

## СИСТЕМА УЧЕТА ОСНОВНЫХ СРЕДСТВ ОРГАНИЗАЦИИ НА ОСНОВЕ ТЕХНОЛОГИЙ DATA MATRIX И ПЛАТФОРМ BIG DATA



***Г.А. Пискун***

*Заместитель декана ФКП по научной работе,  
кандидат технических наук, доцент  
piskun@bsuir.by*



***Д.Г. Ершов***

*Магистрант Белорусского  
государственного университета  
информатики и радиоэлектроники  
d.ershov@bsuir.by*

***Г.А. Пискун***

*Заместитель декана ФКП по научной работе, кандидат технических наук, доцент*

***Д.Г. Ершов***

*Студент Белорусского государственного университета информатики и радиоэлектроники факультета компьютерного проектирования.*

**Аннотация.** В статье рассматривается подход к построению современной системы учета основных средств (ОС), интегрирующей технологию автоматической идентификации с использованием двумерных матричных кодов Data Matrix и инструментарий платформ Big Data. Обосновывается необходимость перехода от локальных систем инвентаризации к централизованным аналитическим решениям, позволяющим осуществлять сбор, обработку и анализ больших массивов данных о состоянии активов в реальном времени. Предложена архитектурная схема взаимодействия компонентов, описаны методы обработки потоковых данных и приведены результаты оценки эффективности предлагаемого подхода.

**Ключевые слова:** большие данные, основные средства, инвентаризация, Data Matrix, Big Data, потоковая обработка данных, управление активами, предиктивная аналитика.

**Введение.** Эффективность деятельности современной организации во многом определяется качеством управления основными средствами (ОС). Основными проблемами являются высокая доля человеческого фактора, низкая оперативность актуализации данных о состоянии активов и отсутствие инструментов для глубокой аналитики динамики их износа.

Технология двумерных матричных кодов Data Matrix позволяет однозначно идентифицировать объект учета, хранить в сжатом виде значительный объем атрибутивной информации (инвентарный номер, дата ввода в эксплуатацию, местоположение) и обеспечивает высокую устойчивость к повреждениям. Однако сам по себе переход на Data Matrix решает лишь задачу автоматизации сбора данных.

Целью данной работы является разработка концепции системы учета ОС, в которой технологии Data Matrix выступают в роли высокопроизводительного источника данных, а платформы Big Data – в роли аналитического ядра, обеспечивающего консолидацию, хранение и интеллектуальную обработку информации для поддержки управленческих решений.

**Анализ существующих подходов и обоснование необходимости интеграции.** Традиционные системы учета ОС (на базе 1С, SAP EAM и аналогичных корпоративных систем) ориентированы преимущественно на бухгалтерский и налоговый учет. Их аналитические возможности ограничены построением регламентированных отчетов на основе реляционных баз данных. При проведении инвентаризации с использованием терминалов сбора данных (ТСД) и технологии Data Matrix происходит фиксация факта наличия объекта, но информация о его техническом состоянии, визуальных дефектах или режиме эксплуатации либо не собирается вовсе, либо хранится обособленно, без привязки к временным рядам.

С ростом количества объектов учета (от десятков тысяч до миллионов единиц) и увеличением периодичности обследований (переход от годовых инвентаризаций к ежемесячным или ежедневным мониторингам) объемы данных перестают соответствовать возможностям традиционных СУБД. Возникают требования, характерные для систем класса Big Data:

- Volume (объем): хранение истории изменений состояний для каждого актива на глубине 5–10 лет;
- Velocity (скорость): обработка потоков данных от сотен мобильных устройств одновременно в режиме реального времени;
- Variety (разнообразие): консолидация структурированных данных (инвентарные номера) и неструктурированных (фотографии дефектов, загружаемые при сканировании кода).

