



ПРАКТИКА ПОСТРОЕНИЯ КОРПОРАТИВНОГО ХРАНИЛИЩА ДАННЫХ В НЕФТЕГАЗОВОМ СЕКТОРЕ

© И.Р. Шеховцова,
2025



И.Р. Шеховцова

Международная компания ООО «Газпром Интернэшнл Лимитед», РФ, Санкт-Петербург

Электронный адрес: I.Shekhovtsova@gazprom-international.com

Введение. Эффективное принятие управленческих решений во многом зависит от наличия оперативного доступа к актуальной информации о состоянии активов. Данные о таких бизнес-структурах, как правило, фрагментированы и взаимозависимы, что требует системного подхода к их консолидации и интерпретации.

Цель. В статье рассматривается опыт построения корпоративного хранилища данных на основе разрозненных данных операционных систем-источников для возможности комплексного анализа состояния активов.

Материалы и методы. Для достижения цели определен перечень ключевых показателей компании на разных уровнях управления, проанализированы достоинства и недостатки архитектурных подходов Уильяма Инмона и Ральфа Кимбола к построению корпоративного хранилища данных, исследованы трудности по устранению несоответствий в моделях систем-источников и обеспечению заданного уровня информационной безопасности в компонентах архитектуры.

Результаты. Приведен реальный пример построения гибридной архитектуры корпоративного хранилища данных на основе подходов Уильяма Инмона и Ральфа Кимбола, настроены ETL-процессы по загрузке в хранилище и насыщению витрин данных, разработан модуль загрузки отсутствующих в системах данных, проведены мониторинг и анализ рисков архитектуры хранилища.

Заключение. Построение корпоративного хранилища данных действительно сопряжено со множеством вызовов в области проработки архитектуры и настройки ETL-процессов. Реальные примеры интеграции, в том числе приведенный в настоящей статье, служат ценным ориентиром для понимания реализованных стратегий, показывают практическое применение интеграционных фреймворков, технологий и методологий, а также предлагают информацию о приобретенных уроках и ключевых факторах, способствовавших успешным проектам интеграции.

Ключевые слова: корпоративное хранилище данных, интеграция данных, витрины данных, ETL, дашборд, риск.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Для цитирования: Шеховцова И.Р. Практика построения корпоративного хранилища данных в нефтегазовом секторе. ПРОНЕФТЬ. Профессионально о нефти. 2025;10(1):98–107. <https://doi.org/10.51890/2587-7399-2025-10-1-98-107>

Статья поступила в редакцию 02.12.2024

Принята к публикации 24.01.2025

Опубликована 31.03.2025

PRACTICE OF BUILDING A CORPORATE DATA WAREHOUSE IN THE OIL AND GAS SECTOR

Irina R. Shekhovtsova

Gazprom International Limited ILLC, Saint Petersburg, RF

E-mail: I.Shekhovtsova@gazprom-international.com

Introduction. Effective managerial decision-making largely depends on having timely access to current information about asset status. Data on such business structures are typically fragmented and interdependent, necessitating a systematic approach to their consolidation and interpretation.

Aim. The article examines the experience of building a corporate data warehouse based on disparate data from operational source systems to enable comprehensive analysis of asset status.

Materials and methods. To achieve the goal, a list of key company performance indicators at different management levels was determined, the advantages and disadvantages of the architectural approaches of William Inmon and Ralph Kimball to building a corporate data warehouse were analyzed, and the difficulties in eliminating inconsistencies in source system models and ensuring the required level of information security in architecture components were investigated.

Results. A real example of constructing a hybrid architecture for a corporate data warehouse based on the approaches of William Inmon and Ralph Kimball is provided. ETL processes for loading data into the warehouse and populating data marts have been configured. A module for loading data absent in the systems has been developed. Monitoring and risk analysis of the constructed architecture have been conducted.

Conclusion. Building a corporate data warehouse indeed presents numerous challenges in architecture development and ETL process configuration. Real-world integration examples, including the one provided in this article, serve as valuable benchmarks for understanding implemented strategies. They demonstrate the practical application of integration frameworks, technologies, and methodologies, and offer insights into lessons learned and key factors that contributed to successful integration projects.

Keywords: corporate data warehouse, data integration, data marts, ETL, dashboard, risk

Conflict of Interest: the author declares no conflict of interest.

