

Цифровые двойники: стратегии и подходы к созданию систем экологического мониторинга

В. П. ПОТАПОВ*, С. Е. ПОПОВ, Е. Л. СЧАСТЛИВЦЕВ

Федеральный исследовательский центр информационных и вычислительных технологий,
650025, Кемерово, Россия

*Контактный автор: Потапов Вадим Петрович, e-mail: vadimptpv@gmail.com

Поступила 05 апреля 2023 г., доработана 18 апреля 2023 г., принята в печать 24 апреля 2023 г.

Целью настоящей работы является исследование возможности использования концепции цифровых двойников применительно к созданию систем экологического мониторинга. Приведена их классификация и сделан вывод о возможности применения определенного типа цифровых двойников к системам экологического мониторинга, представляющего собой комплекс средств сбора, структурированного хранения, последующей обработки и анализа информационных потоков, обеспечивающих данные о состоянии окружающей среды. В качестве базового элемента для создания цифрового двойника предложено использовать цифровую фабрику данных, являющуюся прямым аналогом цифровых теней и обеспечивающую непосредственное взаимодействие со средой. Определены основные инструменты и конструкции, реализующие цифровую фабрику данных, и делается вывод о возможности применения разработанного подхода к созданию цифрового двойника для экологического мониторинга. Предложенный подход реализован в системе экологического мониторинга предприятий угольной промышленности.

Ключевые слова: цифровой двойник, цифровая фабрика, контейнеризация, мониторинг, информационные потоки, оркестрация.

Цитирование: Потапов В.П., Попов С.Е., Счастливцев Е.Л. Цифровые двойники: стратегии и подходы к созданию систем экологического мониторинга. Вычислительные технологии. 2023; 28(3):167–181. DOI:10.25743/ICT.2023.28.3.010.

Введение

За последние 10 лет создано значительное число информационных технологий (ИТ), которые можно отнести к новому поколению [1], существенно изменяющих саму парадигму обработки информации. К ним следует отнести:

- большие данные (Big Data) [2];
- Интернет вещей (умные датчики и сенсоры — IoT) [3];
- облачные вычисления (cloud computing) [4];
- методы искусственного интеллекта (самообучение на основе потоков данных) (machine learning, deep learning, reinforcement learning) [5];
- науку о данных (data science — DS) [6];
- киберфизические системы (cyber-physical systems — CPS) [7];
- другие технологии, например цифровые двойники (digital twins) [8].

Несмотря на тот факт, что каждая из технологий достаточно самостоятельна сама по себе, созданы платформы, которые могут объединять их определенное подмножество, позволяя тем самым получать новые знания из постоянно нарастающих потоков разнородных данных, что и способствовало разработке новых подходов. В настоящей работе мы рассмотрим вопросы создания цифровых двойников и фабрик данных, использующих некоторые из указанных выше технологий.

1. Стратегии создания цифровых двойников

Концепцию цифровых двойников впервые представил публике в 2002 г. Майкл Гривз, профессор Мичиганского университета. В докладе, посвященном управлению жизненным циклом продукта (PLM), он рассказал о возможностях, открывающихся при создании виртуального пространства, которое дублировало бы реальное пространство и обменивалось с ним информацией [9]. Дальнейшему развитию этого подхода способствовали публикации [10, 11]. Начиная с 2015 г. большое количество организаций и фирм, занимающихся новыми информационными технологиями, обратили на цифровые двойники пристальное внимание, и сегодня уже существует несколько концепций (стратегий) их создания, для чего рассмотрим некоторые базовые понятия новых ИТ.

Аналитики Forrester [12] определяют цифрового двойника как процесс создания реального физического объекта в абстрактной цифровой форме, который выступает в качестве посредника для любой связи с реальным устройством. Технология цифровых двойников заключается не только в тестировании продукта или его компонентов. Цифровой двойник способен анализировать и моделировать условия реального мира, реагировать на операционные изменения, используя данные о принципах функционирования частей системы и ответов на окружающие условия, а также показания датчиков, установленных в реальном мире. Цифровой двойник функционирует главным образом как агент целого ряда реальных экспертов, устройств и средств мониторинга. Базовая концепция двойников не сложна для понимания: мониторинг физического объекта осуществляется на основе замкнутого цикла информационного обмена между ним и его виртуальной моделью (тем самым цифровым двойником) (рис. 1).

Наиболее распространенная сегодня классификация включает три типа цифровых двойников [13].

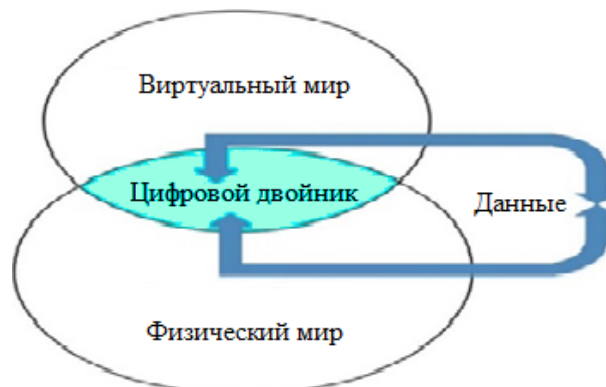


Рис. 1. Концептуальная схема цифрового двойника

Fig. 1. Conceptual scheme of the digital twin

Двойник-прототип (DTP) — это прототип некоторого физического объекта, включающий информацию, необходимую для создания физической версии объекта и описания его свойств. Для производственных процессов он может включать 3D-модель объекта, условия для его создания, технологические регламенты, необходимые материалы и т. п. Чаще всего используется в машиностроении, авиастроении и других отраслях, где может осуществляться сборка целого из разнообразных частей.

Двойник-экземпляр (DTI) применяется в случае, когда имеется конкретный физический объект, с которым двойник остается связанным на протяжении всего времени жизни. Содержит 3D-модель с характерными параметрами, спецификацию на материалы, используемые при создании физического объекта, спецификацию технологических процессов и операций, которые применяются при его создании, а также при любых тестовых испытаниях, архивы сервисного обслуживания, включая замену отдельных деталей, текущие и прогнозируемые значения параметров мониторинга, возникающие в процессе производства.