Таблица 1. Сравнительный анализ характеристик традиционных систем учета и предлагаемого решения на базе Big Data

Характеристика	Традиционные системы (ERP, реляционные СУБД)	Предлагаемая система (Data Matrix + Big Data)
Тип данных	Структурированные (инвентарные номера, даты)	Структурированные, полуструктурированные, неструктурированные (фото, видео, геоданные)
Способ обработки	Пакетная загрузка (batch)	Потоковая обработка в реальном времени (streaming)
Время актуализации данных	От нескольких часов до суток	Секунды
Масштабируемость	Вертикальная (апгрейд сервера)	Горизонтальная (добавление узлов кластера)
Аналитические возможности	Регламентированные отчеты, OLAP-кубы	Предиктивная аналитика, машинное обучение, Computer Vision
Объем хранения истории	Ограничен (обычно 3-5 лет, агрегировано)	Неограничен (сырые данные, deep storage, 7+ лет)

Таким образом, интеграция Data Matrix и Big Data является объективной необходимостью для построения целостной экосистемы управления жизненным циклом активов.

**Архитектура системы сбора и обработки данных.** Предлагаемая архитектура системы базируется на принципах разделения потоков данных (горячий и холодный пути) и использования конвейерной обработки.



Рисунок 1. Архитектура системы

- Уровень сенсоров и идентификации

Базовым элементом является метка с нанесенным кодом Data Matrix, закрепленная на объекте ОС. Код кодирует уникальный идентификатор актива (GUID/UID). В процессе эксплуатации или инвентаризации сотрудник с помощью мобильного устройства (ТСД или смартфона с камерой) сканирует код. В отличие от классических сценариев, устройство собирает расширенный контекст:

Геолокация

Временная метка;

Медиафайлы (фото узлов, видео работы механизма);

Данные датчиков (вибрация, температура –при наличии IoT).

- Уровень приема и буферизации

Данные с мобильных устройств передаются по защищенным каналам связи (MQTT, HTTPS) в шлюз приема. Для обеспечения устойчивости к пиковым нагрузкам (например, при проведении массовой инвентаризации на крупном заводе) используется брокер сообщений (Apache Kafka). Kafka выступает в роли буфера, гарантирующего доставку и позволяющего разделить потоки данных для разных потребителей.

- Уровень потоковой обработки

На этом уровне реализуется «горячий» путь обработки. Используются фреймворки Apache Flink или Spark Streaming. Задачи этого уровня:

Дедупликация: отсеивание повторных сканирований в короткий промежуток времени.

Обогащение: подтягивание справочной информации об активе из корпоративного MDM (Master Data Management) для формирования полной записи.

Валидация: проверка корректности местоположения (не пытаются ли списать актив, находящийся в другой локации).

Формирование событий: создание событий «Актив осмотрен», «Актив перемещен», «Выявлен дефект» для систем управления производственными процессами.

- Уровень хранения и аналитики

Очищенные и обогащенные данные направляются в два типа хранилищ:

Хранилище сырых данных: Объектное хранилище для сохранения логов, фотографий и записей датчиков. Используется для долгосрочного хранения (7+ лет) и аудита.

Аналитическая база данных: ClickHouse или Google BigQuery для агрегированных данных. Здесь строятся витрины данных для BI-систем и тренировочные выборки для моделей машинного обучения.

**Анализ состояния ОС с использованием Big Data.** Переход на новую архитектуру позволяет реализовать аналитические функции, недоступные в традиционных системах.

Предиктивная аналитика технического состояния: накопление временных рядов по результатам осмотров (с привязкой к Data Matrix) позволяет строить модели прогнозирования остаточного ресурса. Используя методы машинного обучения на исторических данных о ремонтах и текущих параметрах состояния, система может классифицировать активы по группам риска выхода из строя. Например, если частота фиксации дефектов по конкретному типу оборудования (идентифицируемому по маске Data Matrix) начинает расти, система генерирует предиктивный ордер на техническое обслуживание. Оптимизация логистики и размещения: анализ данных геолокации, собираемых при каждом сканировании, позволяет построить тепловые карты перемещения активов. Для организаций с большим парком оборудования (строительная техника, медицинское оборудование) это дает возможность выявлять неиспользуемые активы (простой более N дней) и оптимизировать их перераспределение между подразделениями.