For citation: Shekhovtsova I.R. Practice of building a corporate data warehouse in the oil and gas sector. PRONEFT. Professionally about oil. 2025;10(1):98–107. <https://doi.org/10.51890/2587-7399-2025-10-1-98-107>

Manuscript received 02.12.2024

Accepted 24.01.2025

Published 31.03.2025

ВВЕДЕНИЕ

В современной деятельности нефтегазовой компании эффективное принятие управленческих решений во многом зависит от наличия оперативного доступа к актуальной информации о состоянии активов. Данные о таких бизнес-структурах, как правило, фрагментированы ввиду своей многоаспектности (ресурсно-сырьевой, производственно-технологический, экономический и прочие аспекты) и взаимозависимы, что требует системного подхода к их консолидации и интерпретации.

С развитием концепции Business Intelligence (BI) требования к представляемой информации и ее источникам непрерывно возрастают. Во-первых, традиционные способы предоставления данных (например, отчеты в Excel) становятся менее эффективными по сравнению с интерактивными панелями (дашбордами). Во-вторых, информация, на основе которой строится аналитика, должна представлять собой «единую версию истины» (single version of the truth), исключая внутренне разногласия между подразделениями компании [1]. Использование несовместимых данных может привести к неоптимальным управленческим решениям и, как следствие, к финансовым и операционным потерям.

В ответ на эти вызовы всё большее значение приобретает внедрение корпоративного хранилища данных (КХД), который стал одним из центральных компонентов современной BI-инфраструктуры.

ВЫЗОВЫ НА ПУТИ К УСПЕШНОМУ ВНЕДРЕНИЮ КОРПОРАТИВНОГО ХРАНИЛИЩА ДАННЫХ

Внедрение КХД является нетривиальной задачей, сопряженной со множеством вызовов и требующей комплексного подхода к глубокой проработке архитектуры и настройке процессов извлечения, преобразования и загрузки данных (ETL, от англ. «extract, transform, load» — дословно «извлечение,

преобразование, загрузка») из операционных систем, выступающих в качестве источников данных.

В этом контексте операционные системы понимаются как системы обработки транзакций (OLTP, от англ. «online transaction processing systems» — дословно «обработка транзакций в реальном времени»), поддерживающих повседневные бизнес-процессы компании. Вследствие разнообразия структур данных, форматов и протоколов коммуникации каждой такой системы возникает необходимость устранения несоответствий при записи в хранилище. Проблема интеграции данных усугубляется отсутствием универсальных решений для разрешения конфликтов данных [2]. Как показывают исследования [3, 4, 5], подходы, успешные в одном проекте, могут оказаться неэффективными в другом (речь идет о «собственных» реализациях ETL-процессов без использования специализированного программного обеспечения). Это делает задачу интеграции уникальной для каждого конкретного случая, требуя разработки адаптивных механизмов для устранения несоответствий.

ДЛЯ СОЗДАНИЯ КОРПОРАТИВНОГО ХРАНИЛИЩА ДАННЫХ НА ОСНОВЕ РАЗРОЗНЕННЫХ ДАННЫХ ОПЕРАЦИОННЫХ СИСТЕМ-ИСТОЧНИКОВ ПРИВЕДЕН ПРИМЕР ПОСТРОЕНИЯ ГИБРИДНОЙ АРХИТЕКТУРЫ ДАННЫХ НА ОСНОВЕ ПОДХОДОВ УИЛЬЯМА ИНМОНА И РАЛЬФА КИМБОЛА.

Исследования [6, 7, 8] выражают значительный интерес к проблемам обеспечения информационной безопасности (ИБ) КХД. Концептуально необходима проработка следующих аспектов:

- безопасность на уровне инфраструктуры и сетевых взаимодействий (обеспечение функционирования всех компонентов КХД в контуре доверенной среды компании);
- безопасность данных на уровне хранения (шифрование данных, резервное копирование);

- безопасность данных на уровне обработки и интеграции (контроль целостности, валидация входящих данных, согласованность форматов, управление метаданными);
- управление доступом и аутентификация на всех уровнях архитектуры КХД (модель управления доступом в соответствии с принципом минимальных прав, логирование и аудит действий пользователей);
- защита от внешних атак (защита от SQL-инъекций, исправление уязвимостей);
- анализ и управление рисками.

Предпринятые меры и стратегии позволят минимизировать риски, связанные с реализацией угроз нарушения ИБ КХД, в том числе зарегистрированных в Банке угроз Федеральной службы по техническому и экспортному контролю (ФСТЭК)¹.

