УДК 004:69.003:658.011.8 DOI: 10.33622/0869-7019.2020.07.46-51

Хранилища данных в строительной отрасли: принципы построения, области использования

Александр Ильич КОНИКОВ, кандидат технических наук, доцент, e-mail: KonikovAI@mgsu.ru

ФГБОУ ВО «Национальный исследовательский Московский государственный строительный университет» (НИУ МГСУ), 129337 Москва, Ярославское ш., 26

Аннотация. *Хранилища данных входят в состав передовых информационных технологий, основанных на интеллектуальной обработке информации, используются во многих отраслях экономики, в том числе и в строительной отрасли. Применение информационных технологий на основе хранилища данных позволяет повысить качество выполнения работ на этапах проектирования, строительства и эксплуатации объектов, причем на каждом из этапов имеется своя специфика использования хранилища данных. На их основе создаются самые передовые цифровые технологии: интеллектуальная обработка данных Data Mining, технология «Большие данные» и др. Облачные технологии открыли новые подходы к использованию хранилища данных. С помощью облачных служб SaaS, PaaS и других удалось в значительной мере устранить главный недостаток, свойственный традиционным хранилищам данных, — их высокую стоимость. Поэтому облачные технологии хранения и обработки информации могут позволить себе иметь и небольшие предприятия строительной отрасли, когда в силу экономических причин использовать традиционные информационные технологии нецелесообразно. В статье исследуются возможности применения хранилища данных в строительной отрасли и новые горизонты, которые открываются в связи с внедрением облачных технологий.*

Ключевые слова: *строительство, хранилище данных, облачные технологии, Big Data, Data Mining, системы поддержки принятия решений, SaaS, PaaS.*

DATA WAREHOUSE IN THE CONSTRUCTION: PRINCIPLES OF CONSTRUCTION, AREAS OF USE

Alexander I. KONIKOV, e-mail: KonikovAI@mgsu.ru

Moscow State University of Civil Engineering (National Research University), Yaroslavskoe shosse, 26, Moscow 129337, Russian Federation

Abstract. *Data warehouses are part of advanced information technologies based on intelligent information processing and are used in many sectors of the economy, including the construction industry. The use of information technologies based on the data warehouse makes it possible to improve the quality of work at the stages of design, construction and operation of objects, and each of the stages has its own specifics of using the data warehouse. The most advanced digital technologies are created on the basis of the data warehouse: OLAP-analysis, Data Mining, Big Data, etc. Cloud technologies have opened new approaches to the use of data warehousing, the use of cloud services SaaS, PaaS, etc. makes it possible to overcome the main drawback inherent in traditional data warehouses — high cost. Therefore, cloud technologies of storage and information processing can be used at small enterprises of the construction industry — where, for economic reasons, the use of traditional DW is impractical. The article explores the possibilities of using a data warehouse in the construction industry and new horizons that are opened in connection with the introduction of cloud technologies.*

Key words: *construction, data warehouse, cloud technologies, Big Data, Data Mining, DSS, SaaS, PaaS.*

Информационные технологии, использующие хранилище данных (Data Warehouse — DW), эффективно применяются во многих отраслях экономики, в частности в строительной отрасли [1]. На их основе строятся многие современные информационные системы: интеллектуальная обработка данных (Data Mining), системы поддержки принятия решений, обработка больших данных (Big Data) и др. Применение этих систем позволяет вывести на новый качественный уровень многие бизнес-процессы в проектных и строительных организациях, повысить основные экономические и технические показатели.

Концепция «Хранилище данных» — DW изначально была хорошо продумана, ее появлению предшествовал многолетний и плодотворный этап исследований в области систем управления базами данных, поддержки принятия решений и др. Существует ряд «классических» принципов хранения информации в DW:

- *проблемно-предметная ориентация.* Данные объединяются, хранятся в соответствии с областями, которые они описывают, а не с приложениями, которые они используют;
- *интегрированность.* Данные объединены так, чтобы они удовлетворяли всем требованиям предприятия в целом, а не единственной функции бизнеса;
- *некорректируемость.* Дан-

© КОНИКОВ А. И., 2020

ные в хранилище данных не создаются, т. е. поступают из внешних источников, не корректируются и не удаляются;

• *зависимость от времени.* Данные в хранилище точны и корректны только тогда, когда они привязаны к промежутку или моменту времени.