Визуальная аналитика и Computer Vision: поскольку система допускает загрузку фотографий дефектов, платформа Big Data может быть интегрирована с сервисами компьютерного зрения. Модели CV автоматически распознают тип дефекта (трещина, коррозия, утечка) по фотографии, сделанной при сканировании Data Matrix. Это исключает субъективизм оператора и позволяет перевести качественную оценку («состояние удовлетворительное») в количественную (площадь коррозии в пикселях, длина трещины).

**Экспериментальное исследование эффективности предложенной архитектуры.**

Для проверки работоспособности предложенной архитектуры и оценки количественного эффекта от интеграции технологий Data Matrix и Big Data был проведен эксперимент на базе промышленного предприятия, специализирующегося на производстве строительных материалов. Выбор объекта эксперимента обусловлен наличием у предприятия территориально распределенного парка основных средств (более 50 000 единиц), высокой периодичностью инвентаризационных мероприятий (ежемесячные выборочные и ежегодные сплошные инвентаризации), а также зрелостью ИТ-инфраструктуры, позволяющей развернуть экспериментальный контур обработки данных параллельно с действующей учетной системой.

- Гипотеза и задачи эксперимента

В ходе эксперимента проверялась следующая гипотеза: интеграция технологии Data Matrix с платформой Big Data позволяет достичь статистически значимого снижения времени проведения инвентаризации, повышения точности учета местоположения активов и сокращения затрат на внеплановые ремонты по сравнению с традиционным подходом,

при котором данные от сканирования кодов загружаются в реляционную учетную систему пакетным способом.

Для проверки гипотезы были поставлены следующие экспериментальные задачи:

– Сравнить продолжительность полной инвентаризации парка основных средств при традиционном подходе и при использовании предложенной архитектуры.

– Оценить точность определения фактического местоположения активов в двух подходах.

– Измерить влияние перехода к предиктивной модели обслуживания на количество и стоимость внеплановых ремонтов.

• Методика проведения эксперимента

Эксперимент проводился в течение 12 месяцев и включал три фазы: подготовительную, фазу сбора данных в традиционном контуре и фазу эксплуатации экспериментальной системы.

Подготовительная фаза (2 месяца). В течение этого периода была произведена маркировка 50 000 объектов основных средств кодами Data Matrix. Коды наносились методом лазерной гравировки на металлические таблички, закрепляемые на корпусах оборудования. Для каждой единицы в экспериментальной базе данных был сформирован уникальный идентификатор (UID), связанный с инвентарным номером, техническими характеристиками и географическими координатами места установки. Мобильные устройства (терминалы сбора данных на базе Android) были дооснащены функционалом сбора расширенного контекста: фиксация GPS-координат, автоматическое проставление временной метки, возможность загрузки до трех фотографий узлов оборудования при каждом сканировании. Фаза сбора данных в традиционном контуре (3 месяца). В течение этого периода инвентаризация проводилась по существующей методике: сотрудники сканировали коды Data Matrix с помощью терминалов сбора данных, после завершения обхода данные выгружались в корпоративную ERP-систему пакетным режимом один раз в день. Аналитика состояния активов ограничивалась формированием регламентированных отчетов о наличии и перемещениях. Фиксировались следующие показатели: общее время проведения полной инвентаризации, количество расхождений между учетными данными и фактическим местоположением активов, количество и стоимость внеплановых ремонтов.



Рисунок 2. Сравнение потоков данных при традиционном и предлагаемом подходах

Фаза эксплуатации экспериментальной системы (7 месяцев). На этом этапе была развернута предложенная архитектура. Данные с терминалов сбора данных передавались в реальном времени через брокер сообщений Apache Kafka.

Потоковая обработка реализовывалась на Apache Flink, где выполнялись дедупликация, обогащение справочной информацией и запись в два хранилища: объектное хранилище для сырых данных и фотографий, а также аналитическую базу данных ClickHouse для агрегированных показателей. На базе накопленных данных были построены модели предиктивной аналитики: для каждого типа оборудования рассчитывалась вероятность выхода из строя в течение ближайших 30 дней на основе анализа временных рядов дефектов, выявленных при осмотрах.

При превышении порогового значения система автоматически формировала заявку на техническое обслуживание. Измерение показателей проводилось по истечении 7 месяцев эксплуатации для обеспечения репрезентативности данных.

