

# Особенности разработки проекта информационной системы цифрового двойника буровой установки на основе технологии больших данных, машинного обучения и Интернета Вещей

П.А. Болдырев

**Аннотация.** В статье рассматривается принципиальная схема разработки информационной системы цифрового двойника буровой установки на основе элементов четвертой промышленной революции («Индустрия 4.0»). Исследуются подсистемы с использованием алгоритмов искусственного интеллекта, технологии интернета вещей (IoT–InternetofThings), больших данных (BigData). Анализируется взаимосвязь подсистем с решением задач информационной системы. Излагаются характеристики аппаратных средств, а также способы хранения и представления данных информационной системы. Отдельное внимание уделяется моделированию типов информационных запросов и способов передачи данных. Описывается перечень технических и программных средств, с помощью которых есть принципиальная возможность реализации проекта. Приводятся числовые метрики параметров, на основе которых прогнозируется количественно-качественные характеристики проекта информационной системы.

**Ключевые слова:** большие данные, буровая установка, машинное обучение интернет вещей, проект информационной системы, цифровой двойник.

## I. ВВЕДЕНИЕ

Цифровой двойник оборудования промышленного предприятия, в котором используются методы машинного обучения, является одним из ярких представителей так называемой четвертой промышленной революции («Индустрия 4.0»), внедрение, которого позволяет кратно увеличить качество и безопасность производства, ускорить ряд технологических процессов, значительно снизить издержки.

На данный момент имеются противоречия в понимании признаков цифрового двойника, что приводит к существенным разночтениям и даже некорректным формулировкам определения этого понятия. Так же существуют определенные сложности при построении понятийного аппарата, касающегося

искусственного интеллекта и машинного обучения, что предопределяет необходимость его проработки для дальнейшего использования в практических целях. Информационные технологии стремительно меняются, предлагая экономическим субъектам доступ ко все более дешевым, быстродействующим, легким, устойчивым устройствам, удобным протоколам связи, способными работать во все более широких областях человеческой деятельности. Те предприятия, которые уже внедрили технологии четвертой промышленной революции, в полной мере оценили их количественные и качественные преимущества, позволяющие быть конкурентоспособными на рынке, предлагать сотрудникам более комфортные и безопасные условия труда, а также обеспечивать лучшую социальную и экологическую ответственность перед обществом [6]. Исходя из сказанного, к основным задачам, рассматриваемым в данной статье относятся:

- 1) рассмотрение основных понятий, связанных с четвертой промышленной революцией;
- 2) описание принципиальной схема работы цифрового двойника буровой установки;
- 3) анализ аппаратных и программных средств реализации цифрового двойника буровой установки;
- 4) конкретизация параметрических данных цифрового двойника.

## II. ОСНОВНЫЕ ПОНЯТИЯ ЦИФРОВОГО ДВОЙНИКА, БОЛЬШИХ ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ

Для точного и наиболее полного понимания прежде всего необходимо определить понятия «цифровой двойник» и «машинное обучение».

Ретроспективный анализ позволяет уверенно утверждать, что приоритет определения понятия «цифровой двойник» по праву принадлежит Майклу Гривзу [1], который впервые ввел его в публичный оборот в 2002 году на конференции Общества инженеров-технологов и звучит это определение следующим образом: «цифровой двойник представляет

собой набор виртуальных информационных конструкций, которые полностью описывают потенциальный или реальный физически производимый продукт от микроатомного уровня до макрогеометрического уровня» [17].

Заслуживает внимания также определение данного понятия, предложенное аэрокосмическим инженером Карен Уиллкокс: «цифровой двойник — это набор связанных вычислительных моделей, которые со временем эволюционируют, чтобы постоянно представлять структуру, поведение и контекст уникального физического актива» [18].

По определению крупнейшего сервиса облачных вычислений Amazon Web Services, «цифровой двойник – это виртуальная модель физического объекта. Он охватывает жизненный цикл объекта и использует данные в реальном времени, отправленные с датчиков объекта, для моделирования поведения и мониторинга операций» [9]. В то же время Высшая школа инжиниринга московского инженерно-физического института дает следующее определение: «Цифровой двойник – это неразрывно связанные между собой виртуальные представления физического объекта (продукта или процесса) в каждый момент времени. Цифровые двойники используются на протяжении всего жизненного цикла объекта для моделирования, прогнозирования и оптимизации продукта и системы производства» [11].

Таким образом, главными элементами цифровых двойников являются:

- реально существующий физический объект;
- виртуальная модель вышеуказанного реально существующего объекта;
- параметрическая информация, считываемая с реально существующего объекта и обновляющая в реальном времени виртуальную модель;
- оборудование для считывания параметрической информации;
- программно-аппаратный комплекс для обработки параметрической информации с предиктивным моделированием.