Двойник-агрегатор (DTA) — распределенная информационно-вычислительная система, организующая доступ ко всем цифровым двойникам-экземплярам, которая может посылать им потоковые (активные и пассивные) запросы, получая необходимую информацию для последующего воздействия на их состояние. Иными словами, цифровые двойники предполагают управление связью между пограничными устройствами и внутренними системами с последующим зеркальным отражением изменений в виртуальной модели устройства и передачей агрегированной информации лицам, принимающим решения. Это наиболее интересный цифровой двойник с точки зрения построения систем мониторинга.

На основе приведенной классификации цифровых двойников разработаны подходы к созданию реальных систем для различных приложений ведущими мировыми корпорациями, такими как Oracle, IBM, Microsoft, General Electric, Ansys, которые стали лидерами в этой новой области [14, 15]. Значительный объем исследований и конкретные разработки, выполненные зарубежными фирмами в рамках технологий цифровых двойников, позволяют сделать вывод о том, что это направление сегодня открывает возможность значительно повысить конкурентность как самих производств, так и выпускаемой продукции. К сожалению, в нашей стране в этой области хотя и проводятся работы в рамках проекта “Цифровая экономика”, однако их объем явно недостаточен.

Наиболее значимой разработкой мирового уровня является выполненная с применением фабрик будущего (factories of the future) система комплексных технологических решений, ключевым элементом которых становятся “умные” математические модели и “умные” цифровые двойники (smart digital twins) объектов (изделий, продуктов) производства и технологических (производственных) процессов. Концепция фабрик будущего реализуется в рамках дорожной карты “Технет” Национальной технологической инициативы (НТИ) и мегапроекта “Фабрики будущего”, нацеленного на развитие и повышение конкурентоспособности отечественной высокотехнологичной промышленности за счет решения инженерно-технологических проблем-вызовов (industrial challenge problems) государственного значения, которые не удается решить предприятиям с помощью традиционных подходов [16].

В рамках этого проекта разработана технология, объединившая методы DMU (digital mock-up, цифровые прототипы), CAD (computer-aided design, компьютерное проектирование), CAE (computer-aided engineering, конечно-элементные модели), PLM (product lifecycle management, управление жизненным циклом продукта) и вошедшая в кривую

зрелости технологий Гартнера в 2016 г. [17]. Используемые математические модели, или “умные” модели, агрегируют в себе все знания, которые применяются при создании продукта:

- фундаментальные науки и законы;
- геометрические (CAD) и вычислительные инженерные (CAE) полномасштабные модели реальных объектов и физико-механических процессов;
- полные данные о материалах, из которых изготавливается изделие;
- информацию об эксплуатационных режимах, включая информацию, которая обеспечивает заданное поведение конструкции в тех или иных ситуациях (так называемое программируемое поведение);
- данные о технологиях производства, как отдельных элементов, так и конструкций в целом;
- прочие параметры.

Полученный “умный” цифровой двойник, создаваемый в результате численного моделирования и оптимизации на основе “умных” больших данных, имеет мало общего с простой геометрической 3D-моделью и кинематическими расчетами, с которыми зачастую ассоциируются цифровое проектирование и моделирование. При создании умной модели задается многоуровневая матрица целевых показателей и ресурсных ограничений (временных, финансовых, технологических, производственных и т. д.). В сложной технической системе количество показателей может достигать десятков или даже сотен тысяч, в отличие от традиционного варианта, где их около 500. Далее, на основе выполнения виртуальных испытаний формируется итоговый цифровой двойник (smart digital twin) реального объекта, который ведет себя с высокой степенью точности так же, как и реальный объект на всех этапах жизненного цикла. Такой новый подход, отвечающий глобальным трендам четвертой промышленной революции и цифровой экономики, позволяет проводить разработки значительно быстрее, дешевле и эффективнее.

В настоящей работе широко использовано понятие цифровой тени, которая определяется как система связей и зависимостей, описывающих поведение реального объекта, как правило, в нормальных условиях работы и содержащихся в избыточных больших данных, получаемых с реального объекта при помощи технологий промышленного Интернета. Цифровая тень способна предсказать поведение реального объекта только в тех условиях, в которых осуществлялся сбор данных, но она не позволяет моделировать ситуации, в которых реальный объект не эксплуатировался. Это интересное понятие, которое пока редко используется при разработке цифровых двойников, но несомненно может применяться в различных областях, в том числе и научных исследованиях, особенно там, где объекты работают в стационарных режимах, отклонений от которых обычно не происходит. При использовании цифровых теней существенно упрощаются многие аспекты создания цифровых двойников, а также экономятся вычислительные ресурсы.

По словам руководителя проекта — проф. Алексея Боровкова [18], “говоря о цифровых тенях, необходимо обращать внимание на генерацию данных. Обычно все говорят про хранение, передачу, обработку, защиту информации, но ключевой элемент в этой цепочке — генерация данных, потому что именно этот этап определяет объем задач по их обработке. Многие мировые компании уже генерируют больше информации, чем могут обработать, извлекая оттуда важные сведения. Генерировать нужно лишь содержательные данные, очищенные от мусора и шума. Нецелесообразно сплошь обклеивать производство датчиками, потому что это неминуемо приведет к терабайтам, которые

мы не будем успевать обрабатывать. И цифровой двойник здесь может помочь: он позволяет указать на критические зоны производства и критические характеристики производимого продукта, чтобы собирать только такую нужную информацию.” С нашей точки зрения, именно в процессе создания цифровой тени можно решать такие задачи, как создание систем мониторинга на основе технологий цифровых двойников.

Современные системы мониторинга формируются главным образом на основе инфраструктуры пространственных данных в соответствии с распоряжением Правительства РФ № 1157-р от 21 августа 2006 г., которая определяется как информационно-телекоммуникационная система, обеспечивающая доступ пользователей к национальным распределенным ресурсам пространственных данных, а также распространение и обмен в сети Интернет в целях их производства и использования [19]. Наиболее общее определение системы мониторинга как системы регламентированного сбора данных, организации их структурированного хранения, обработки и анализа с целью принятия необходимых управляющих решений сегодня должно перетерпеть некоторые изменения, в том числе вызванные появлением технологий цифровых двойников. В современной интерпретации любые системы мониторинга являются пассивными приемниками пространственных данных, а принятие решений на их основе никак не связывается с изменением их параметров.