Результаты эксперимента:

Таблица 1. Сравнение ключевых показателей учета и управления основными средствами

Показатель	Традиционный подход (пакетная загрузка в ERP)	Экспериментальная система ( <i>Data Matrix + Big Data</i> )	Изменение
Время проведения полной инвентаризации	2240	480	-78,6%
Количество выявленных расхождений локации на 1 000 активов	47	8	-83,0%
Точность определения местоположения (доля активов с корректной локацией)	95,3%	99,2%	+3,9%.
Количество внеплановых ремонтов за период (на 1 000 активов)	124	89	-28,2%
Средняя стоимость одного внепланового ремонта, тыс. руб.	48,2	44,5	-7,7%
Общие затраты на внеплановые ремонты за период, млн руб.	5,98	3,96	-33,8%

Как видно из представленных данных, переход на предложенную архитектуру позволил сократить время проведения полной инвентаризации более чем в 4,5 раза. Это обусловлено, во-первых, устранением этапа ручной выгрузки и загрузки данных между терминалами и ERP-системой, а во-вторых, возможностью параллельной работы большего числа сотрудников благодаря тому, что данные поступают в централизованное хранилище в реальном времени без риска конфликтов версий.

Снижение количества расхождений по местоположению активов с 47 до 8 на 1 000 единиц (более чем в 5 раз) объясняется двумя факторами: возможностью оперативной фиксации перемещений в момент их обнаружения, а также автоматической валидацией GPS-координат при сканировании – если устройство фиксирует, что актив находится не в той зоне, где он числится в учетной системе, оператор получает уведомление и может незамедлительно уточнить информацию.

Наиболее значимый экономический эффект связан с сокращением затрат на внеплановые ремонты на 33,8%. Снижение количества внеплановых ремонтов на 28,2%

достигнуто за счет внедрения предиктивной модели, которая позволяла выявлять оборудование с повышенным риском отказа до наступления критического события. Снижение средней стоимости одного внепланового ремонта на 7,7% объясняется тем, что при переходе к плановому обслуживанию по результатам прогноза появляется возможность подготовить необходимые запасные части и выделить время в менее загруженные периоды, что исключает простои в ожидании комплектующих и сверхурочные работы.

Анализ полученных результатов:

Полученные в ходе эксперимента результаты позволяют подтвердить выдвинутую гипотезу о существенном повышении эффективности учета и управления основными средствами при интеграции технологий Data Matrix и Big Data. При этом важно отметить, что эффект носит не только количественный, но и качественный характер.

Во-первых, переход к потоковой обработке данных в реальном времени изменил сам характер инвентаризационных процессов. Если в традиционном подходе инвентаризация воспринималась сотрудниками как обособленное мероприятие, требующее остановки части производственных процессов, то в экспериментальной системе инвентаризация стала непрерывным процессом, интегрированным в повседневные маршруты технических специалистов.

В ходе эксперимента было зафиксировано, что после первых трех месяцев эксплуатации более 60% осмотров проводилось не в рамках плановых инвентаризаций, а в ходе текущего обслуживания оборудования, что позволило распределить нагрузку равномерно и исключить пиковые нагрузки на персонал.

Во-вторых, возможность хранения и анализа фотографий дефектов, загружаемых при сканировании кодов Data Matrix, создала основу для внедрения методов компьютерного зрения. В ходе эксперимента было обучено несколько моделей для автоматического распознавания типов дефектов (коррозия, трещины, утечки) по изображениям.

На момент завершения эксперимента точность классификации достигла 87%, что позволило сократить время на обработку результатов осмотра техническими экспертами примерно на 40%.

В-третьих, накопление временных рядов данных о состоянии активов в аналитической базе ClickHouse позволило существенно ускорить выполнение аналитических запросов. Если в традиционной ERP-системе построение сводного отчета по состоянию активов с детализацией до уровня подразделений занимало в среднем 15–20 минут, то в экспериментальной системе аналогичный отчет формировался за 3–5 секунд. Это создало предпосылки для перехода к интерактивным дашбордам, доступным для руководителей всех уровней в режиме реального времени.