Внедрение системы DW — проект дорогостоящий, который могут позволить себе крупные строительные организации, и тогда на его базе можно развернуть информационные системы, существенно повышающие качество бизнес-процессов. К таким системам относятся:

• поддержка принятия решений — СППР;

• оперативный многомерный OLAP-анализ;

• углубленный интеллектуальный анализ Data Mining;

• быстрый анализ больших объемов разнотипных данных Big Data и др.

В строительной отрасли характер использования DW в значительной мере зависит от конкретного этапа — проектирования, строительства, эксплуатации *(рис.1).*

Рис. 1. ***Использование систем DW на этапах проектирования, строительства, эксплуатации объектов строительства***

На этапе **проектирования** крупного объекта строительства требуется хранить в DW в упорядоченном виде огромные объемы разнотипной информации: технической, экономической, нормативной и т. д. При этом на базе DW возможно организовать функционирование следующих систем.

Система поддержки принятия решений — СППР (Decision Support System — DSS). Цель этой автоматизированной системы — помочь руководству принимать решения в сложных условиях. Для выбора решений используют интеллектуальный анализ данных, поиск знаний в базах данных, нейронные сети и другие современные подходы.

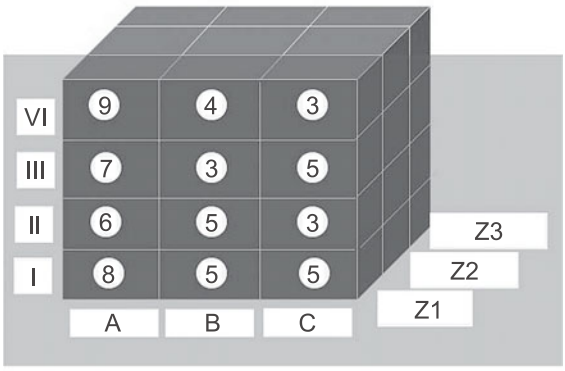
Экспертные системы — компьютерная система, способная в определенном смысле заменить специалиста-эксперта в разрешении проблемной ситуации. Современные экспертные системы зачастую строятся в рамках концепции искусственного интеллекта. Важнейшая часть экспертной системы — база знаний, т. е. совокупность фактов и правил логического вывода в выбранной предметной области деятельности.

Углубленный анализ данных (Data Mining). Согласно определению, цитируемому в работе [2], Data Mining — «набор методов и алгоритмов для автоматического извлечения из сырых данных потенциально полезных закономерностей, которые были бы не тривиальны и до этого не известны» [2]. Сырая информация для добычи данных — это, в нашем случае, базы данных с рабочей информацией строительной компании. Математический аппарат основан на применении деревьев решений, искусственных нейронных сетей, нечеткой логики и др. Информация, найденная в процессе применения методов Data Mining, должна быть нетривиальной и ранее неизвестной (например, информация «средний расход цемента» этому условию не удовлетворяет). Полученные знания должны быть применимы и на новых данных с некоторой степенью достоверности. Например, если выявлена закономерность: «стоимость строительных работ ниже, если соблюдается ряд конкретных условий» — это знание может принести экономическую пользу. Важнейшее условие успешного использования Data Mining состоит в наглядном пред-

ставлении результатов, что позволяет работать с Data Mining сотрудникам без специальной математической подготовки.

