

BIG DATA В ОБЛАСТИ ЭКОЛОГИИ

Стрельцов А.Д., Войтехович Д.К.

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

Научный руководитель: Мигалевич С.А. – магистр техн. наук, старший преподаватель, начальник центра информатизации и инновационных разработок.

Аннотация. В данной научной работе рассматриваются подходы к проектированию систем обработки больших данных для мониторинга окружающей среды, описывая примеры успешной реализации таких систем. С увеличением объема экологических данных, собранных с помощью датчиков, спутников и других источников, интеграция технологий обработки больших данных в области экологии становится ключевой задачей. Данные об экологическом состоянии, такие как уровень загрязнения воздуха, температуры и влажности, обладают высоким объемом и низкой плотностью значений, что создает сложности для анализа и принятия решений.

Ключевые слова: большие данные, экология, мониторинг окружающей среды, управление данными, автоматизация

Введение. Технический прогресс в настоящее время напрямую зависит от данных, а именно: от особенностей их оборота, правовых режимов в отношении различных видов данных. Государства, а также компании, которые раньше остальных поняли ценность данных и использовали это в процессе своей деятельности, стали доминирующими на соответствующих рынках и даже в мировых масштабах. Субъекты публичного сектора также используют сквозные технологии на основе данных, в том числе и Big Data. Ограничения, вызванные пандемией коронавирусной инфекции и активное внедрение дистанционных рабочих мест лишь ускорили процесс цифровой трансформации. Но еще большая потребность в количестве данных провоцирует новые вызовы, которые не могут остаться без ответа [1].

Применение технологий Big Data в борьбе с загрязнением. Алгоритмы машинного обучения анализируют исторические данные о загрязнении воздуха и воды, собранные с различных источников, включая спутниковые изображения и наземные сенсоры. Эти алгоритмы могут обнаруживать сложные взаимосвязи между различными факторами, такими как промышленные выбросы, движение транспорта, погодные условия и уровни загрязнения.

Данные алгоритмы анализируются в реальном времени, что позволяет оперативно реагировать на ухудшение экологической ситуации. Например, система мониторинга качества воздуха AirVisual [2] использует данные с тысяч станций по всему миру для составления актуальных карт загрязнения воздуха. Использование данных алгоритмов помогают прогнозировать будущие изменения качества воздуха и воды в различных регионах, позволяя властям и организациям принимать активные меры по предотвращению или снижению вредного воздействия на окружающую среду.

Анализ Больших Данных находит применение и в области управления отходами. Современные технологии позволяют отслеживать потоки мусора, определять наиболее загруженные пункты сбора и переработки, а также анализировать эффективность систем рециклинга. Инновационные стартапы, такие как RecycleTrack Systems, используют данные для оптимизации маршрутов сбора отходов и повышения доли переработки, что способствует сокращению объемов мусора на свалках и улучшению экологической обстановки в целом (рисунок 1) [3].

