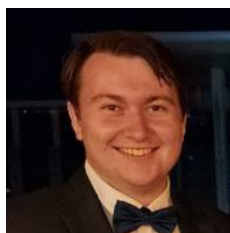


УДК 004.622

## BIG DATA: ПРОБЛЕМЫ И ТЕХНОЛОГИИ



**Д.И. Альховик**  
Студент 1 курса, кафедры  
информатики  
25350006@study.bsuir.by



**В.Д. Владымец**  
Ассистент кафедры  
информатики, инженер  
программист ОИАСУ ЦИИР  
БГУИР  
v.vladymtsev@bsuir.by



**А.Н. Марков**  
Старший преподаватель,  
заместитель начальника  
Центра информатизации и  
инновационных разработок  
БГУИР  
a.n.markov@bsuir.by

**Д.И. Альховик**

Студент 1 курса специальности «Информатика и технологии программирования» БГУИР

**В.Д. Владымец**

Ассистент кафедры информатики, инженер-программист ОИАСУ ЦИИР БГУИР

**А.Н. Марков**

Старший преподаватель кафедры ПОИТ, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

**Аннотация.** Данный доклад рассматривает тему Big Data и его применение в различных областях, таких как медицина, финансы, транспорт, розничная торговля, государственное управление и интернет вещей. В докладе приведены конкретные примеры использования Big Data в каждой из этих областей, а также описаны потенциальные выгоды и преимущества, которые может принести использование данных большого объема. В заключении доклада подчеркивается, что примеры применения Big Data продолжают развиваться и расширяться, открывая новые возможности для улучшения бизнес-процессов и повышения эффективности в различных областях.

**Ключевые слова:** Big Data, анализ данных, медицина, финансы, транспорт, розничная торговля, государственное управление, IoT, примеры использования.

### Введение.

В современном мире мы сталкиваемся с огромными объемами данных, которые постоянно генерируются из различных источников. Эти данные могут содержать ценную информацию, но для их обработки и анализа необходимы специальные технологии и инструменты. В связи с этим, концепция Big Data становится все более актуальной и востребованной.

Однако, работа с Big Data также включает в себя ряд проблем и вызовов, связанных с объемом, скоростью, разнообразием и качеством данных, а также безопасностью и доступностью информации. Необходимость обработки больших объемов данных ставит перед нами задачу разработки новых технологий и инструментов, которые позволят эффективно работать с Big Data.

В данном докладе мы рассмотрим основные проблемы и вызовы, связанные с Big Data, а также технологии и инструменты, которые используются для работы с такими данными. Мы также рассмотрим примеры применения Big Data в различных областях, таких как медицина, банковская сфера, ритейл, транспорт и наука. В конце доклада мы обсудим значение Big Data в будущем и дадим рекомендации для дальнейших исследований.

### Проблемы Big Data.

Big Data представляет собой набор данных, которые могут быть слишком большими, сложными и разнообразными для обработки традиционными методами. Объем и скорость

создания данных постоянно растет, что приводит к ряду проблем, связанных с их обработкой и анализом. Ниже мы рассмотрим некоторые из основных проблем, связанных с Big Data:

1. Объем данных: с постоянным ростом количества данных, с которыми мы работаем, существует необходимость в большем объеме вычислительных ресурсов и хранилищ для их сохранения и обработки.