Технологии Big Data [3–9]. Данное направление относительно новое и основано на концепции «Хранилище данных», интеллектуальном анализе данных. В определенном смысле оно развивает идеи и методы технологии Data Mining. Для краткой и наглядной характеристики технологии «Большие данные» традиционно выделяют «три V»: объем (volume) — обрабатываются большие объемы данных; скорость (velocity) — высокоскоростная обработка и оперативное получение результатов; многообразие (variety) — возможность обработки различных типов структурированных и слабоструктурированных данных. Наряду с методами класса Data Mining (кластерный анализ, регрессионный анализ и др.) в Big Data используется ряд новых подходов. К ним можно отнести: краудсорсинг — привлечение к решению проблем широкого круга лиц на добровольных началах; смешение и интеграцию данных — набор приемов, которые дают возможность интегрировать разнородные данные из разных источников для глубинного анализа и пр. Внедрение новых подходов и позволило выделить Big Data в качестве отдельного направления.

Ситуационный центр представляет собой отдельное помещение, оснащенное устройствами визуализации (набор мониторов, в том числе с большим экраном), современными средствами телекоммуникаций (быстрый интернет, устройства для проведения видеоконференций, проводные и беспроводные каналы связи и т. п.) и интеллектуальными средствами обработки информации (СППР, экспертные системы, Big Data и др.) [5].



A — высотные здания; B — здания средней этажности; C — малоэтажные здания

Рис. 2. ***Пример использования куба данных для OLAP-анализа***

На этапе **строительства** крупного объекта также требуется хранить и обрабатывать большие объемы информации, связанные непосредственно со строительным производством. При принятии решений в сложных ситуациях могут потребоваться СППР и экспертные системы, о которых говорилось выше. Эффективное использование этих систем возможно, если они позволяют получать результат достаточно быстро. В плане оперативности наиболее логичное решение — применение OLAP-технологии, или оперативной аналитической обработки данных, представленных в виде многомерного куба. Принцип OLAP-анализа упрощенно можно пояснить на конкретном примере *(рис. 2).*

По оси *X* указан тип здания (A–C), по оси *Y* — номер квартала (I–IV), по оси *Z* — застройщик (*Z1–Z3*). В клетках куба показана информация о вложенных средствах (допустим, в млн р.). Проведя сечение куба по оси *Z*, можно получить и проанализировать данные, например, для застройщика (*Z1*). Выполняя сечение по разным плоскостям и различным параметрам, можно проводить быстрый анализ ситуации.

Для крупных строек зачастую требуется проанализировать в короткий срок огромное количество разнотипных характеристик и параметров. Здесь может оказаться полезной обработка данных по технологии Big Data. Для очень крупных объектов строительства может быть оправданно решение «Ситуационный центр». Несколько «нивелировать» затраты на данное решение можно при многократном использовании оборудования такого центра для других строек.

Разумеется, на этапе строительства DW могут применять для более простых задач: например, по соответствующим запросам быстро получать нужную информацию, однако только в этом качестве использование DW экономически нецелесообразно.

На этапе **эксплуатации** крупного объекта строительства также имеется большой объем разнотипной информации — показания датчиков состояния здания (с привязкой параметров ко времени), данные, касающиеся технической, нормативной документации и т. д. Существуют предпосылки для использования:

- систем OLAP-анализа, которые позволяют оперативно получить развернутую информацию по эксплуатируемому объекту, сгруппированную по различным признакам;
- систем поддержки принятия решений, например, когда следует решить вопрос о капитальном ремонте.

Для крупных строительных сооружений может идти речь об использовании технологий Data Mining и Big Data в случае, если следует проанализировать хара-

ктер и закономерности выхода из строя оборудования.

Есть еще один важный момент для использования DW на этапе эксплуатации: хранимая информация о состоянии объекта сама по себе ценная. Эта информация может быть полезна для проектировщиков новых объектов строительства, причем ценность информации, по мере пополнения данными, возрастает.

Применение многих компьютерных систем, в том числе систем на базе DW, зачастую сдерживали экономические факторы. Новый импульс для внедрения компьютерных систем в практику строительных организаций получен в результате появления облачных технологий. Действительно, эти технологии, помимо прочего, позволяют решить проблему материальных затрат на внедрение и эксплуатацию системы, что особенно важно для небольших строительных организаций.