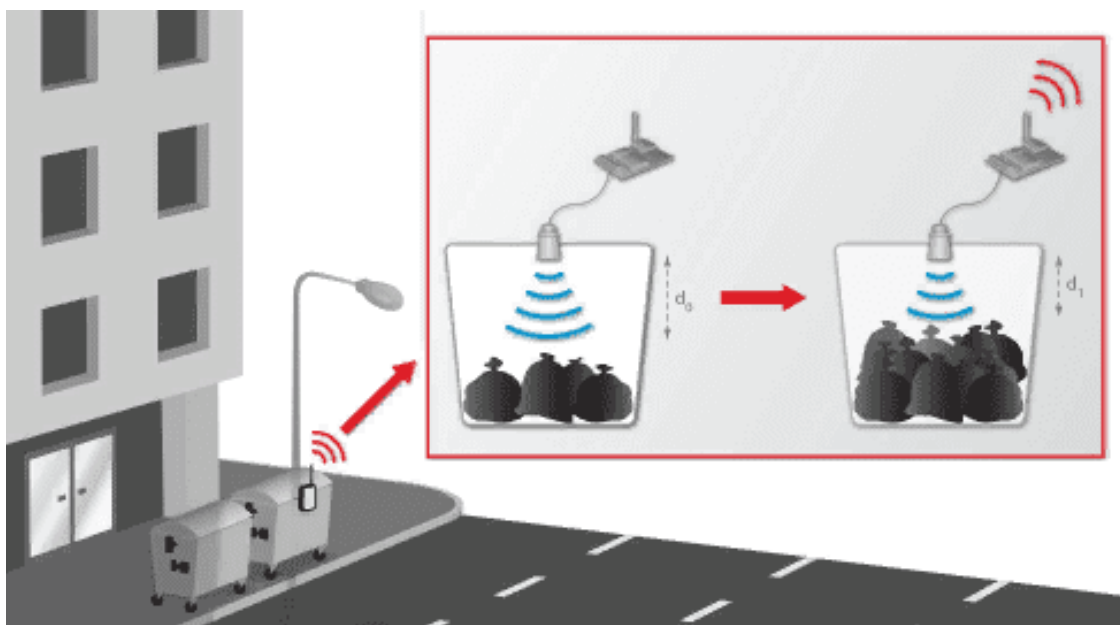


Рисунок 1 – Умный мусорный бак с IoT-датчиком

Современные мусорные контейнеры могут не только сообщать о наполнении, но и сами разбираться с отходами. Например, в Нью-Йорке, Женеве, Дублине, а также некоторых других городах Европы и США, установлены высокотехнологичные урны, которые самостоятельно прессуют мусор и упаковывают его. Примерная стоимость одной урны около \$4 000. Благодаря этим устройствам Филадельфии удалось сократить количество мусоросборочных рейсов почти в 8 раз: с 17 до 2 [4].

Предсказание экологических катастроф. Алгоритмы машинного обучения способны предсказывать экологические катастрофы, такие как лесные пожары, наводнения и засухи, на основе анализа метеорологических данных и изменений в окружающей среде. Это позволяет своевременно предпринимать меры для минимизации потенциального ущерба для населения и экосистем.

В целом, применение алгоритмов машинного обучения к Большим Данным в экологии не только улучшает понимание текущего состояния окружающей среды, но и открывает новые возможности для предотвращения будущих экологических проблем. Эти технологии обеспечивают более эффективное и целенаправленное использование ресурсов для защиты планеты.

Примеры успешных реализаций. Одним из успешных примеров интеграции больших данных в экологию является проект по мониторингу качества воздуха в крупных городах. Система собирает данные с тысяч датчиков, анализирует их и предоставляет информацию о состоянии воздуха в реальном времени. Это позволяет местным органам власти быстро реагировать на изменения и принимать меры по улучшению экологической ситуации. Насколько актуальным и востребованным является применение технологии обработки больших данных в области экологии, можно наблюдать на примере ряда зарубежных стран, в которых Big Data активно используется и находится на достаточно высоком уровне развития [5] (таблица 1).

Таблица 1 – Применение технологии обработки больших данных в зарубежных странах

Критерий	США	Канада	Великобритания	Австралия	Нидерланды	Германия
В каком году начали использовать	Начало 2000-х	2010	2005	2010	2008	2010
Финансирование в начале внедрения	\$100 млн	\$50 млн	£30 млн	AUD 40 млн	€25 млн	€30 млн
Финансирование на момент 2022 года	\$150 млн	\$75 млн	£50 млн	AUD 60 млн	€40 млн	€50 млн
Снижение уровня загрязнения после внедрения	30%	20%	25%	15%	20%	30%

Недостатки использования Big Data. Отсутствие стандартов в области больших данных в экологии представляет собой значительную проблему, затрудняющую сбор, обработку и анализ данных. Одним из основных аспектов является применение различных методов сбора данных. Исследователи могут использовать спутниковые снимки, полевые исследования и сенсоры, что приводит к несоответствиям в качестве и типах данных. Отсутствие единого протокола для сбора данных также может затруднить точность исследований.

Ещё одной серьёзной проблемой является нехватка метаданных, которые содержат информацию о контексте, условиях сбора и обработки данных. Часто метаданные не предоставляются или имеют неполную информацию, что может привести к неправильному пониманию данных. Если исследователь не понимает, как и когда данные были собраны, это может привести к неправильным выводам.

Проблемы совместимости программного обеспечения также играют важную роль. Различные инструменты и платформы для анализа данных могут не поддерживать одни и те же форматы, что создает дополнительные проблемы для совместного использования данных. Это затрудняет совместные исследования и обмен информацией между организациями и учеными.

Неоднородность в анализе и интерпретации данных также является причиной отсутствия стандартов. Это может привести к тому, что разные исследователи будут использовать разные методы анализа, что усложняет сравнение результатов и может привести к разным выводам из одних и тех же данных.

