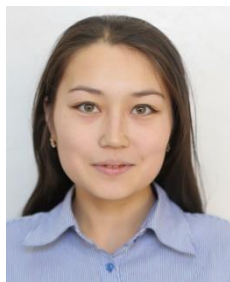


УДК 004.67

ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ BIG DATA В ПРОМЫШЛЕННОСТИ КАЗАХСТАНА



А.С. Амирова

Докторант Евразийского национального университета им.Л.Н.Гумилева



А.Т.Тохметов

Доцент кафедры информационных систем Евразийского национального университета им.Л.Н.Гумилева, кандидат физико-математических наук

Евразийский национальный университет им.Л.Н.Гумилева, Республика Казахстан

E-mail: akzhibek.amirova@gmail.com

А.С. Амирова

Окончила Евразийский национальный университет им. Л.Н.Гумилева. Докторант ЕНУ им. Л.Н.Гумилева.

А.Т.Тохметов

Доцент кафедры информационных систем Евразийского национального университета им.Л.Н.Гумилева, кандидат физико-математических наук

Аннотация. В статье сделан обзор технологий Big Data, используемых на предприятиях Казахстана, рассматриваются проблемы и перспективы дальнейшего развития и использования систем обработки больших данных в промышленности Казахстана. Проведен анализ проблем внедрения технологий Big Data в промышленной сфере.

Ключевые слова: большие данные, анализ данных, промышленный интернет, цифровизация промышленности.

Введение. Одной из важнейших задач в рамках реализации Государственной программы «Цифровой Казахстан» является цифровизация промышленности, где акцент будет сделан на развитие собственных технологий и компетенций, популяризацию цифровых технологий, таких как Big Data, искусственный интеллект, машинное обучение и промышленный интернет вещей.

В современных условиях организации создают большое количество неструктурированных данных, таких как текстовые документы, изображения, видеозаписи, машинные коды, таблицы. Вся эта информация хранится во множестве репозиториях, порой даже за пределами организации. Компании могут иметь доступ к огромному массиву собственных данных и не иметь необходимых инструментов, которые могли бы установить взаимосвязи между этими данными и сделать на их основе значимые выводы. Традиционные методы анализа информации не могут угнаться за огромными объемами постоянно растущих и обновляемых данных, что в итоге и открывает дорогу технологиям Big Data.

Большие данные (Big Data) – это массивы данных, генерируемых действиями, производимыми электронным путем либо в результате коммуникации между электронными устройствами.

Можно выделить следующие особенности технологий Big Data [1, 2]:

- работа с информацией огромного объема и разнообразного состава;
- информация весьма часто обновляется и находится в разных источниках;
- качественно отличающийся метод открывающей аналитики для выявления практических знаний, которые непосредственно монетизируются в прибыль;
- наглядное отображение отчетов и возможности сценарного анализа («что, если...»);
- цель применения технологий Big Data – увеличение эффективности работы, создание новых продуктов и повышение конкурентоспособности.

Технологии Big Data могут быть полезны для решения следующих задач [3]:

- прогнозирование рыночной ситуации;
- маркетинг и оптимизация продаж;
- эффективное сегментирование клиентов;
- совершенствование товаров и услуг;
- принятие более обоснованных управленческих решений на основе анализа Big Data;
- оптимизация портфеля инвестиций;
- повышение производительности труда;
- эффективная логистика;
- мониторинг состояния основных фондов.

Согласно отчету компании McKinsey «Global Institute, Big data: The next frontier for innovation, competition, and productivity» [4], большие данные стали важным фактором производства наряду с трудовыми и капитальными ресурсами. Использование больших данных станет основой конкурентного преимущества и роста компаний.

Согласно прогнозу, опубликованному в докладе аналитической фирмы «Эра данных 2025» к 2025 году объем всех данных во всем мире составит 163 зеттабайт. Это в 10 раз больше, чем общий объем данных по состоянию на 2016 год (рисунок 1) [5].