Практически система не знает, что происходит с результатами анализа, которые она генерирует для управления внешней средой, так как не получает сигналов от самой среды, за которой проводится наблюдение. И этот факт существенным образом сказывается на качестве управляющих решений, так как они выполняются с запаздыванием даже в случае систем, работающих в режиме реального времени. Кроме того, современные системы мониторинга, в большинстве своем рассчитанные на традиционные пространственные данные, получаемые посредством точечных замеров, и их анализ достаточно проблематично распространить на большие площади, что является характерной особенностью различных горнопромышленных систем мониторинга (геомеханического, технологического, экологического и др.). Отсутствие связи с различными математическими моделями и современными методами обработки и анализа данных еще более усугубляет данную ситуацию.

По нашему мнению, использование технологий цифровых двойников поможет в создании современных систем мониторинга, наиболее адекватно описывающих процессы наблюдения за сложными природными объектами, а также за контролем над ними с учетом современных тенденций развития средств измерения, методов искусственного интеллекта и новейших математических моделей (имеются в виду 3D-модели в гетерогенных средах). Для конкретизации декларируемого подхода рассмотрим создание конкретных систем экологического мониторинга предприятий горнопромышленного комплекса (на примере угольной промышленности) с использованием технологии цифровых двойников. Общая концептуальная схема предполагаемой к созданию системы мониторинга с использованием технологий цифровых двойников показана на рис. 2.

Как следует из рисунка, цифровой двойник системы экологического мониторинга относится к ДТА-типу, он представляет собой распределенную информационно-вычислительную систему, состоящую из нескольких взаимно связанных блоков, организующих доступ ко всем элементам цифрового двойника. Система может посылать потоковые (активные и пассивные) запросы, получая необходимую информацию для последующего воздействия на их состояние, на это указывают наличие обратной связи с конечным блоком принятия решений и двойные связи между всеми элементами двойника. До-

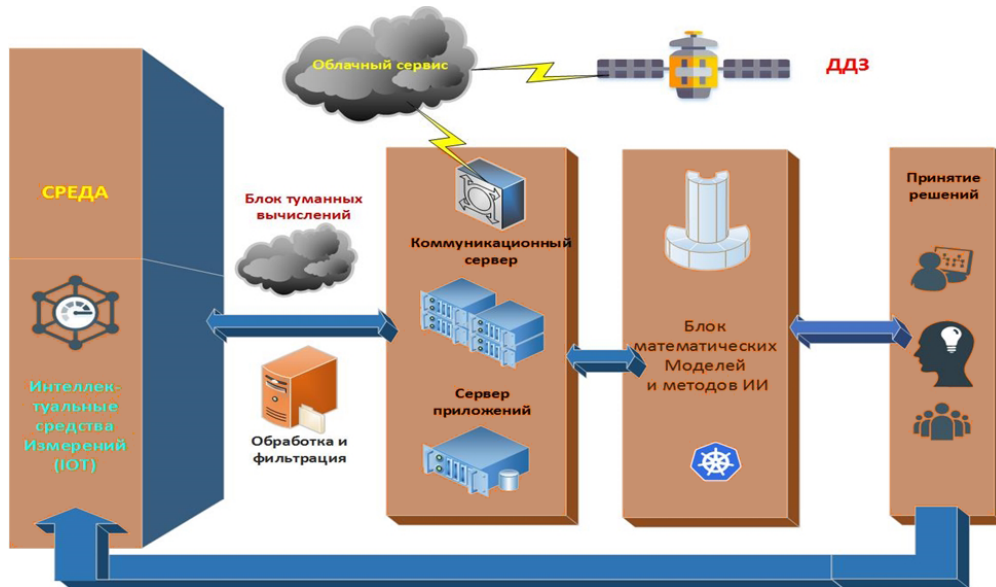


Рис. 2. Концептуальная схема цифрового двойника системы экологического мониторинга
 Fig. 2. Conceptual scheme of the digital twin for the environmental monitoring system

полнительно введенный блок туманных вычислений предназначен для оптимизации потоков данных, он генерирует содержательные данные от пороговых устройств типа интеллектуальных датчиков, проводя предварительную обработку информации, например фильтруя шумы на изображениях.

2. Создание цифровой фабрики данных как базового элемента цифрового двойника

С точки зрения архитектуры цифровой двойник представляет собой интеграцию цифровой фабрики данных и цифровой платформы моделирования, которая включает конвейеры, состоящие из численных моделей и наборов методов искусственного интеллекта. Рассмотрим подробно структуру цифровой фабрики данных и ее отдельных элементов, необходимых для включения их в цифровой двойник.

Сформулированный нами подход, основанный на создании цифровых фабрик данных, позволяет достаточно гибко определить общую платформу для использования стека данных, не привязываясь к конкретной реализации отдельных элементов такой системы и изменяя их при необходимости. Цифровая фабрика является своеобразным интегратором для получения, обработки, хранения и анализа потоков данных, оркестрируя и распределяя их по конвейерам соответственно специфике разрабатываемой системы. С другой стороны, в терминах создания цифровых двойников-агрегаторов она представляет собой в определенном смысле цифровую тень и является их основным базовым элементом.

В наиболее общем виде цифровая фабрика данных — цельная концептуальная архитектура управления информацией с полным и гибким интерфейсом взаимодействия с ней. Это единая и согласованная архитектура управления данными, которая обеспечивает беспрепятственный доступ к данным и их обработку [20].

Главная особенность современных цифровых фабрик данных заключается в интенсивном использовании подходов и инструментария Big Data и AI (искусственного ин-

теллекта), а также machine learning (машинного обучения) для управления и обработки получаемых данных. Важно отметить, что под цифровой фабрикой данных обычно понимают замкнутую (автономную) экосистему, которая используется для максимально эффективного доступа к корпоративным данным, а не определенную площадку от конкретного производителя программного обеспечения.