Именно наличие цифровых информационных потоков делает возможным практическое применения цифровых двойников для решения актуальных задач при разработке новых продуктов, обслуживании уже существующих. И чем больше параметров считывается, тем более комплексные и сложные задачи можно решать. В связи с этим видится необходимым рассмотрение понятия машинного обучения, которое в большинстве случаев включается в разработку цифровых двойников [4].

Согласно определению специалистов Microsoft, машинное обучение – это использование математических моделей данных, которые помогают компьютеру обучаться без непосредственных

инструкций [9]. Другой лидер в области внедрения цифровых решений Oracle дает следующее определение: «Машинное обучение – это направление искусственного интеллекта, сосредоточенное на создании систем, которые обучаются и развиваются на основе получаемых ими данных» [8].

Следовательно, в основе машинного обучения находятся технологии, способные особым образом обрабатывать получаемые массивы данных для дальнейшего обучения систем искусственного интеллекта. Конечным результатом такого обучения является польза, извлекаемая в процессе применения обученного искусственного интеллекта на конкретном объекте и решения конкретных проблем и задач.

В настоящий момент существует множество подходов к определению больших данных. С одной стороны, это набор технологий, инструментов, методов и подходов, предназначенных для решения проблем обработки больших объемов данных, с другой стороны [14], под большими данными часто понимают просто объем данных (в том числе структурированные данные, медиа и случайные объекты), который невозможно обработать обычными (стандартными) высокопроизводительными методами [16]. Однако количественная характеристика уже не является решающей на сегодняшний момент, ключевой аспект связан с возможностью извлекать из этих данных новую ценность – новые данные, которые можно применить [7].

### III. СОДЕРЖАНИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ (ИС) ЦИФРОВОГО ДВОЙНИКА (ЦД) БУРОВОЙ УСТАНОВКИ

Стоит сразу отметить, что технология ЦД в горнодобывающей промышленности может иметь множество применений и спецификаций [3]. В то же время ЦД для буровых установок нефтегазовой сферы имеет свои особенности [2].

Для того, чтобы ЦД был полноценным, необходимо выделить 5 базовых подсистем, обязательно присутствующих в ИС:

1) Физическое оборудование – буровая установка, а также передающие и принимающие датчики интернета вещей. Их, в свою очередь, тоже можно декомпозировать на отдельные составляющие, однако в рамках текущего этапа анализа это избыточно.

2) Центр обработки данных, включающий базу данных, в том числе с применением облачных вычислений [20].

3) Искусственный интеллект – алгоритмы машинного обучения [16].

4) Клиентское программное обеспечение (ПО) – программная среда, оснащенная пользовательским интерфейсом, позволяющая принимать всю необходимую информацию о системе и управлять физическим оборудованием.

5) Информационный поток, циркулирующий между вышеперечисленными компонентами системы и обеспечивающий их аналитическую и управленческую взаимосвязь.

Данную систему можно декомпозировать на множество подсистем и элементарных составных частей, однако ограничимся наиболее значимыми из них. Для больше наглядности представим не просто систему, но принципиальную схему ее работы, продемонстрированную на рис. 1.



Рис. 1. Принципиальная схема работы ИС ЦД буровой установки на основе машинного обучения

Наглядное представление узлов и их взаимное расположение изображено на рис.1.

Узлами изучаемой системы являются:

- пользователи (операторы, системные администраторы, администраторы СУБД и др.);
- физическое оборудование с подсистемой датчиков;
- центр обработки данных;
- база данных;
- искусственный интеллект.

На рис. 1 также наглядно продемонстрировано направление информационных потоков. Их характер и направление можно охарактеризовать следующим образом:

- сырые данные (BigData), представляющие собой необработанные сплошные информационные массивы, исходящие от датчиков физического оборудования к центру обработки данных (ЦОД);
- подготовленные данные – обработанные в ЦОД сырые данные, избавленные от информационных «выбросов» и шумов, с компенсированными пропусками, отправляемые от базы данных ЦОД к алгоритмам машинного обучения (искусственного интеллекта);
- готовое решение, являющее собой суммированный результат всех вычислений искусственного интеллекта, упакованное в виде понятных для пользователя системы показателей по работе физического оборудования как в реальном времени, так и с дополнительными предиктивными параметрами;

– управленческая информация – это та совокупность пользовательских команд, которые формируются на основе готового решения от искусственного интеллекта и направляются от клиентского программного обеспечения к физическому оборудованию для соответствующего изменения параметров его работы с целью ее оптимизации.

С одной стороны, было бы несправедливо не упомянуть важнейшую часть подобных систем – людей (операторов, инженеров и т.д.), однако в данном случае предполагается, во-первых, их априорное участие, а во-вторых, представленная система вполне может работать и в полностью автономном режиме.

Для анализа задач, выполняемых посредством программно-технических средств информационной системы, рассмотрим таблицу 1.