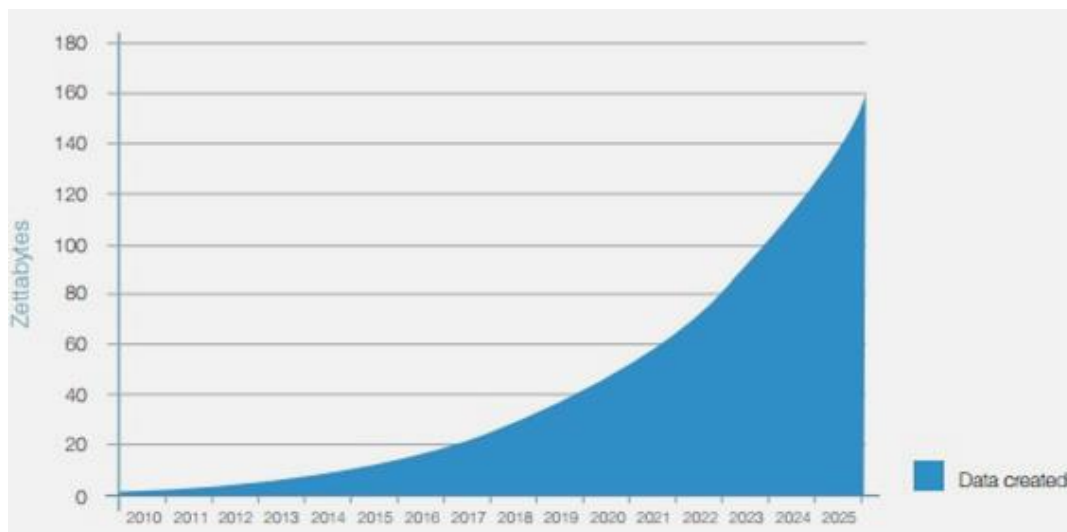


Рисунок 1. – Динамика роста общего объема данных в мире к 2025 году.

Несмотря на небольшой срок существования технологии Больших данных, уже есть оценки эффективного использования на реальных примерах. Один из самых высоких показателей относится к энергетике. Так, по оценкам аналитиков, аналитические технологии BigData способны на 99 % повысить точность распределения мощностей генераторов [6].

Приведем несколько примеров применения технологии BigData в различных отраслях [7]:

–энергетика (влияние погоды на генерацию энергии, анализ данных от «умных» счетчиков, исследовательские инфраструктуры для эффективного использования энергии в зданиях);

–наука (Большой Адронный коллайдер - общеевропейская инфраструктура для оценки качества при тестировании наноматериалов);

–E-Commerce (анализ поведения покупательских моделей, моделирование поведения клиентов);

–транспорт (влияние погоды и трафика на доставку и потребление грузов);

–колл-центр (анализ расшифровок разговоров для понимания поведения клиентов);

–финансы (решения по финансовым рискам, анализ мнения клиентов);

–ИТ (анализ логов от разных транзакционных систем);

–телекоммуникационные сети (анализ операций и сбоев сети).

В Казахстане на 2018 г. 80% предприятий обрабатывающей промышленности и 60% предприятий добывающей промышленности находятся на уровне полуавтоматизированных операций, или, говоря иначе, на уровне индустриализации 2.0, либо на этапе перехода к автоматизированному производству.

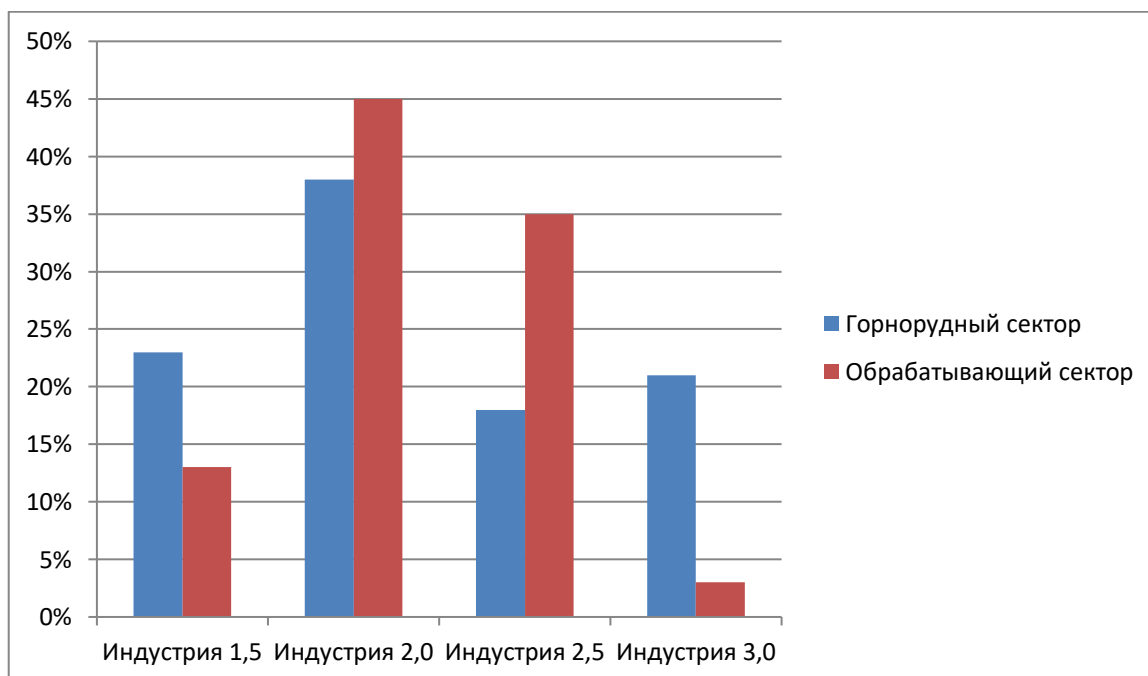


Рисунок 2. – Состояние промышленности Казахстана на 2018 г [8].