Современные цифровые фабрики данных позволяют эффективно справляться с основными задачами в плане получения, хранения и обработки разрозненной информации. С их помощью такую информацию легче искать, обрабатывать, структурировать и интегрировать с другими системами ИТ-инфраструктуры. В части безопасности данных цифровые фабрики также выгодно выделяются на фоне альтернативных вариантов, так как позволяют:

- обеспечить надежную защиту информации;
- реализовать управление информацией стандартными открытыми интерфейсами API;
- максимально гибко и тонко настраивать доступ к информации для отдельных категорий пользователей сети.

Архитектура цифровых фабрик данных нацелена на максимальную прозрачность процессов анализа, обновления, интеграции, маршрутизации, а также трансформации данных в соответствии с конкретными требованиями бизнеса. Компания Forrester Research [21] определила архитектуру, которая помогает обрабатывать, преобразовывать, защищать и координировать данные из разрозненных источников данных для обеспечения надежного представления корпоративных данных в реальном времени, и назвала ее фабрикой больших данных. Вследствие типовой архитектуры, построенной на шести различных уровнях (рис. 3), она охватывает потоковую передачу данных, гибридные и многооблачные развертывания.

Цифровая фабрика данных благодаря своей архитектуре как системный интегратор обеспечивает оперативное реагирование на события, высокий уровень прогнозируемости, оптимизацию обработки и обслуживания ресурсов. На текущий момент указанные технологии представляют собой тренд в области Big Data и корпоративного ИТ-сектора, а не готовые технологические решения. На практике для сквозной интеграции и управления хранилищами данных используются конкретные технологии Big Data: Apache Kafka, Spark, Hadoop, Hive, NiFi, AirFlow и прочие средства для сбора, обработки, маршрутизации и преобразования пакетных и потоковых данных в различных форматах.

Помимо упомянутых и других инструментов Big Data концепция цифровых фабрик данных может быть дополнена семантическими графами, которые позволяют определять, стандартизировать и согласовывать все входящие данные в бизнес-терминах, понятных конечным пользователям. Сама цифровая фабрика данных по максимуму использует весь потенциал облачных технологий, виртуализируя все компоненты ИТ-инфраструктуры — от наборов информации до программных приложений. Подобная сервисная модель частично соответствует DevOps-подходу [22], а потому инструменты контейнеризации (такие как Docker, Kubernetes [23]) также можно отнести к средствам для создания цифровых фабрик данных.

Архитектура цифровых фабрик данных идеально укладывается в концепцию Data Ops [24], на основе которой можно организовать быстрое реагирование для любых изменений в данных, получить максимально высокий уровень прогнозирования, оптимизировать процессы хранения, обработки и обслуживания ресурсов, работающих с ин-

формацией. Надо иметь в виду, что процессы, в которых задействованы Data Ops и Dev Ops, значительно различаются. В DataOps используется гораздо большее количество инструментов, так как приходится иметь дело с множеством источников и моделей данных, управляя информационными потоками (оркестрация), корпоративными озерами (Data Lakes) и “песочницами” данных (Sandbox Management), формируя воспроизводимые среды для работы с Big Data. Поэтому под фабрикой данных мы подразумеваем определенный набор инструментов и процессов:

- на конкретных шагах обработки информации используется машинное обучение (от анализа получаемых данных до оптимизации алгоритмов их обработки);
- все потребители и источники данных получают сквозную интеграцию (в том числе и базы/хранилища данных, озера данных) при помощи соответствующих интерфейсов API;
- вместо монолитных программных платформ применяются микросервисные архитектуры;
- используется наибольшее число возможных облачных решений;
- информационные потоки оркестрируются;
- качество информации повышается после унификации и виртуализации;
- независимо от типа данных к ним обеспечивается быстрый доступ (из баз данных, хранилищ данных, корпоративных озер данных и т. д.);
- обеспечивается безопасный доступ внутри компании (разных групп пользователей) для обработки информации параллельно с гибкой настройкой прав доступа к информации для каждой группы клиентов на корпоративном уровне.

Таким образом, можно сделать вывод, что цифровая фабрика данных благодаря приведенным характеристикам и процессам соответствует концепции Data Ops.

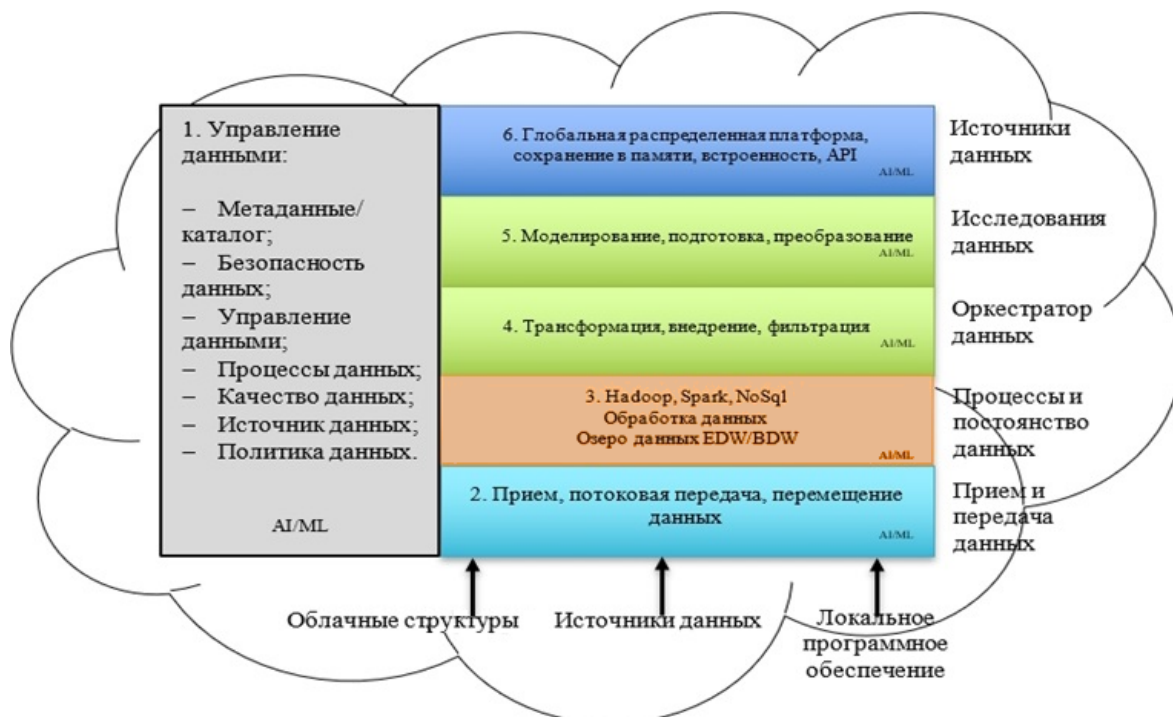


Рис. 3. Типовая архитектура цифровой фабрики данных
Fig. 3. Typical digital data factory architecture

Архитектура цифровых фабрик данных идеально укладывается в технологии цифрового двойника, при помощи которого можно организовать быстрое реагирование на любые изменения в данных, получить максимально высокий уровень прогнозирования, оптимизировать процессы хранения, обработки и обслуживания различных информационных ресурсов.

