.belstat.gov.by/o-belstate_2/sistema-belstata/tsentralnyi-apparat-belstata_4/zadachi-i-funktsii-belstata/. – Дата доступа: 05.03.2024.
3. Кригер, А. Б. Цифровые технологии мониторинга развития аутсорсинга в регионах дальневосточного федерального округа/ А. Б. Кригер, В. В. Ивин, А. С. Боначева // Современные технологии управления, 2021. – С.1-8.
4. Магеррамов, З. Т. Big Data: проблемы, методы анализа, алгоритмы / З. Т. Магеррамов, В.Г. Абдуллаев, А. З. Магеррамова // Радиоэлектроника и информатика. 2017. – С.42-53.