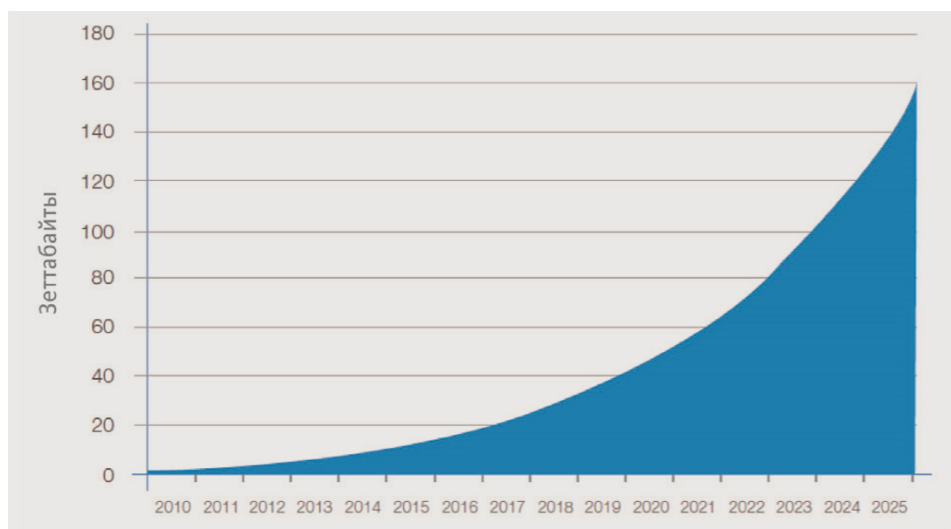


Рисунок 1 – Рост объема данных в мире.

2. Скорость обработки данных: некоторые приложения, такие как интернет-поиск или реклама в режиме реального времени, требуют мгновенной обработки и анализа данных. Это ставит перед нами задачу оптимизации производительности и скорости обработки данных.

3. Разнообразие данных: данные могут поступать из различных источников и в различных форматах. Это усложняет процесс их обработки и анализа, так как необходимы различные инструменты и методы для работы с различными типами данных.

4. Качество данных: не все данные являются полезными или точными. Некоторые могут быть неполными, некорректными или содержать ошибки. Это может привести к неточным выводам и решениям, основанным на этих данных.

5. Безопасность данных: с увеличением объема данных также возрастает риск их утечки, взлома или кражи. Это может привести к серьезным последствиям для компаний и организаций, которые хранят и обрабатывают конфиденциальные данные.

6. Недостаток опытных специалистов: работа с Big Data требует специализированных знаний и навыков. Однако, на данный момент спрос на таких специалистов значительно превышает их количество, что создает дополнительные проблемы в работе с Big Data.

### **Технологии обработки Big Data.**

Технологии обработки Big Data развиваются с высокой скоростью, чтобы обеспечить эффективную и быструю обработку большого объема данных. В данном контексте, стоит отметить такие технологии как:

Apache Hadoop [1]: распределенная система хранения и обработки данных, которая использует большое количество узлов серверов для распределения задач по обработке данных. Она позволяет обрабатывать большие объемы данных параллельно и снижает время обработки.

Apache Spark [7]: открытая платформа обработки больших данных, которая позволяет выполнять вычисления в памяти, что ускоряет обработку данных и улучшает производительность. Она также поддерживает обработку потоков данных, машинное обучение и графовые вычисления.

NoSQL базы данных [4]: это альтернатива реляционным базам данных, которая может обрабатывать неструктурированные данные и гораздо лучше масштабируется для работы с большими объемами данных. Примеры NoSQL баз данных включают MongoDB и Cassandra.

HBase [1]: это распределенная база данных, которая предназначена для работы с большими объемами структурированных данных. Она позволяет хранить, обновлять и извлекать данные в режиме реального времени, а также обеспечивает масштабируемость и отказоустойчивость.

Apache Flink [4]: это открытая платформа обработки потоков данных и батч-процессинга. Она позволяет обрабатывать данные в реальном времени, обеспечивает высокую производительность и масштабируемость.

Эти технологии являются только некоторыми из инструментов, которые используются для обработки больших объемов данных, и их выбор зависит от конкретных требований и задач, которые необходимо решить.

### **Примеры использования Big Data.**

Примеры использования Big Data разнообразны и могут включать в себя различные отрасли и области. Некоторые примеры использования Big Data включают:

1. Медицина: использование данных о здоровье, пациентах и их истории болезни для улучшения диагностики и лечения, а также оптимизации системы здравоохранения.

2. Финансы: анализ данных о транзакциях, потребительском поведении и финансовых рынках для прогнозирования трендов и управления рисками.

3. Транспорт: использование данных о движении транспорта и поведении водителей для улучшения безопасности дорожного движения и оптимизации транспортных потоков.