Таблица 1. Взаимосвязь подсистем и решаемых задач ИС

Подсистема	Задача
Физическое оборудование	Анализ работы физического оборудования с помощью датчиков
Центр обработки данных	Подготовка всех данных, поступающих с датчиков
Искусственный интеллект	Обработка и анализ данных с центра обработки данных Отправка готового решения пользователю
Клиентское ПО	Представление готового решения пользователю
Информационный поток	Обеспечение взаимосвязи между всеми компонентами системы

Используя данные таблицы 1 выделим, что к программным средствам, обеспечивающим эффективное функционирование ИС можно отнести следующие:

- клиентское приложение, созданное на базе игрового движка UnrealEngine 4.27, позволяющий управлять трехмерным цифровым двойником посредством графического интерфейса (например, с возможностью подключения устройств виртуальной реальности);
- программный комплекс Microsoft Azure со встроенными сервисами (либо его аналог), с помощью которого осуществляется прием информации с датчиков физического оборудования, а также развертка дополнительных сервисов;
- система управления базами данных (СУБД) на технологиях NoSQL;
- дополнительные программные пакеты для формирования дополнительного контента: 3d-редактор 3dsMax 2022, графический редактор нанесения физически достоверных материалов Substance Painter

7.1, звуковой редактор SamplitudePro X6 suite, редактор векторной графики CorelDraw 2023, онлайн-сервис для отрисовки пользовательского интерфейса Figma;

– также стоит отметить необходимость наличия операционной системы, например, Microsoft Windows 10, и необходимых драйверов для функционирования всех компонентов компьютера, например, видео-, аудиосистем и т.д.

#### IV. ХАРАКТЕРИСТИКИ АППАРАТНЫХ СРЕДСТВ В ИС ЦД БУРОВОЙ УСТАНОВКИ

К аппаратным средствам можно отнести следующие:

– физическое оборудование, составляющее основу цифрового двойника;

– датчики и сенсоры, закрепляемы на физическое оборудование;

– средства коммуникации между всеми устройствами;

– устройства для запуска клиентского приложения, а также управления ЦОД, базой данных и алгоритмов машинного обучения: серверы, персональные компьютеры, ноутбуки, планшеты, смартфоны[5].

Для определения пропускной способности системы будем исходить из следующих целевых параметров проекта:

– количество пользователей системы будет сравнительно небольшим (до 50 человек в пиковых значениях);

– решаются задачи по приему сырых данных из 4 датчиков оборудования со скоростью около 20 операций в секунду с каждого, последующая их обработка и пропуск через нейросеть с формированием системы поддержки принятия решений, а также управление физическим оборудованием через другие датчики;

– формируемый размер данных в день – не более 1 Гб, сложность обработки данных средняя;

– используется стандартное оборудование: обычные офисные ПК, серверы, датчики и другие устройства, доступные к приобретению на широком рынке;

– необходима высокая скорость обработки информационных потоков и обеспечение высокой надежности работы системы в целом.

Таким образом, пропускная способность проектируемой системы будет следующей:

- скорость приема данных: 80 операций в секунду (20 операций в секунду с каждого из 4 датчиков);

- размер формируемых данных в день: порядка 1 Гб, что означает пропускную способность примерно в 12.5 Мбит/с;

- скорость обработки данных и пропуска через нейросеть: зависит от сложности нейросети и объема данных, но должна быть достаточной для обработки 80 операций в секунду.

Разумеется, в процессе тестирования необходимо будет убедиться, что оборудование соответствует требованиям скорости и надежности работы системы, и при необходимости провести дополнительную настройку или модернизацию оборудования.

#### V. МОДЕЛИРОВАНИЕ ТИПОВ ИНФОРМАЦИОННЫХ ЗАПРОСОВ И СПОСОБОВ ПЕРЕДАЧИ ДАННЫХ

Проектом подразумевается наличие следующих типов информационных запросов:

– на получение информации о состоянии системы: это могут быть запросы на получение информации о доступности системы, загрузке процессора, использовании памяти, объеме свободного места на диске, состоянии сетевых соединений и т.д.;

– на доступ к данным: это могут быть запросы на получение, изменение или удаление данных, находящихся в базе данных или других хранилищах данных, связанных с системой;

– на выполнение операций: это могут быть запросы на выполнение различных операций, связанных с управлением системой, таких как установка параметров, запуск или остановка служб, перезагрузка системы и пр.;

– на выполнение вычислений: например, это могут быть запросы на выполнение различных вычислительных задач, связанных с обработкой данных, анализом данных, прогнозированием;

– на получение отчетов: это могут быть запросы на получение различных отчетов, связанных с работой системы, таких как отчеты о загрузке процессора, отчеты о доступности системы, отчеты о работе сетевых соединений и т.д.;

– на получение помощи: это могут быть запросы на получение помощи и поддержки от системных администраторов или технической поддержки, связанные с устранением проблем, возникающих в работе системы.