Рассмотрим ряд общих моментов, касающихся облачных технологий. Термин «облачные технологии» служит в качестве метафоры, когда от пользователя скрыта, т. е. закрыта облаком, сложная инфраструктура сетевого доступа к вычислительным ресурсам.

Потребители облачных вычислений могут значительно сократить расходы на инфраструктуру информационных технологий и гибко реагировать на изменения вычислительных потребностей.

Для объяснения экономической составляющей облачных подходов уместно провести аналогию с услугами водо- или электроснабжения. В развитых инфраструктурах услуги предоставляются по соответствующим коммунальным сетям и оплачиваются по мере потребления. Это гораздо выгодней и удобней, чем разработка каждым потребителем собственного водозабора или монтирование своей электроустановки.

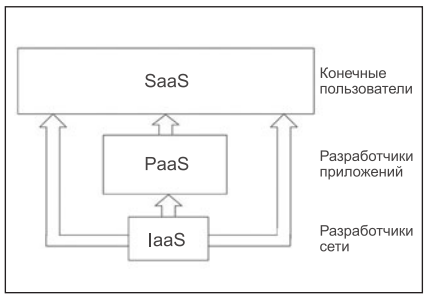


Рис. 3. ***Облачные службы: SaaS, PaaS и IaaS***

Важное преимущество «облачного» подхода — использование новых версий программных продуктов, возможность доступа с любого устройства, подключенного к интернету, и организации совместной работы с данными и др. К основным недостаткам относят проблему безопасности при работе с «облаком», особенно в отношении конфиденциальных и приватных данных. Иногда для строительных организаций этот вопрос весьма актуален.

Определены обязательные характеристики облачных вычислений:

- самообслуживание по требованию — потребитель сам определяет свои вычислительные потребности: серверное время, скорости доступа и пр.;
- объединение ресурсов — поставщик услуг объединяет ресурсы для обслуживания большего числа потребителей и перераспределения мощностей;
- эластичность — услуги могут быть расширены, сужены и т. д.

В настоящее время есть несколько моделей облачных вычислений [10, 11].

Программное обеспечение как услуга (SaaS). Модель, в которой потребителю предоставляется возможность использования прикладного программного обеспечения (ППО) провайдера, работающего в облачной инфраструктуре. ППО доступно посредством различных клиентских устройств или через интерфейсы тонких клиентов, например, веб-браузер.

Платформа как услуга (PaaS). Потребитель получает возможность использовать облачную инфраструктуру для базового программного обеспечения и размещения на нем собственных или существующих приложений.

Инфраструктура как услуга (IaaS). Имеется возможность с помощью облачной инфраструктуры самостоятельно управлять набором физических ресурсов, таких как серверы, сетевое оборудование, накопители и т. п.

Есть и другие службы, например, рабочее место как услуга WaaS — позволяет компании применять облачные вычисления для организации рабочих мест своих сотрудников; данные как услуга DaaS — предоставляют пользователю дисковое пространство для хранения больших объемов информации.

Наибольший интерес представляют услуги SaaS, PaaS и IaaS *(рис. 3).* Они предназначены для различных групп пользователей облачных услуг: конечные пользователи — SaaS; разработчики приложений — PaaS; разработчики сети — IaaS.

Теперь перейдем непосредственно к вопросу «хранилище данных» на основе облачных технологий. Здесь возможно решение тех же задач, что и в «обычных» DW (интеллектуальный анализ Data Mining и др.). Однако выбирают иные подходы к решению — использование облачных служб SaaS, PaaS и т. д. В само понятие «хранилище данных» вкладывается другое содержание. Облачное хранилище данных (cloud storage) представляет собой модель онлайн-хранилища, в котором данные хранятся на распределенных в сети серверах, предоставляемых в пользование клиентам третьей стороной. В отличие от модели хранения данных на собственных серверах внутренняя структура

клиенту не видна, закрыта «облаком».