4. Розничная торговля: анализ данных о покупках, поведении потребителей и тенденциях рынка для улучшения продуктового ассортимента, ценообразования и управления запасами.

5. Государственное управление: использование данных о социально-экономическом развитии, демографических трендах и других параметрах для принятия решений в области политики и управления ресурсами.

6. Интернет вещей (IoT): использование данных, собираемых от различных устройств и датчиков, для управления умными домами, городами и промышленными системами [8-13].

Примеры использования Big Data продолжают развиваться и расширяться, открывая новые возможности для улучшения бизнес-процессов и повышения эффективности в различных областях.

### **Заключение.**

Big Data является важной технологией в современном мире, позволяющей организациям собирать, хранить и анализировать большие объемы данных. Примеры использования Big Data разнообразны и могут быть найдены в различных отраслях и областях, включая медицину, финансы, транспорт, розничную торговлю, государственное управление и IoT. Big Data помогает организациям принимать более обоснованные решения, повышать эффективность бизнес-процессов и улучшать качество продуктов и услуг. Однако, использование Big Data также представляет вызовы, такие как необходимость обеспечения безопасности и конфиденциальности данных, а также необходимость наличия высококвалифицированных специалистов для работы с данными. Несмотря на эти вызовы, использование Big Data остается важной стратегией для организаций, желающих оставаться конкурентоспособными в современном цифровом мире.

### **Список литературы**

- [1] IBM. "Big data analytics." IBM. <https://www.ibm.com/analytics/big-data-analytics>  
[2] Amazon. "Big data." Amazon Web Services. <https://aws.amazon.com/big-data/> [3] Google. "What is big data?" Google Cloud. <https://cloud.google.com/big-data/what-is-big-data>  
[4] C. Mohan. "Big data: Promises, challenges and opportunities." IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 6, 2014, pp. 1526-1537.

- [5] V. G. Cerf and R. E. Kahn. "A protocol for packet network intercommunication." IEEE Transactions on Communications, vol. 22, no. 5, 1974, pp. 637-648.
- [6] D. L. Mills. "Internet time synchronization: the Network Time Protocol." IEEE Transactions on Communications, vol. 39, no. 10, 1991, pp. 1482-1493.
- [7] M. Zaharia, et al. "Apache Spark: a unified engine for big data processing." Communications of the ACM, vol. 59, no. 11, 2016, pp. 56-65.
- [8] McKinsey & Company. "The age of analytics: Competing in a data-driven world." McKinsey & Company, December 2016.
- [9] Forbes. "10 powerful examples of big data analytics in healthcare."
- [10] IBM. "Big data in banking: How to use big data analytics."
- [11] KD Nuggets. "7 innovative examples of big data in action."
- [12] Gartner. "Gartner Top 10 Strategic Technology Trends for 2022."
- [13] SAS. "Big data analytics." SAS

## **BIG DATA: PROBLEMS AND TECHNOLOGIES**

***D. I. Alkhovik***

*Student of BSUIR, software  
engineer*

***V.D. Vladymtsev***

*Bechalar engineer system-  
programmer, Department of  
Integrated Automated Control  
Systems*

***A.N. Markov***

*Senior lecturer of the department,  
Deputy head of the Center for  
Informatization and Innovative  
Developments*

*Center for Informatization and Development of the Belarusian University of State Informatics and  
Radioelectronics, Republic of Belarus*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*

*E-mail: 25350006@study.bsuir.by, v.vladymtsev@bsuir.by, a.n.markov@bsuir.by*

**Annotation.** This report examines the topic of Big Data and its application in various fields, such as medicine, finance, transport, retail, public administration and the Internet of Things. The report provides specific examples of the use of Big Data in each of these areas, as well as describes the potential benefits and advantages that the use of large-volume data can bring. In conclusion, the report emphasizes that the examples of Big Data applications continue to develop and expand, opening up new opportunities for improving business processes and increasing efficiency in various areas.

**Keywords:** Big Data, data analysis, medicine, finance, transport, retail, public administration, IoT, usage examples.