В рамках предлагаемого подхода на основе цифровых фабрик данных можно реализовать конкретный сценарий работы. Определяются источники данных и протоколы доступа к ним, которые передаются, оркестрируются, конвейеризуются, поступают в хранилище данных для последующей аналитической обработки.

В общем случае можно выделить достаточно большую группу однородных технологий, связанных с получением требуемой информации. При появлении новых источников данных нет необходимости изменять отдельные блоки или сервисы, а следует лишь настраиваться на протоколы передачи и форматы, если они отличаются от существующих в системе.

Технологию создания цифровой фабрики данных как платформы цифрового двойника можно представить следующим образом (рис. 4). Она формируется из типовых элементов, таких как контейнеры (объединяемые в зависимости от типа данных и протоколов их передачи в конвейеры), хранилища данных, блок оркестрации, вычислительные блоки по обработке и анализу данных. На рисунке показаны основные элементы цифровой фабрики данных, а также информационные связи между ними. Каждый из элементов выполняет определенные функции, реализуемые в зависимости от типа данных через множество контейнеров.

Отдельный контейнер представляет собой инкапсулируемый модуль обработки данных, информационно связанный с другими. Кластер контейнеров образует конвейер, в результате работы которого реализуется обработка конкретного типа данных. Промежуточные результаты обработки, а также сами исходные данные помещаются в хранилище, из которого они в дальнейшем могут быть получены на любом этапе работы системы. Это позволяет, с одной стороны, локализовать методы обработки данных, а с другой — осуществлять их эффективное масштабирование для переноса в другие

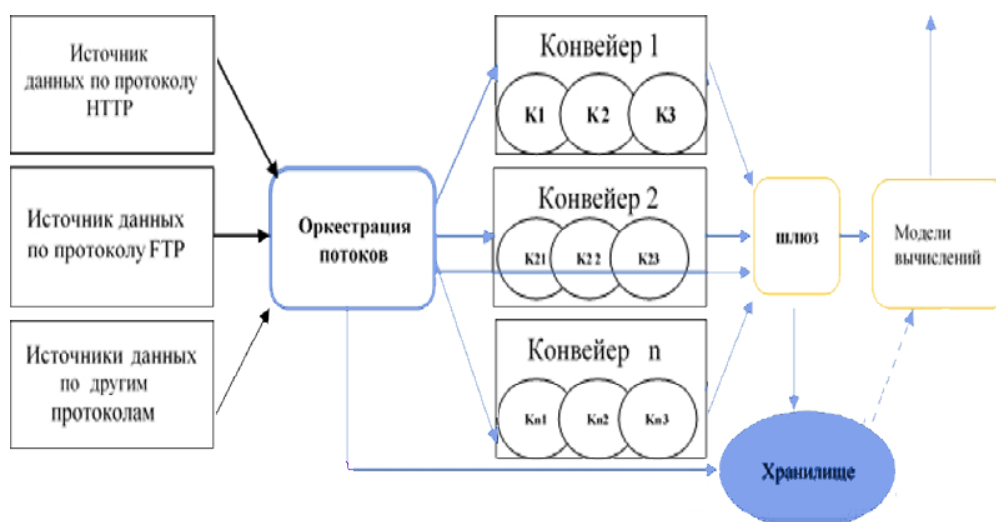


Рис. 4. Концептуальная схема создания цифровой фабрики данных
Fig. 4. Conceptual framework for the creation of a digital data factory

системы, поскольку любой контейнер имеет свой собственный вход и выход. При формировании цифровой фабрики данных фактически все компоненты переводятся на уровень микросервисов. Блок вычислительных моделей также представляет собой набор модулей, реализующих расчетные модели, для передачи результатов анализа на вышележащий уровень. Он готовит цифровые потоки для цифрового двойника, который должен организовывать обратную связь с источниками данных в целях оптимального управления всей системой.

Ключевым элементом цифровой фабрики является компонент, реализующий процедуру ETL (extract, transform, load — “извлечение”, “преобразование”, “загрузка”). В ее рамках данные из нескольких систем-источников проходят ряд стадий преобразования и попадают в систему-приемник (рис. 5).

1. *Процесс загрузки.* Его задача — найти и загрузить в ETL-данные произвольного формата для дальнейшей обработки.
2. *Процесс валидации и очистки данных.* На этом этапе данные последовательно проверяются на корректность и полноту, отсутствие пропущенных значений, составляется отчет об ошибках для исправления.
3. *Процесс взаимодействия данных с целевой моделью.* На этом этапе к валидированной таблице пристраиваются еще n столбцов (по количеству справочников целевой модели данных), а потом по таблицам для каждой дополненной ячейки и для каждой строки проставляются значения целевых справочников.
4. *Процесс агрегации данных.* Этот процесс нужен из-за разности детализации данных в OLTP- и OLAP-системах.
5. *Выгрузка в целевую систему* — технический процесс использования коннектора и передачи данных в целевую систему.