АРХИТЕКТУРНЫЕ ПОДХОДЫ К ПОСТРОЕНИЮ КОРПОРАТИВНОГО ХРАНИЛИЩА ДАННЫХ

Глобально, существуют два фундаментальных подхода к проектированию корпоративных хранилищ данных, каждый из которых воплощает свою философию в отношении структуры и использования данных: архитектуры Уильяма Инмона (William H. Inmon) и Ральфа Кимбола (Ralph Kimball). Архитектура, предложенная Уильямом Инмоном (рис. 1), предполагает построение централизованного хранилища данных (enterprise data warehouse)

¹ Перечень угроз. Банк данных угроз безопасности информации Федеральной службы по техническому и экспортному контролю (ФСТЭК); 2024 [обновлено 23 сентября 2024; процитировано 24 сентября 2024]. Доступно: <https://bdu.fstec.ru/threat>

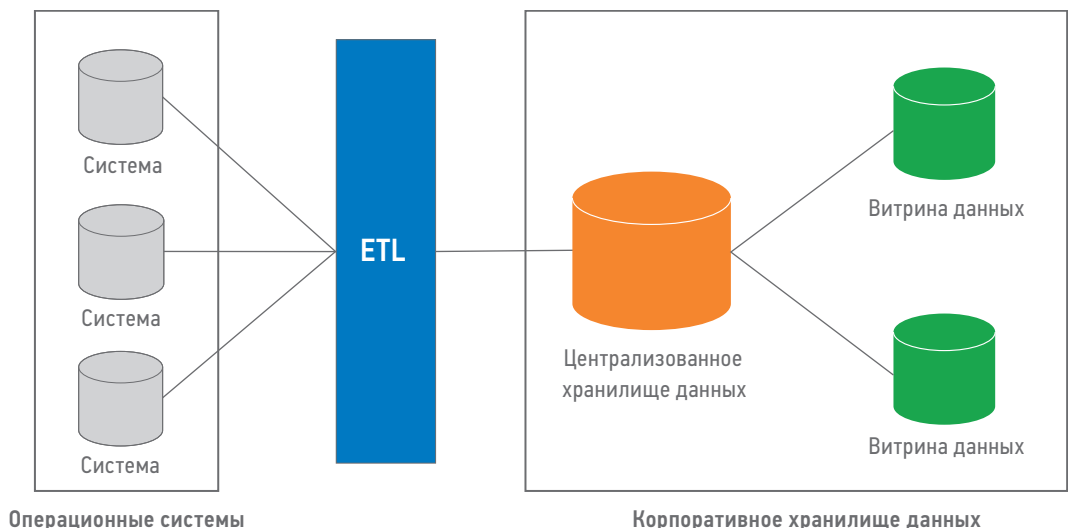


Рис. 1. Архитектура хранилища, предложенная Инмоном. Составлен автором
Fig. 1. Inmon's proposed data warehouse architecture. Prepared by the author

с нормализованной структурой, являющегося единой точкой сбора верифицированных данных из операционных систем, и последующее их распределение в витрины (data marts), ориентированные на запросы руководства и конкретных подразделений [9]. Очевидно, что по крайней мере часть КХД, и зачастую довольно значительная, должна быть создана до того, как можно будет предоставить витрины данных. Иной подход предлагает Ральф Кимбол (рис. 2), идея которого заключается в создании множества локализованных витрин данных (data marts), отвечающих аналитическим запросам отдельных подразделений [9]. В основе этой архитектуры лежит размерная модель (dimensional model), состоящая из таблиц фактов и измерений: таблицы фактов содержат релевантные метрики (чаще всего агрегированные), а таблицы измерений обеспечивают к ним контекст, предоставляя согласованные срезы. Подобные архитектуры могут привести к противоречию данных в таблицах фактов для разных витрин, хотя отличаются гибкостью и оперативностью в реализации. В рамках современных требований к обработке и анализу данных, выделенного бюджета и имеющихся ресурсов оба подхода могут рассматриваться как взаимодополняющие.