Облачными хранилищами стали такие известные интернет-сервисы, как Dropbox, Google Drive, Яндекс.Диск, Облако Mail.Ru и др.

Благодаря отмеченным выше достоинствам облачные хранилища данных могут весьма успешно применяться в строительной отрасли:

• при хранении больших объемов данных. Хранимая информация может относиться к различным областям: характеристики датчиков состояния здания (с временной привязкой), нормативные документы и т. д.;

• при интеллектуальном анализе данных (Data Mining и пр.) в разных строительных организациях, включая и небольшие. Ранее в силу экономических причин это было невозможно;

• при обработке данных, необходимых для OLAP-анализа (используется для решения оперативных задач на этапе строительства и эксплуатации).

По мере развития облачных технологий, приведенный выше список будет дополняться. Несомненно, что грамотное применение облачных технологий хранения и обработки информации во многом определяет эффективность работы современного предприятия, в том числе строительной отрасли.

Вывод

В работе исследуется одно из фундаментальных понятий в области информационных технологий — хранилище данных (DW). Рассмотрение проводится в контексте возможного использования DW в строительной отрасли. Даны принципы построения DW. Указано, что применение информационных технологий на основе хранилища данных позволяет повысить качество выполнения работ на этапах проектирования, строительства и эксплуатации объектов. Отмечено, что на базе DW возможно внедрение самых современных цифровых технологий: систем поддержки принятия решений (СППР), интеллектуальный углубленный анализ данных Data Mining, анализ «большие данные» Big Data, оперативный анализ данных OLAP и др. Дан краткий анализ этих систем, рассмотрены области рационального применения. Подчеркнуто, что на каждом из этапов строительного производства имеется своя специфика использования DW, и, соответственно, применяются свои информационные системы.

Отмечено, что облачные технологии открыли новые возможности использования хранилища данных и в строительстве. Появилась возможность в значительной мере преодолеть проблемы экономического характера, что позволило внедрять хранилища данных не только в крупных, но и в небольших строительных организациях.

Показано, что при облачной реализации изменился смысл самого понятия «хранилище данных». Уточнен смысл этого понятия, указаны обязательные характеристики облачных вычислений (самообслуживание по требованию, эластичность и др.). Даны модели облачных вычислений — программное обеспечение как услуга SaaS, платформа как услуга PaaS и т. д., приведены группы пользователей для каждой из облачных услуг.

ЛИТЕРАТУРА

1. *Konikov A., Kulikova E., Stifeeva O.* Research of the possibilities of application of the Data Warehouse in the construction area [Исследование возможностей использования хранилища данных в строительстве] // MATEC Web of Conferences. 2018. Vol. 251. P. 03062.
2. *Дюк В. А., Флегонтов А. В., Фомина И. К.* Применение технологий интеллектуального анализа данных в естественнонаучных, технических и гуманитарных областях. URL: https://cyberleninka.ru/article/n/primenenie-tehnologiy-intellektualnogo-analiza-dannyh-v-estestvennonauchnyh-tehnicheskih-i-gumanitarnyh-oblastyah/viewer (дата обращения: 6.02.2020).
3. *Konikov A., Konikov G.* Big Data is a powerful tool for improving the environment in the construction business [Big Data – мощный инструмент для улучшения окружающей среды в строительном бизнесе] // IOP Conference. Series: Earth and Environmental Science. 2017. Vol. 90. P. 012184.
4. *Коников А. И.* Перспективные направления в области информационных систем управления строительством // Промышленное и гражданское строительство. 2019. № 6. С. 64–69.
5. *Коников А. И.* Ситуационный центр управления эксплуатацией зданий // Промышленное и гражданское строительство. 2018. № 7. С. 84–87.
6. *Ivanov N., Gnevanov M.* Big data: perspectives of using in urban planning and management [Большие данные: перспективы использования в городском планировании и управлении] // MATEC Web of Conferences. 2018. Vol. 170. P. 01107.
7. *Гневанов М. В., Иванов Н. А* Технологии «больших данных» (Big Data) и их применение в градостроительном планировании // Промышленное и гражданское строительство. 2018. № 4. С. 83–87.
8. *Волынсков В. Э.* «Большие данные» (BIG DATA) в градостроительстве // Academia. Архитектура и строительство. 2017. № 2. С. 29–32.
9. *Kitchin R.* The data revolution. Big Data, Open Data, Data infrastructures and their consequences [Революция данных. Большие данные, открытые данные, ин-

фраструктура данных и их последствия]. Londres, SAGE Publ., 2014. 240 p.