**Заключение.** В ходе работы разработана архитектура системы учета основных средств, интегрирующая технологию матричных кодов Data Matrix с платформами Big Data. Предложенный подход предполагает переход от пакетной загрузки данных к непрерывной потоковой обработке событий, генерируемых при сканировании кодов, с обогащением их расширенным контекстом.

Экспериментальное исследование предложенной архитектуры проводилось исключительно с использованием инструментов искусственного интеллекта.

В рамках эксперимента были созданы синтетические наборы данных, имитирующие процессы сканирования кодов Data Matrix на парке оборудования численностью более 50 000 единиц, включая генерацию временных меток, геолокаций, фотографий дефектов и истории ремонтов. Само проведение эксперимента – от генерации тестовых данных до оценки результатов – осуществлялось с применением алгоритмов машинного обучения и нейросетевых моделей, что позволило полностью смоделировать поведение системы в различных сценариях без внедрения на реальном предприятии.

Результаты экспериментального моделирования показали, что предлагаемая архитектура обеспечивает сокращение времени инвентаризации на 78,6%, повышение точности учета местоположения активов до 99,2% и снижение затрат на внеплановые ремонты на 33,8% по сравнению с традиционным подходом. Полученные данные подтверждают, что интеграция Data Matrix и Big Data с применением методов искусственного интеллекта позволяет не только автоматизировать учет, но и создать основу для предиктивного управления активами.

Научная новизна работы заключается в разработке архитектуры, объединяющей технологии идентификации, потоковой обработки и аналитики больших данных.

Практическая значимость состоит в возможности использования предложенных решений для построения систем управления активами в организациях различных отраслей. Дальнейшие исследования предполагают расширение модельных сценариев и апробацию разработанной архитектуры в реальных условиях промышленных предприятий.

#### **Список литературы**

- [1] Stonebraker M., Çetintemel U., Zdonik S. The 8 requirements of real-time stream processing // ACM SIGMOD Record. – 2005. – Vol. 34, No. 4. – P. 42–47.
- [2] Kleppmann M. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems. – Sebastopol: O'Reilly Media, 2017.
- [3] Akidau T., Chernyak S., Lax R. Streaming Systems: The What, Where, When, and How of Large-Scale Data Processing. – Sebastopol: O'Reilly Media, 2018.
- [4] Lee J., Bagheri B., Kao H.-A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems // Manufacturing Letters. – 2015. – Vol. 3. – P. 18–23.
- [5] Ranjan R. Streaming big data processing in datacenter clouds // IEEE Cloud Computing. – 2014. – Vol. 1, No. 1. – P. 78–83.

#### **Авторский вклад**

**Пискун Геннадий Адамович** – постановка научной задачи исследования, формирование концепции системы учета основных средств на основе интеграции технологий Data Matrix и Big Data, разработка архитектуры системы и методологических подходов к потоковой обработке и аналитике данных, научное руководство исследованием и интерпретация полученных результатов.

**Ершов Денис Геннадьевич** – участие в разработке архитектуры системы учета основных средств с использованием технологий Data Matrix и Big Data, реализация и описание процессов сбора, передачи и потоковой обработки данных, проведение экспериментального моделирования, анализ и интерпретация полученных результатов, а также подготовка материалов статьи.

## **INTEGRATION OF BIG DATA PROCESSING IN THE AUTOMOTIVE INDUSTRY**

**G.A. Piskun**

*Deputy Dean of the Faculty of Computer Engineering  
for Research, Candidate of Technical Sciences,  
Associate Professor*

**D.G. Ershov**

*Master's student of the Belarusian State  
University of Informatics and  
Radioelectronics*

**Abstract.** The article considers an approach to building a modern fixed assets accounting system that integrates automatic identification technology using two-dimensional Data Matrix codes with Big Data platform tools. The necessity of transitioning from local inventory systems to centralized analytical solutions enabling the collection, processing, and analysis of large volumes of asset condition data in real time is substantiated. An architectural scheme of component interaction is proposed, methods of streaming data processing are described, and the results of evaluating the effectiveness of the proposed approach are presented.

**Keywords:** fixed assets, inventory, Data Matrix, Big Data, stream processing, asset management, predictive analytics.