Для решения этих проблем необходимо разработать и внедрить стандарты для сбора, хранения и анализа данных, что поможет улучшить качество и совместимость данных. Создание платформ для обмена данными и метаданными между исследователями может способствовать более широкому использованию и проверке данных. Обучение исследователей стандартам и методологиям повысит общую квалификацию в области работы с большими данными. Таким образом, активное участие научного сообщества в разработке единой системы стандартов является ключевым для более эффективного использования данных и улучшения качества экологических исследований.

Уже существует успешный пример использования стандартов. Компания IBM использует стандарт ISO 14001[6] для формирования общих методов и решений, постоянного совершенствования и согласованности во всем мире. Результатом этой работы является более эффективная и действенная система экологического менеджмента (EMS). Принятие стандарта ISO 14001 помогло IBM поддерживать экологическое лидерство,

превращаясь из производителя вертикально интегрированных систем в поставщика высоко значимых инноваций и услуг со стратегическими задачами, включающими в себя облачные вычисления, аналитические, мобильные, социальные технологии и технологии безопасности. С учетом того, что бизнес компании IBM менялся с годами, менялась и общая EMS, так как она постоянно выявляет и управляет процессом экологического взаимодействия с привлечением новых возможностей для бизнеса. Таким образом, в период с 1990 г. по 2014 г. Внедрение ISO 14001 привели к сбережению 6,8 млн. МВт-ч расходов электроэнергии и связанных с ней выбросом 4,2 млн. тонн CO₂. В течение 2014 года IBM достигла результата в энергосбережении к показателю 6,7 % от полного использования [7].

Заключение. Большие Данные открывают перед нами новые горизонты в понимании и решении экологических проблем, став передовым инструментом в борьбе за сохранение окружающей среды. Их потенциал в этом направлении кажется безграничным, начиная от мониторинга и анализа состояния экосистем до предсказания и предотвращения экологических катастроф.

Проекты, основанные на анализе Больших Данных, уже демонстрируют свою эффективность в улучшении качества воздуха в городах, оптимизации систем управления отходами и повышении эффективности использования природных ресурсов.

Список литературы

1. Технологии больших данных (big data) для охраны окружающей среды [Электронный ресурс]. Режим доступа: <https://lawinfo.ru/articles/1966/tehnologii-bolsix-dannyx-big-data-dlya-okhrany-okruzhayushhei-sredy>. Дата доступа: 20.03.2025.
2. Качество воздуха вокруг вас [Электронный ресурс]. Режим доступа: <https://www.iqair.com/ru/>. Дата доступа: 20.03.2025.
3. Большие Данные (Big Data) в борьбе с загрязнением [Электронный ресурс]. Режим доступа: <https://trendmind.space/techtrends/big-data-v-ekologii/>. Дата доступа: 20.03.2025.
4. Эко-Big Data в большом городе: как технологии делают мегаполис чище [Электронный ресурс]. Режим доступа: <https://bigdataschool.ru/blog/iot-big-data-ml-gorod-ekologiya.html>. Дата доступа: 20.03.2025.
5. Использование Big Data в международном бизнесе [Электронный ресурс]. Режим доступа: <https://www.mathnet.ru/links/0a804ff05d31e5b56c46481451dbff06/tisp521.pdf>. Дата доступа: 20.03.2025.
6. Сертификат системы экологического менеджмента ISO 14001 [Электронный ресурс]. Режим доступа: <https://standartno.by/services/sistema-ekologicheskogo-menedzhmenta-iso-14000/>. Дата доступа: 20.03.2025.
7. Компания IBM использует стандарт ISO 14001 для ведения стабильного бизнеса [Электронный ресурс]. Режим доступа: <https://www.iso.org/ru/news/2015/11/Ref2015.html>. Дата доступа: 20.03.2025.

UDC 621.3.049.77–048.24:537.2

BIG DATA IN THE SPHERES OF ECOLOGY

Streltsov A.D., Voytehovich D.K.

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

Migalevich S.A. – Master of Sci., Senior Lecturer, Head of the Center for Informatization and Innovative Developments

Annotation. This scientific paper examines approaches to the design of big data processing systems for environmental monitoring, describing examples of successful implementation of such systems. With the increasing volume of environmental data collected by sensors, satellites and other sources, the integration of big data processing technologies in the field of ecology is becoming a key task. Environmental data, such as air pollution, temperature, and humidity, have a high volume and low density of values, which creates difficulties for analysis and decision-making.

Keywords: big data, ecology, environmental monitoring, data management, automation