На данный момент рассматриваются два базовых способа передачи данных:

– кабельная передача данных: это метод передачи данных, при котором информация передается по проводам, подключенным к устройствам. Кабельная передача данных может быть осуществлена с использованием различных типов кабелей, таких как витая пара, коаксиальный кабель, оптоволокно;

– беспроводная передача данных: это метод передачи данных, при котором информация передается по радиоволнам, без использования проводов. Беспроводная передача данных может быть осуществлена с использованием различных технологий, таких как Wi-Fi, Bluetooth, NFC.

К наиболее важным и применимым в проекте методам, и средствам обработки данных относятся следующие:

– интернет вещей (Internet of Things, IoT): это технология, которая позволяет собирать данные с различных устройств и оборудования, подключенных к Интернету. IoT-сенсоры могут быть установлены на оборудовании для сбора данных о его работе и состоянии;

– аналитика данных (Data Analytics): это методы и технологии, используемые для анализа больших объемов данных. Данная технология может быть использована для анализа данных, собранных с IoT-сенсоров, и определения оптимальных параметров работы оборудования;

– машинное обучение (Machine Learning): это методы и технологии, используемые для обучения компьютерных систем на основе данных. Может быть использовано для определения оптимальных параметров работы оборудования и предсказания возможных сбоев;

– базы данных (Databases): это системы управления базами данных, используемые для хранения и организации данных. В случае промышленного оборудования, базы данных могут быть использованы для хранения данных, собранных с IoT-сенсоров, и их последующей обработки.

## VI. ХРАНЕНИЕ И ПРЕДСТАВЛЕНИЕ ДАННЫХ В ИС ЦД БУРОВОЙ УСТАНОВКИ

В проекте предусмотрено два способа хранения данных:

– локальное хранение – это метод хранения данных, при котором данные хранятся на локальных устройствах. Средствами хранения здесь выступают жесткие диски, флеш-накопители и т.д. Локальное хранение обеспечивает быстрый доступ к данным и может быть более безопасным, чем облачное хранение;

– хранение на физических носителях – это метод хранения данных, при котором данные сохраняются на физических носителях. Средства хранения в этом случае – внешние диски, мобильные флеш-накопители и др. Хранение на физических носителях может быть полезным для сохранения критически важных данных, но может быть менее удобным для доступа к данным.

Представление данных является важным аспектом работы ИС. В проекте предусмотрены следующие методы представления данных:

– моделирование производственных процессов с использованием цифровых двойников оборудования. Моделирование производственных процессов позволяет оптимизировать работу оборудования и улучшить производительность производственных процессов;

– репликация оборудования, при котором цифровой двойник создается в виде точной копии реального оборудования. Репликация оборудования позволяет точно моделировать работу оборудования и использовать его для анализа и оптимизации производственных процессов;

– визуализация данных в графическом виде, который позволяет быстро и наглядно анализировать данные. В случае цифровых двойников промышленного оборудования, визуализация данных может быть использована для отображения параметров работы оборудования и его состояния;

– в качестве дополнительного метода рассматривается виртуальная реальность – технология, которая позволяет создавать виртуальное пространство и взаимодействовать с ним в реальном времени. В случае цифровых двойников промышленного оборудования, виртуальная реальность может быть использована для визуализации работы оборудования и его состояния.

К возможным средствам предоставления данных можно отнести:

– Microsoft Power BI, Unreal Engine, Microsoft Azure Machine Learning, Siemens Plant Simulation.

Некоторые из наиболее подходящих для проекта методов и средств ввода-вывода данных на примере включают в себя:

– IoT-сенсоры: это устройства, которые используются для сбора данных с оборудования и окружающей среды, позволяя получить информацию о состоянии оборудования, параметрах его работы, возникновении и влиянии внешних факторов;

– M2M-коммуникации, которые позволяют устройствам обмениваться данными между собой без участия человека и могут быть использованы для передачи данных между различными компонентами оборудования;

– протоколы обмена данными: это стандарты, которые определяют формат и способ передачи данных между устройствами. В случае цифровых двойников, протоколы обмена данными могут быть использованы для передачи данных между IoT-сенсорами и другими компонентами оборудования;

– системы управления производством – это программные системы, которые используются для управления производственными процессами. Они могут быть использованы для передачи данных между различными компонентами оборудования и для управления производственными процессами на основе этих данных.

К возможным средствам ввода-вывода данных можно отнести IoT-сенсоры Microsoft Azure IoT Suite, M2M-коммуникацию MQTT, протоколы обмена данными Modbus, системы управления производством Siemens Simatic IT.

## VII. ПРОТОТИП РАБОТЫ ЦИФРОВОГО ДВОЙНИКА БУРОВОЙ УСТАНОВКИ

В рамках исследовательской работы была реализована программная реализация прототипа ЦД буровой установки и, в частности, наглядно показаны параметрические данные, которые представляют наибольший интерес у бизнес-заказчиков подобных систем: промышленных предприятий нефтегазовой и машиностроительной отраслей. Вкратце укажем, как происходит работа данного прототипа.