Однако в современных условиях такой подход трансформируется в вариацию ETL, т. е. наиболее важными становятся шаги EL (extract и load — выгрузка и загрузка данных). Операция T (transform — преобразование данных) может быть опциональна и выполняется по мере возникновения задач анализа данных. Одним из инструментов, реализующих данный подход, является программное обеспечение с открытым исходным кодом Airbyte (рис. 6). Оно реализует архитектуру, в рамках которой исполнение ETL-процесса достигается за счет взаимодействия веб-интерфейса пользователя, планиров-

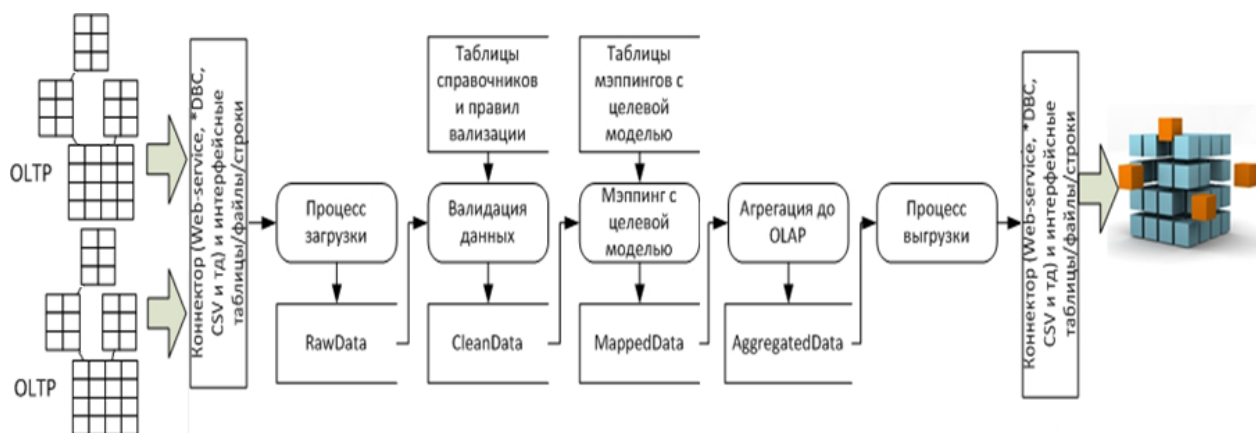


Рис. 5. Общая схема ELT-обработки потоков данных
Fig. 5. General scheme of ELT processing of data flows

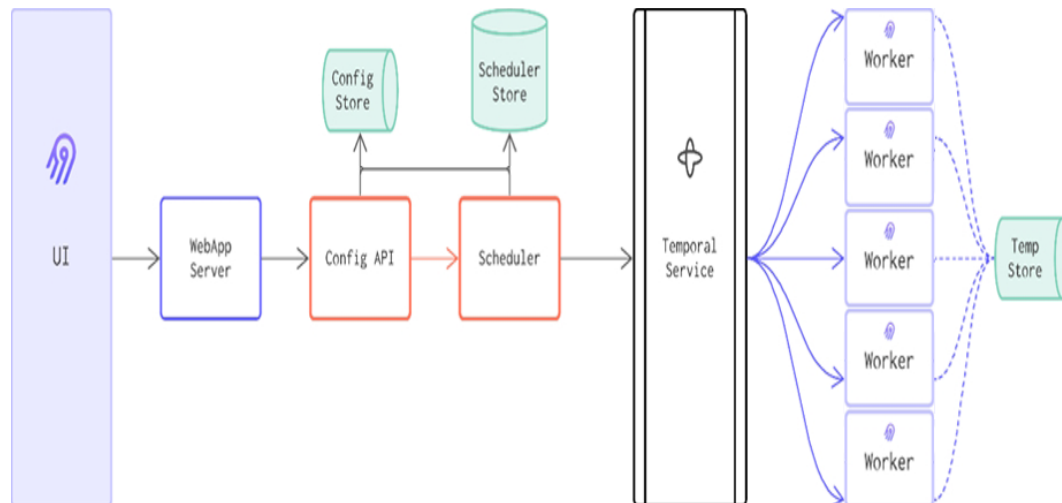


Рис. 6. Модель элемента цифровой фабрики с использованием Airbyte
 Fig. 6. Model of an element for the digital factory using Airbyte

щика, набора коннекторов источников и приемников данных и обработчиков заданий. Все компоненты реализуются в виде контейнеров Docker с заранее определенным набором команд для их запуска.

Таким образом, при разработке цифровой фабрики данных можно учитывать опыт, полученный при анализе ранее созданного программного обеспечения. В частности, подход к разбиению системы на компоненты и их контейнеризацию, т. е. выделение модулей, запуск которых должен осуществляться по требованию других элементов системы, и формирование их API на основе набора опций запуска.

Процесс создания системы экологического мониторинга можно рассматривать как формирование цифрового двойника третьего типа, использующего при создании цифровой платформы описанные выше подходы конструирования цифровой фабрики как базового элемента системы.

Сама платформа цифрового двойника состоит из нескольких взаимно увязанных блоков, определяющих (применительно к системе экологического мониторинга) взаимодействие с внешней средой на уровне получения, обработки и хранения потоков информации о ее состоянии, последующее многофункциональное моделирование отдельных ее элементов и принятие решений на основе некоторой интегральной оценки.

Цифровая фабрика обеспечивает получение информационных потоков и их предварительную обработку с формированием соответствующих структур их хранения. Интеграция фабрики с блоками цифрового моделирования и формирование управляющих решений обеспечивают многофункциональную работу цифрового двойника.

Заключение

Проведен анализ существующих подходов к созданию цифровых двойников с использованием одной из самых известных классификаций их типов. Основное внимание уделено созданию систем мониторинга с помощью цифровых двойников. Предложено использование цифровых фабрик как естественного интегратора информационных потоков для двойников третьего типа. При этом для мониторинга может быть реализована специальная цифровая платформа, которая объединяет данные из различных источников.

Эти данные предварительно обрабатываются и хранятся на уровне фабрики. Последующая интеграция с математическими моделями (в том числе и с методами искусственного интеллекта) дает возможность создания многофункционального цифрового двойника, обеспечивающего необходимые решения по управлению состоянием окружающей среды. Обычно цифровой двойник является более сложной структурой, так как он создается на основе конкретных параметров и характеристик объекта, а также может включать в себя моделирование процессов и взаимодействий с другими объектами.