По данным специалистов, если в мире технологии BigData стала популярной в 2000-е, то в Казахстане лишь в конце 2016 года. В Казахстане крупным заказчиком по обработке данных выступает государство, за период независимости в стране собралось огромное количество данных, их нужно обрабатывать и использовать, чтобы контролировать и конкурировать вне страны. Так, например, одно из крупных внедрений технологии BigData будет осуществлено в министерстве финансов Республики Казахстан. Также это направление стратегически важно для статистики и прогнозирования экономики [9]. В Министерстве здравоохранения РК уже начали создавать единую базу данных пациентов с цифровой историей заболеваний. При первом обращении пациента в поликлинику, данные на него сразу

же вводятся в компьютер, далее по сети в единую базу. Когда такая картотека данных будет полностью оцифрована, можно прогнозировать улучшение работы системы здравоохранения.

Как известно область применения технологии Big Data огромна. Например, используя большие данные, вы можете узнать о предпочтениях клиентов, эффективности производства или провести анализ рисков финансовой операции. Большие данные – это уже устоявшаяся область технологий, даже, несмотря на ее относительно молодой возраст, который получил широкое распространение во многих сферах бизнеса и играет важную роль в развитии компаний. Благодаря внедрению передовых технологий, бизнес-лидеры могут не только принимать более эффективные решения, но и совершенствовать производственные процессы и решать ключевые бизнес-проблемы. Например, своевременное получение важной информации позволяет повысить производительность и повысить безопасность на предприятии.

Одной из перспективных отраслей промышленности Казахстана в плане внедрения технологии больших данных является горнорудная отрасль, в совершенствовании которой может помочь использование больших данных и применение индустриального Интернета вещей (IIoT). «Применение» IIoT подразумевает под собой внедрение ряда датчиков на всех этапах переработки горнорудного материала. Это позволит систематизировать данные производственного процесса и поднять уровень переработки.

Big Data находятся на службе у маркетинговых подразделений казахстанских компаний – например, у АО «Казпочта», телекоммуникационных фирм, страховых компаний, ритейлеров. Они изучают поведение покупателей для выстраивания эффективной стратегии и продвижения на рынке предложений, которые потенциально будут пользоваться большим спросом.

Исходя из анализа данных, собранных благодаря технологии Больших данных, в столице РК – Нур-Султане уже функционирует система «умных светофоров», позволившая, по данным столичного акимата, сократить заторы на дорогах на 21% [10].

Благодаря Big Data можно, в частности, отследить автомобиль, нарушивший правила дорожного движения, по камерам видеонаблюдения. Потенциал есть также в фоссилизации грузоперевозок, анализа пропускной способности дорог, пассажиропотоков. Big Data могут стать большим помощником и в сфере оказания государственных услуг, и с точки зрения информированности граждан о проводимой политике, и оценки реализации выдвинутых инициатив для корректировки и улучшения работы соответствующих структур. Интересный пример – применение МВД РК технологии Big Data в анализе и предупреждении правонарушений: в реальном времени на карте городов отображается «криминогенность» того или иного района: данные о правонарушениях для такой карты находятся в открытом доступе на сайте Генпрокуратуры, а информация о новых случаях обновляется в течение недели. В дальнейшем использование имеющейся информации правоохранительными органами может сыграть значительную роль в профилактике преступлений.

С целью изучения и практического применения технологий Big Data в Казахстане на базе АО «Национальные информационные технологии» была создана Лаборатория «Больших Данных». В настоящее время здесь ведутся исследовательские работы по анализу профиля пользователя портала eGov. Основными источниками данных являются: портал электронного правительства, центры обслуживания населения, электронные обращения граждан через почтовые сервисы, СМС-сообщения и звонки в Единый контакт-центр, социальные сети. Например, в настоящее время портал электронного правительства eGov содержит около 15 ТБайт информации, которая, в свою очередь, ежемесячно пополняется на 1 ТБайт. Кроме этого, в планах лаборатории Больших данных обрабатывать не менее 3 ТБайтов данных из других открытых источников [11]. Мировые достижения «в числах» показывают, что технология Больших данных активно развивается и уже достигла положительных результатов.