10. *Межуева Т. А.* Постановка задач облачных вычислений // Вестник МГСУ. 2011. № 6. С. 444–447.

11. *Холод И. И.* Архитектура «облака» интеллектуального анализа данных на основе библиотеки алгоритмов с блочной структурой // Известия СПбГЭТУ «ЛЭТИ». 2014. № 6. С. 34–40.

R E F E R E N C E S

1. Konikov A., Kulikova E., Stifeeva O. Research of the possibilities of application of the Data Warehouse in the construction area. *MATEC Web of Conferences*, 2018, vol. 251, p. 03062.
2. Duk V. A., Flegontov A. V., Fomina I. K. *Primenenie tehnologiy intellektualnogo analiza dannyh v estestvennonauchnyh, tehnicheskih I gumanitarnyh oblastyah* [Application of data mining technologies in natural-science, technical and humanitarian fields]. Available at: https://cyberleninka.ru/article/n/primenenie-tehnologiy-intellektualnogo-analiza-dannyh-v-estestvennonauchnyh-tehnicheskih-i-gumanitarnyh-oblastyah/viewer (accessed 6.02.2020). (In Russian).
3. Konikov A., Konikov G. Big Data is a powerful tool for improving the environment in the construction business. *IOP Conference Series: Earth and Environmental Science*, 2017, vol. 90, p. 012184.
4. Konikov A. I. Promising areas in the field of information systems of construction management. *Promyshlennoye i grazhdanskoye stroitel'stvo*, 2019, no. 6, pp. 64–69. DOI: 10.33622/0869-7019.2019.06.64-69. (In Russian).
5. Konikov A. I. Situational control center of buildings operation. *Promyshlennoe i grazhdanskoe stroitel'stvo*, 2018, no. 7, pp. 84–87. (In Russian).
6. Ivanov N., Gnevanov M. Big data: perspectives of using in urban planning and management. *MATEC Web of Conferences*, 2018, vol. 170, p. 01107.
7. Gnevanov M. V., Ivanov N. A. Application of technologies "Big Data" in urban planning. *Promyshlennoe i grazhdanskoe stroitel'stvo*, 2018, no. 4, pp. 83–87. (In Russian).
8. Volynskov V. E. Big Data in urban planning. *Academia. Arhitektura i stroitelstvo*, 2017, no. 2, pp. 29–32. (In Russian).
9. Kitchin R. *The data revolution. Big Data, Open Data, Data infrastructures and their consequences.* Londres, SAGE Publ., 2014. 240 p.
10. Mezhueva T. A. Statement of cloud computing problems. *Vestnik MGSU*, 2011, no. 6, pp. 444–447. (In Russian).
11. Holod I. I. The architecture of the cloud of data mining based on a library of algorithms with a block structure. *Izvestiya SPbGETU "LETI"*, 2014, no. 6, pp. 34–40. (In Russian).

Для цитирования: *Коников А. И.* Хранилища данных в строительной отрасли: принципы построения, области использования // Промышленное и гражданское строительство. 2020. № 7. С. 46–51. DOI: 10.33622/0869-7019.2020.07.46-51.

For citation: Konikov A. I. Data Warehouse in the Construction: Principles of Construction, Areas of Use. *Promyshlennoe i grazhdanskoe stroitel'stvo* [Industrial and Civil Engineering], 2020, no. 7, pp. 46–51. (In Russian). DOI: 10.33622/0869-7019.2020.07.46-51.