В то же время цифровая фабрика данных может быть более простой структурой, так как создается на основе данных из различных источников и может не содержать всех деталей и характеристик объекта. Однако это будет зависеть от конкретной задачи, для которой создается цифровой двойник. На основе изложенного подхода реализована система экологического мониторинга, которая сегодня работает на двух угольных предприятиях Кузбасса. Предполагается в дальнейшем развивать систему как с точки зрения ее функциональных возможностей, так и количества измеряемых параметров, поступающих в цифровой двойник.

Благодарности. Данная работа возникла и развивалась по предложению академика Ю.И. Шокина и находилась под его пристальным вниманием много лет, за что авторы работы приносят ему свою благодарность.

Список литературы

- [1] **Crelin J.** Principles of information technology. Grey House Publishing, Inc.; 2020: 410. ISBN:9781642656954.
- [2] **Karimi H.A.** Big Data: techniques and technologies in geoinformatics. CRC Press; 2014: 306.
- [3] **Dow C.** Mastering IoT: build modern IoT solutions that secure and monitor your IoT infrastructure. Packt Publishing; 2019: 782.
- [4] **Ranjan R., Mitra K., Jayaraman P.P., Wang L., Zomaya A.Y.** Handbook of integration of cloud computing, cyber physical systems and Internet of Things. Springer; 2020: 331. ISBN:978-3-030-43794-7.
- [5] **Acharjya D.P., Mitra A., Zaman N.** Deep learning in data analytics: recent techniques, practices and applications. Springer; 2022: 271. ISBN:978-3-030-75854-7.
- [6] **Polkowski Z., Kumar M.S., Vasilev J.** Data science in engineering and management. CRC Press; 2022: 152. ISBN:978-1-032-10625-0.
- [7] **Banerjee J.S., Bhattacharyya S., Obaid A.J., Yeh W.-C.** Intelligent cyberphysical systems security for Industry 4.0: applications, challenges and management. CRC Press; 2022: 284. ISBN:978-1-032-14835-9.
- [8] **Pal S.K., Mishra D., Pal A., Dutta S., Chakravarty D., Pal S.** Digital twin — fundamental concepts to applications in advanced manufacturing. Springer; 2022: 495. ISBN:978-3-030-81814-2.
- [9] **Grieves M.W.** Digital twin: manufacturing excellence through virtual factory replication. LLC; 2014: 7. Available at: https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication.

-
- [10] **Tao F., Sui F., Liu A., Qi Q.** Digital twin-driven product design framework. *International Journal of Production Research*. 2019; (57):3935–3953. DOI:10.1080/00207543.2018.1443229. Available at: <https://www.tandfonline.com/doi/abs/10.1080/00207543.2018.1443229?journalCode=tprs20>.
- [11] **Datta S.** Emergence of digital twins. *Journal of Innovation Management*. 2017; (5):14–34.
- [12] **Miller P.** Digital twins combine enterprise data and IoT to drive new business value. *Principal analyst*. 15, 2019. Available at: <https://go.forrester.com/blogs/digital-twin-iotnew-business-value>.
- [13] **Mottura S., Viganó G., Greci L., Sacco M., Carpanzano E.** New challenges in collaborative virtual factory design. *International Conference on Information Technology for Balanced Automation Systems*. Springer; 2008: 17–24. Available at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1088.9509&rep=rep1&type=pdf>.
- [14] **Ducree J.** Digital twin: an oracle for efficient crowdsourcing of research and technology development through blockchain. *Oracle Corporation Preprint*. 2021: 15. DOI:10.20944/preprints202110.0148.v1. Available at: <https://www.preprints.org/manuscript/202110.0148/v1>.
- [15] The promise of a digital twin strategy. Best practices for designers and manufacturers of products and industrial equipment. *Microsoft Services*. 2022: 23. Available at: <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE1IMIi>.
- [16] **Kukushkin K., Ryabov Yu., Borovkov A.** Digital twins: a systematic literature review based on data analysis and topic modeling. *MDPI, Data*. 2022; 7(12):21.
- [17] **Unhelkar B., Gonsalves T.** Artificial intelligence for business optimization: research and applications. *CRC Press*; 2021: 325. ISBN:978-1-032-02886-6.
- [18] **Прохоров А., Лысачев М.** Цифровой двойник. Анализ, тренды, мировой опыт / Под ред. А. Боровкова. М.: ООО “АльянсПринт”; 2020: 401.
- [19] Инфраструктура пространственных данных. Требования к информационному обеспечению. *ГОСТ Р 58571-2019*. 2019: 17.
- [20] **Maryse C.** Digital organizations manufacturing: scripts, performativity and semiopolitics. *Wiley*; 2018: 318. ISBN:978-1-119-52766-4.
- [21] **Dolgui A., Bernard A., Lemoine D., von Cieminski G., Romero D.** Advances in production management systems. Artificial intelligence for sustainable and resilient production systems. *IFIP WG 5.7 International Conference, APMS 2021, Nantes, France, September 5–9, 2021. Proceedings, pt II*. Springer; 2021: 730. ISBN:978-3-030-85901-5. Available at: <https://link.springer.com/book/10.1007/978-3-030-85902-2>.
- [22] **Sharma S.** The DevOps adoption playbook: a guide to adopting DevOps in a multi-speed IT enterprise. *John Wiley & Sons*; 2017: 416.
- [23] **Deepak V.** Kubernetes microservices with Docker. *Apress*; 2016: 456.
- [24] **Alok K.** Practical full stack machine learning: a guide to build reliable, reusable, and production-ready full stack ML solutions. *BPB Publications*; 2022: 751. ISBN:978-93-91030-42-1.
- [25] **Crickard P.** Data engineering with Python. *Packt Publishing*; 2020: 357.
-

Digital twins: strategies and approaches to creating environmental monitoring systems

ПОТАПОВ ВАДИМ П. *, ПОПОВ СЕМЕОН Е., СЧАСТЛИВЦЕВ ЕВГЕНИ Л.