Приведем рис. 2, на котором видно, что в результате запуска программы на дисплей выводится окно программы, содержащая стартовую (начальную) сцену, крупным планом демонстрирующая буровую установку с вращающимся буром.

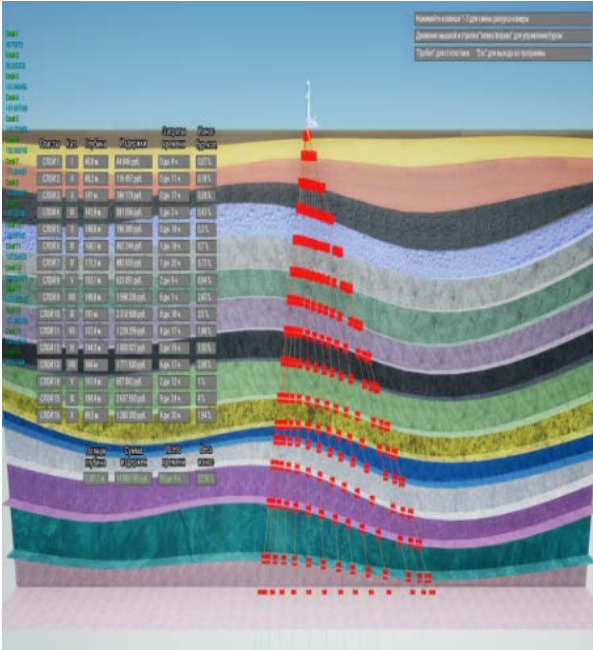


Рис. 2. Скриншот работы прототипа ИС ЦД буровой установки на основе машинного обучения

В правом верхнем углу окна программы, показанного на рис. 2 содержится раздел с описанием управления программой посредством клавиатуры и мышки на стандартном офисном персональном компьютере. Так, клавиши 1,2,3 отвечают за соответствующее позиционирование виртуальных камер в приложении, клавишами «влево» и «вправо» можно перемещать буровую установку в соответствующих направлениях (при этом установлен лимит на расстояние перемещений). Движение мышкой «влево» и «вправо» позволяет управлять направлением виртуального бура буровой установки относительно горизонта. Клавиша «Esc» завершает работу программы, автоматически закрывая ее окно. Клавиша «Пробел» срабатывает только после предварительного нажатия клавиш 1,2,3, позиционирующих виртуальные камеры и отвечает за вызов интерфейса статистики на экран дисплея.

Интерфейс статистики содержит следующую информацию:

1. Пласты – последовательный нумерованный список условных слоев земной коры.

2. Кат. – категории сложности по буримости, от I до X, от самой простой до самой сложной. Каждый слой обладает своей категорией сложности, влияющей на стоимость бурения, затраты времени, износ буровой колонны, возникающие в процессе бурения риски.

3. Глубина – расстояние в метрах, которое необходимо преодолеть буру от начала до конца каждого пласта.

4. Издержки – сумма всех затрат в рублях, необходимых для преодоления каждого пласта. Формула, заложенная в программе:

$$C = L \cdot D \cdot R \cdot (B + P),$$

где  $C$  – издержки в рублях,  $L$  – глубина пласта в метрах,  $D$  – коэффициент категории сложности по буримости,  $R$  – коэффициент риска,  $B$  – базовая стоимость бурения одного метра пласта первой категории сложности по буримости в рублях,  $P$  – стоимость типа применяемой буровой колонны по диаметру в рублях.

5. Затраты времени – количество времени в днях и часах, необходимое на прохождение каждого пласта. Формула, заложенная в программе:

$$T = L \cdot D \cdot t,$$

где  $T$  – затраты времени на прохождение пласта,  $L$  – глубина пласта в метрах,  $D$  – коэффициент категории сложности по буримости,  $t$  – время прохождения одного метра пласта первой категории сложности по буримости в часах.

6. Износ бур.кол. – физический износ бурильной колонны, возникающий в процессе прохождения каждого пласта, выраженный в процентах, где 100% составляет полный износ. Формула, заложенная в программе:

$$Dep = \frac{L \cdot D \cdot 100}{M},$$

где  $Dep$  – износ буровой колонны,  $L$  – глубина пласта в метрах,  $D$  – коэффициент категории сложности по буримости,  $M$  – усредненное расстояние прохождения буровой колонны через пласты первой категории сложности по буримости до полного износа.

В нижней части интерфейса статистики отображается информация о полной глубине, сумме издержек, всего времени и всего износа, где содержатся суммы соответствующих значений по столбцам из основной части статистики.