Применение технологии больших данных в самых разных областях, такие как здравоохранение, банки, розничная торговля, операторы связи и многие из которых стандартизованы или активно стандартизируются как на национальном, так и на международном уровнях. В рамках национальной стандартизации в соответствии с Государственной программой «Информационный Казахстан - 2020» в этом году проводятся работы по разработке проекта СТ РК «Большие данные. Облачные вычисления на основе требований и возможности» (на основе У.3600: Большие данные - облачных вычислений на основе требований и возможностей). Проект СТ РК «Большие данные. Облачные вычисления на основе требований и возможности» предусматривает подход использования облачных вычислений для удовлетворения существующих проблем в использовании больших данных [12].

Big Data уже нашли утилитарное применение в промышленности, но пока внедрение этих технологий в РК сдерживается устаревшим оборудованием, с которого трудно получать информацию, и неготовностью инфраструктуры «читать» и анализировать данные. Основные проблемы с внедрением Big Data находятся на уровне понимания самой сути технологии. Так, в стране до сих пор есть непонимание самого феномена, который путают только с наличием массива данных, а не методами их обработки, анализа и применения. Big Data называют все, что связано с данными и математическими алгоритмами их обработки. В то время как классическое понимание термина предполагает наличие определенного объема информации - речь идет об огромных массивах данных, которых у большинства промышленных предприятий сегодня просто нет. Big Data появятся, когда будут развернуты проекты в области промышленного интернета вещей, когда будут установлены тысячи датчиков, с которых будут собираться терабайты информации. Другими проблемами являются дефицит специалистов с должным уровнем образования, которое дает знания не на уровне теории, а на уровне практики. От качества работы сотрудников, занимающихся глубинной и предикативной аналитикой, напрямую зависит скорость возврата инвестиций в Big Data. Также распространение новых технологий, в том числе Big Data, в Казахстане блокируют следующие факторы: ограниченные бюджеты и обеспокоенность по поводу информационной безопасности. Так, например, по данным Accenture, более 50% опрошенных компаний при внедрении Big Data столкнулись с вопросами информационной безопасности, около 40% – с отсутствием подготовленных кадров, 35% – со сложностями при интеграции с существующими системами. Огромный потенциал уже существующих в организации данных часто не может быть эффективно использован самими специалистами из-за устаревших бизнес-процессов или внутренних регламентов.

Заключение. Несмотря на определенные сложности и проблемы развития, технологии Big Data становятся одним из важнейших направлений формирования новых сервисов, повышения конкурентоспособности сервисных предприятий, создания инновационных маркетинговых инструментов продвижения услуг в экономике Казахстана.

Список литературы

- [1.] Тиндал С. Большие данные: все, что вам необходимо знать. PCWeek/RE, 2012, № 25 (810). URL: <http://www.pcweek.ru/idea/article/detail.php?ID=141962>.
- [2.] Черняк Л. Большие данные - новая теория и практика // Открытые системы. СУБД. - М.: Открытые системы, 2011.- №10.- С. 18-25
- [3.] Инженерный журнал: наука и инновации, 2014, вып. 8. URL: <http://engjournal.ru/catalog/it/asu/1228.html>.
- [4.] https://bigdatawg.nist.gov/pdf/MGI_big_data_full_report.pdf
- [5.] <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>.
- [6.] Обзор: Бизнес-аналитика и большие данные в России.- 2014. - [Электронный ресурс]: complex.imexp.ru
- [7.] Технологии «больших данных» в важнейших отраслях. CITO Research при поддержке QlikView, 2014 г.

[8.] Жукабаева Т.К., Кусаинова А.Т /Технология больших данных (BIG DATA).основные характеристики и перспективы применения. Вычислительная техника. – Астана, 2016 г. – №1. – С.115-132.

PROBLEMS AND PROSPECTS FOR USING BIG DATA IN THE INDUSTRY OF KAZAKHSTAN

A.S.AMIROVA

Postgraduate student of the Eurasian National University named after L.N. Gumilyov

A.T.TOKHMETOV,

Doctor of Physical and Mathematical Sciences Associate Professor of the Information Systems Department of the Eurasian National University named after L.N. Gumilyov, candidate of physical and mathematical sciences

Eurasian National University named after L.N. Gumilyov, Republic of Kazakhstan

E-mail: akzhibek.amirova@gmail.com

Abstract. The article provides an overview of the Big Data technologies used at the enterprises of Kazakhstan, discusses the problems and prospects of further development and use of big data processing systems in the industry of Kazakhstan. The analysis of the problems of introducing Big Data technologies in the industrial sphere is carried out.

Keywords: big data, data analysis, industrial Internet, digitalization of industry.