Federal Research Center for Information and Calculated Technologies, 650025, Kemerovo, Russia

*Corresponding author: Potapov Vadim P., e-mail: vadimptpv@gmail.com

*Received April 05, 2023, revised April 18, 2023, accepted April 24, 2023.***Abstract**

The purpose of this work is to investigate the possibility of using the concept of digital twins applied to the designing of environmental monitoring systems. Their classification is given, and the conclusion about possibility of application of a certain type of digital twins to systems of ecological monitoring representing a complex of means of gathering, structured storage, the subsequent processing and the analysis of the information streams providing data on a condition of environment is made. As a basic element for the design of the twin, it is proposed to use a digital factory, which is a direct analogue of digital shadows, and providing direct interaction with the environment. The basic tools realizing the digital factory are defined and the conclusion about possibility of application of the approach developed in work to design of object-oriented digital twin for ecological monitoring system is made. The proposed approach is implemented for the environmental monitoring system of coal industry enterprises.

Keywords: digital twins, digital factory, containerization, monitoring, information flows, orchestration.

Citation: Potapov V.P., Popov S.E., Schastlivtsev E.L. Digital twins: strategies and approaches to creating environmental monitoring systems. Computational Technologies. 2023; 28(3):167–181. DOI:10.25743/ICT.2023.28.3.010. (In Russ.)

Acknowledgements. This work arose and developed at the suggestion of Academician Yu.I. Shokin, and was under his close attention for many years, for which the authors of the work bring their gratitude to him.

References

1. **Crelin J.** Principles of information technology. Grey House Publishing, Inc.; 2020: 410. ISBN:9781642656954.
2. **Karimi H.A.** Big Data: techniques and technologies in geoinformatics. CRC Press; 2014: 306.
3. **Dow C.** Mastering IoT: build modern IoT solutions that secure and monitor your IoT infrastructure. Packt Publishing; 2019: 782.
4. **Ranjan R., Mitra K., Jayaraman P.P., Wang L., Zomaya A.Y.** Handbook of integration of cloud computing, cyber physical systems and Internet of Things. Springer; 2020: 331. ISBN:978-3-030-43794-7.
5. **Acharjya D.P., Mitra A., Zaman N.** Deep learning in data analytics: recent techniques, practices and applications. Springer; 2022: 271. ISBN:978-3-030-75854-7.
6. **Polkowski Z., Kumar M.S., Vasilev J.** Data science in engineering and management. CRC Press; 2022: 152. ISBN:978-1-032-10625-0.
7. **Banerjee J.S., Bhattacharyya S., Obaid A.J., Yeh W.-C.** Intelligent cyberphysical systems security for Industry 4.0: applications, challenges and management. CRC Press; 2022: 284. ISBN:978-1-032-14835-9.

8. **Pal S.K., Mishra D., Pal A., Dutta S., Chakravarty D., Pal S.** Digital twin — fundamental concepts to applications in advanced manufacturing. Springer; 2022: 495. ISBN:978-3-030-81814-2.
9. **Grieves M.W.** Digital twin: manufacturing excellence through virtual factory replication. LLC; 2014: 7. Available at: https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication.
10. **Tao F., Sui F., Liu A., Qi Q.** Digital twin-driven product design framework. International Journal of Production Research. 2019; (57):3935–3953. DOI:10.1080/00207543.2018.1443229. Available at: <https://www.tandfonline.com/doi/abs/10.1080/00207543.2018.1443229?journalCode=tpres20>.
11. **Datta S.** Emergence of digital twins. Journal of Innovation Management. 2017; (5):14–34.
12. **Miller P.** Digital twins combine enterprise data and IoT to drive new business value. Principal analyst. 2019. Available at: <https://go.forrester.com/blogs/digital-twin-iotnew-business-value>.
13. **Mottura S., Viganó G., Greci L., Sacco M., Carpanzano E.** New challenges in collaborative virtual factory design. International Conference on Information Technology for Balanced Automation Systems. Springer; 2008: 17–24. Available at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1088.9509&rep=rep1&type=pdf>.
14. **Ducree J.** Digital twin: an oracle for efficient crowdsourcing of research and technology development through blockchain. Oracle Corporation Preprint. 2021: 15. DOI:10.20944/preprints202110.0148.v1. Available at: <https://www.preprints.org/manuscript/202110.0148/v1>.
15. The promise of a digital twin strategy. Best practices for designers and manufacturers of products and industrial equipment. Microsoft Services. 2022: 23. Available at: <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE1IMIi>.
16. **Kukushkin K., Ryabov Yu., Borovkov A.** Digital twins: a systematic literature review based on data analysis and topic modeling. MDPI, Data. 2022; 7(12):21.
17. **Unhelkar B., Gonsalves T.** Artificial intelligence for business optimization: research and applications. CRC Press; 2021: 325. ISBN:978-1-032-02886-6.
18. **Prokhorov A., Lysachev M.** Tsifrovoy dvoynik. Analiz, trendy, mirovoy opyt [Digital twin. Analysis, trends, world experience] / A. Borovkov (ed.). Moscow: OOO “Al’yans Print”; 2020: 401. (In Russ.)
19. Infrastruktura prostranstvennykh dannyykh. Trebovaniya k informatsionnomu obespecheniyu [Spatial data infrastructure Requirements for information support]. GOST R 58571. 2019: 17. (In Russ.)
20. **Maryse C.** Digital organizations manufacturing: scripts, performativity and semiopolitics. Wiley; 2018: 318. ISBN:978-1-119-52766-4.
21. **Dolgui A., Bernard A., Lemoine D., von Cieminski G., Romero D.** Advances in production management systems. Artificial intelligence for sustainable and resilient production systems. IFIP WG 5.7 International Conference, APMS 2021, Nantes, France, September 5–9, 2021. Proceedings, pt II. Springer; 2021: 730. ISBN:978-3-030-85901-5. Available at: <https://link.springer.com/book/10.1007/978-3-030-85902-2>.
22. **Sharma S.** The DevOps adoption playbook: a guide to adopting DevOps in a multi-speed IT enterprise. John Wiley & Sons; 2017: 416.
23. **Deepak V.** Kubernetes microservices with Docker. Apress; 2016: 456.
24. **Alok K.** Practical full stack machine learning: a guide to build reliable, reusable, and production-ready full stack ML solutions. BPB Publications; 2022: 751. ISBN:978-93-91030-42-1.
25. **Crickard P.** Data engineering with Python. Packt Publishing; 2020: 357.