Основную часть окна программы занимает визуальное отображение буровой установки, находящейся в самой верхней части схематически расположенных последовательно пластов земли. Каждый слой имеет некоторое искривление, симулирующее неровность реально существующих земных пластов и позволяющее при перемещении виртуального бура, а также при изменении положения буровой установки получать различные значения глубин указанных пластов.

От виртуального бура исходит красно-зеленый луч (наподобие лазерного дальномера), оставляющий при

каждом пересечении нового пласта ярко-красные квадратные метки, позволяющие визуально оценить расстояние прохождения бура и буровой колонны. Общая длина луча символизирует максимальную глубину бурения данной буровой установкой. Данный луч служит для вычисления глубины пластов в метрах, что составляет основу для калькуляции всех формул, результаты которых представлены в окне интерфейса.

Благодаря работе данного прототипа можно достаточно быстро вручную найти подходящие варианты работы буровой установки исходя из меняющихся внешних условий. Однако по-настоящему данная ИС получит благодаря включению алгоритмов машинного обучения, которые позволят автоматически находить оптимальные варианты для каждого указанного параметра, значительно сэкономив время на его ручном подборе [12].

## VIII. ЗАКЛЮЧЕНИЕ

В данном исследовании мы рассмотрели основные понятия и особенности разработки информационной системы при внедрения цифровых двойников буровых установок на основе машинного обучения. Благодаря анализу конкретных метрик на примере данного проекта можно экстраполировать промежуточные результаты на аналогичные системы с применением других промышленных объектов, имеющих в основе физический актив.

Повсеместное применение цифровых двойников на основе машинного обучения уже сегодня становится такой же обыденностью, как и массовое производство автомобилей около века назад. Пока еще остается масса сфер хозяйственно-экономической деятельности, где ощущается нехватка подобных информационных решений, однако на базе опыта внедрения уже имеющихся решений можно достаточно быстро приспособить их для решения практически любой задачи [10].

Благодаря наличию открытых технологий, свободного доступа к информации и базам данных, а также относительно недорогим техническим средствам, реализовать цифровые двойники и машинное обучение сегодня можно буквально в домашних условиях [15].

В результате проведенного нами исследования можно сделать вывод, что информационная система программного обеспечения – это не просто программный код. Это решение потребности заказчика, покрытие его проблем современными технико-технологическими средствами. И, учитывая мультипликативный эффект от потенциального массового внедрения, чем более качественным будет программное решение, тем больше пользы оно принесет как клиенту в частности, так и экономике в целом.

## БИБЛИОГРАФИЯ

1. БЫКОВА В.Н., КИМ Е., ГАБДУЛШАЕВ М.Р., МУСИЕНКО В.О., ОГУРЦОВ А.О., ТУРОВСКАЯ Е.А. *Применение цифрового двойника в нефтегазовой отрасли* // Актуальные проблемы нефти и газа. 2020. Вып. 1(28). С. <https://doi.org/10.29222/ipng.2078-5712.2020-28.art8>
2. ЕРЕМИНН.А., ЕРЕМИНА Л. Н. *Цифровой двойник в нефтегазовом производстве* // Нефть. Газ. Новации; 2018; №12; С.14–17
3. ЗОЛОТАРЕВ О.В. *Применение цифровых двойников в горнодобывающей промышленности: цифровой двойник мельницы измельчения.* – Международный форум SEYMARTEC Mining. Эффективность и безопасность горнодобывающей промышленности, Челябинск, 2019.
4. КОКОРЕВ Д.С., ЮРИН А. А. *Цифровые двойники: понятие, типы и преимущества для бизнеса:* [Электронный ресурс]: URL: <https://cyberleninka.ru/article/n/tsifrovyye-dvoyniki-ponyatie-tipy-i-preimuschestva-dlya-biznesa>.
5. ЛЕОНТЬЕВА И.Н. *Технология «цифровой двойник» как инструмент интеграции между вузами и промышленными предприятиями.* *ПРОНЕФТЬ.* Профессионально о нефти. 2022;7(3): 119-128. <https://doi.org/10.51890/2587-7399-2022-7-3-119-128>
6. МАРКОВА Н. И., УЖЕГОВА А. М. *Цифровая реальность: от модели к образу* // Вопросы управления. 2019. № 2 (38). С. 45—50
7. МАТВЕЕВА А.В. *Большие данные как исследовательская технология: возможности и ограничения применения в современной управленческой* // Общество: социология, психология, педагогика. – 2021. – № 12. – С. 94–103.
8. *Искусственный интеллект (ИИ)* [Электронный ресурс]: – Режим доступа: <https://www.oracle.com/cis/artificial-intelligence/>
9. *Сведения о моделях двойников и их определении в Azure Digital Twins* [Электронный ресурс]: – Режим доступа: <https://learn.microsoft.com/ru-ru/azure/digital-twins/concepts-models>
10. СОСФЕНОВ Д.А. *Цифровой двойник: история возникновения и перспективы развития* // Интеллект. Инновации. Инвестиции. – 2023. – №4. – С. 35-43, <https://doi.org/10.25198/2077-7175-2023-4-35>.
11. *Цифровой двойник* [Электронный ресурс]: – Режим доступа: [https://hes.mephi.ru/?page\\_id=18708](https://hes.mephi.ru/?page_id=18708)
12. ЧЕРНИКОВ А.Д., ЕРЕМИН Н.А., СТОЛЯРОВ В.Е., СБОЕВ А.Г., СЕМЕНОВА-ЧАЩИНА О.К., ФИЦНЕР Л.К. (2020). *Применение методов искусственного интеллекта для выявления и прогнозирования осложнений при строительстве нефтяных и газовых скважин: проблемы и основные направления решения.* Георесурсы, 22(3), с. 87–96. DOI:<https://doi.org/10.18599/grs2020.3.87-96>
13. *Что такое технология цифрового двойника?* [Электронный ресурс]: – Режим доступа: <https://aws.amazon.com/ru/what-is/digital-twin/>
14. ВНОВМИК S. *Digital twin of subsea pipelines: conceptual design integrating IoT, machine learning and data analytics* // Offshore Technology Conference. 6–9 May 2019; Proceedings. Houston, Texas, USA, 2019. Paper OTC-29455-MS. 9 p
15. BÖSCHEN S., HEINRICH C., ROSEN R. *Next Generation Digital Twin* // Proceedings of TMCE 2018, Las Palmas de Gran Canaria, Spain, 7-11 May, 2018.
16. *Digital Twins for Industrial Applications: An Industrial Internet Consortium White Paper.* Version 1. 2020-02-18. 19 p.
17. GRIEVES, M. *Virtually Intelligent Product Systems: Digital and Physical Twins*, in *Complex Systems Engineering: Theory and Practice*, S. Flumerfelt, et al., Editors. 2019, American Institute of Aeronautics and Astronautics. p. 175-200
18. MICHAEL G. KAPTEYN, JACOB V.R. PRETORIUS, AND KAREN E. WILLCOX. *A Probabilistic Graphical Model Foundation for Enabling Predictive Digital Twins at Scale* [Электронный ресурс]: <https://kiwi.oden.utexas.edu/papers/Digital-twin-scale-graphical-model-Kapteyn-Willcox.pdf>
19. NOSHATI, C.L. & SCHUBERT, J.J. *The Role of Machine Learning in Drilling Operations.* A Review, in *Society of Petroleum Engineers.* 2018 [Электронный ресурс]: <https://doi.org/10.2118/191823-18ERM-MS>

20. SINGH, K., YALAMARTY, S.S., KAMYAB, M., & CHEATHAM, C. *Cloud-Based ROP Prediction and Optimization in Real Time Using Supervised Machine Learning*, Unconventional Resources Technology Conference. 2019[Электронный ресурс]: <https://doi.org/10.15530/urtec-2019-343>.

Петр Алексеевич Болдырев, к.т.н., доцент, Оренбургский государственный университет, Оренбург, Россия (e-mail: [uprunis@mail.osu.ru](mailto:uprunis@mail.osu.ru)).



# Features of the development of the digital twin drilling rig information system project based on big data technologies, machine learning, and the Internet of Things

Peter Boldyrev

**Abstract:** The article considers a schematic diagram of the development of an information system for a digital twin of a drilling rig based on elements of the fourth industrial revolution ("Industry 4.0"). Subsystems using artificial intelligence algorithms, Internet of Things technology (IoT-InternetofThings), big data (BigData) are investigated. The interrelation of subsystems with the solution of information system tasks is analyzed. The characteristics of the hardware are described, as well as ways of storing and presenting information system data. Special attention is paid to modeling the types of information requests and data transmission methods. A list of technical and software tools is described, with the help of which there is a fundamental possibility of implementing the project. Numerical metrics of parameters are given, on the basis of which quantitative and qualitative characteristics of the information system project are predicted.

**Keywords:** information system project, digital twin, big data, machine learning, Internet of things, drilling rig.

## REFERENCES

1. BYKOVA V.N., KIM E., GABDULSHAEV M.R., MUSIENKO V.O., OGURCOV A.O., TUROVSKAJA E.A. Primenenie cifrovogo dvojnika v neftegazovoj otrasli // Aktual'nye problemy nefti i gaza. 2020. Vyp. 1(28). S. <https://doi.org/10.29222/ipng.2078-5712.2020-28.art8>
2. EREMINN.A., EREMINA I. N. Cifrovoy dvojniki v neftegazovom proizvodstve // Neft'. Gaz. Novacii; 2018; #12; S.14-17
3. ZOLOTAREV O.V. Primenenie cifrovih dvojnikov v gornodobyvajushhej promyshlennosti: cifrovoy dvojniki mel'nicy izmel'chenija. – Mezhdunarodnyj forum SEYMARTEC Mining. Jeftektivnost' i bezopasnost' gornodobyvajushhej promyshlennosti, Cheljabinsk, 2019.
4. KOKOREV D.S., JURIN A. A. Cifrovye dvojniki: ponjatie, tipy i preimushhestva dlja biznesa: [Jeftektronnyj resurs]; URL: <https://cyberleninka.ru/article/n/tsifrovye-dvoyniki-ponyatie-tipy-i-preimushhestva-dlja-biznesa>.
5. LEONT'EVA I.N. Tehnologija «cifrovoy dvojniki» kak instrument integracii mezhdu vuzami i promyshlennymi predpriyatijami. PRONEFT'. Professional'no o nefti. 2022; 7(3): 119-128. <https://doi.org/10.51890/2587-7399-2022-7-3-119-128>
6. MARKOVA N. I., UZHEGOVA A. M. Cifrovaja real'nost': ot modeli k obrazu // Voprosy upravlenija. 2019. # 2 (38). S. 45—50
7. MATVEEVA A.V. Bol'shie dannye kak issledovatel'skaja tehnologija: vozmozhnosti i ogranichenija primenenija v sovremennoj upravlencheskoj // Obshhestvo: sociologija, psihologija, pedagogika. – 2021. – # 12. – S. 94–103.
8. Iskustvennyj intellekt (II) [Jeftektronnyj resurs]: – Rezhim dostupa: <https://www.oracle.com/cis/artificial-intelligence/>
9. Svedenija o modeljah dvojnikov i ih opredelenii v AzureDigitalTwins [Jeftektronnyj resurs]: – Rezhim dostupa: <https://learn.microsoft.com/ru-ru/azure/digital-twins/concepts-models>
10. SOSFENOV D.A. Cifrovoy dvojniki: istorija vznikovenija i perspektivy razvitiya // Intellekt. Innovacii. Investicii. – 2023. – #4. – S. 35-43, <https://doi.org/10.25198/2077-7175-2023-4-35>.
11. Cifrovoy dvojniki [Jeftektronnyj resurs]: – Rezhim dostupa: [https://hes.mephi.ru/?page\\_id=18708](https://hes.mephi.ru/?page_id=18708)
12. ChERNIKOV A.D., EREMIN N.A., STOLJAROV V.E., SBOEV A.G., SEMENOVA-ChASHHINA O.K., FICNER L.K. (2020). Primenenie metodov iskustvennogo intellekta dlja vyjavlenija i prognozirovanija oslozhnenij pri stroitel'stve neftjanyh i gazovyh skvazhin: problemy i osnovnye napravlenija reshenija. Georesursy, 22(3), s. 87–96. DOI: <https://doi.org/10.18599/grs2020.3.87-96>
13. Chto takoe tehnologija cifrovogo dvojnika? [Jeftektronnyj resurs]: – Rezhim dostupa: <https://aws.amazon.com/ru/what-is/digital-twin/>
14. BHOWMIK S. Digital twin of subsea pipelines: conceptual design integrating IoT, machine learning and data analytics // Offshore Technology Conference. 6–9 May 2019; Proceedings. Houston, Texas, USA, 2019. Paper OTC-29455-MS. 9 p
15. BÖSCHEN S., HEINRICH C., ROSEN R. Next Generation Digital Twin // Proceedings of TMCE 2018, Las Palmas de Gran Canaria, Spain, 7-11 May, 2018.
16. Digital Twins for Industrial Applications: An Industrial Internet Consortium White Paper. Version 1. 2020-02-18. 19 p.
17. GRIEVES, M. Virtually Intelligent Product Systems: Digital and Physical Twins, in Complex Systems Engineering: Theory and Practice, S. Flumerfelt, et al., Editors. 2019, American Institute of Aeronautics and Astronautics. p. 175-200
18. MICHAEL G. KAPTEYN, JACOB V.R. PRETORIUS, AND KAREN E. WILLCOX. A Probabilistic Graphical Model Foundation for Enabling Predictive Digital Twins at Scale [Jeftektronnyj resurs]: <https://kiwi.odon.utexas.edu/papers/Digital-twin-scale-graphical-model-Kapteyn-Willcox.pdf>
19. NOSHATI, C.L. & SCHUBERT, J.J. The Role of Machine Learning in Drilling Operations. A Review, in Society of Petroleum Engineers. 2018 [Jeftektronnyj resurs]: <https://doi.org/10.2118/191823-18ERM-MS>
20. SINGH, K., YALAMARTY, S.S., KAMYAB, M., & CHEATHAM, C. Cloud-Based ROP Prediction and Optimization in Real Time Using Supervised Machine Learning, Unconventional Resources Technology Conference. 2019 [Jeftektronnyj resurs]: <https://doi.org/10.15530/urtec-2019-343>