ОПЫТ ИНТЕГРАЦИИ, РЕАЛИЗОВАННЫЙ В КОМПАНИИ

Этап 1 — Анализ потребностей бизнеса и разработка концептуальной модели хранилища данных

Для определения объема данных, необходимого для поддержки принятия решений,

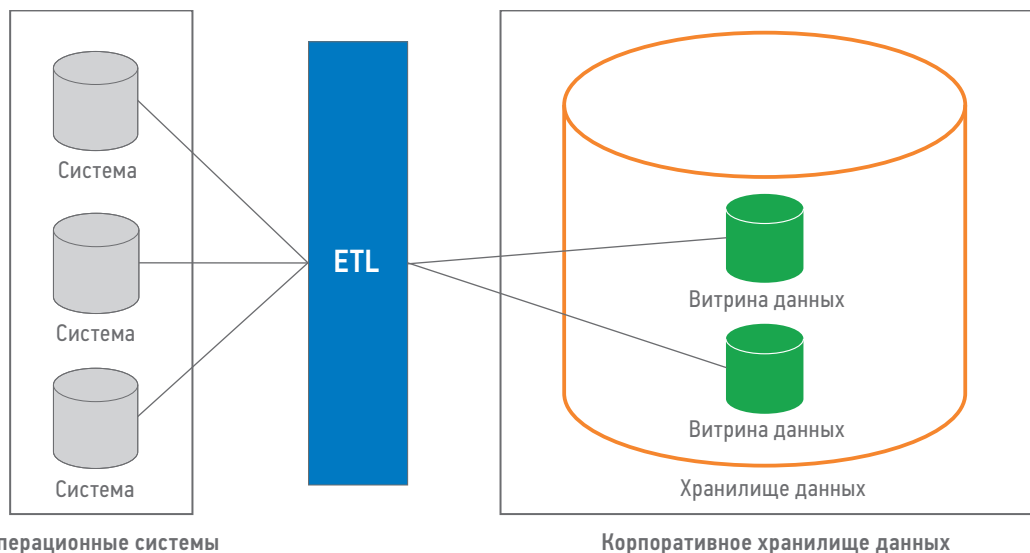


Рис. 2. Архитектура хранилища, предложенная Кимболом. Составлен автором
Fig. 2. Kimball's proposed data warehouse architecture. Prepared by the author

проведено интервью с менеджментом компании и сформирован перечень ключевых показателей на различных уровнях управления. Собранная информация позволила построить концептуальную модель архитектуры хранилища в виде таблицы и документально зафиксировать наименования показателей, частоту их обновления, уровни детализации, ответственные подразделения, источники, форму представления и описание.

При анализе модели были идентифицированы данные, отсутствующие в операционных системах, и запланированы процессы по доработке систем и дополнительному сбору данных.

Тщательная проработка концептуальной модели и вовлеченность в процесс руководства значительно снизили риски, связанные с несоответствием хранилища данных реальным потребностям, и во многом определили успешность будущего внедрения КХД.

Этап 2 — Выбор и реализация архитектуры корпоративного хранилища данных

При построении архитектуры КХД учитывались архитектурные подходы Уильяма Инмона и Ральфа Кимбола.

Отказ от исключительного следования Инмону связан с тем, что создание единого хранилища данных требует внушающих финансовых и временных ресурсов на всем этапе жизненного цикла КХД. Продолжительный период разработки и отсутствие ощутимых бизнесом результатов может привести к снижению доверия к проекту, сопротивлению со стороны пользователей и руководства и отрицательно сказаться на его поддержке и финансировании.

Решение не ограничиваться подходом Кимбола связано со сложностью управления множеством независимых локализованных витрин данных. При такой модели хранилища возрастает риск возникновения несоответствий между витринами из-за использования разнородных источников данных. Осознание ограничений, налагаемых строгим следованием одной методологии, привело к реализации в компании гибридного подхода, позволившего извлечь преимущества из каждой архитектурной парадигмы. В результате в компании была разработана архитектура КХД, представленная двумя уровнями — уровнем аналитического хранилища данных (АХД) и уровнем витрин данных (рис. 3).

С точки зрения представленной концепции, АХД является единым источником верифицированных данных и представляет собой несколько баз данных (БД) с нормализованной структурой, сгруппированных по целям аналитики. Ключевым архитектурным требованием к каждой БД является хранение атомарных данных в виде таблиц фактов, где уровень детализации определяется утвержденными таблицами измерений. На рис. 4 приведен пример структуры одной из БД хранилища.

В свою очередь, витрины создаются на основе данных, поставленных АХД, и оптимизируются для работы с конкретными BI-инструментами. На этом уровне применяется размерное моделирование, однако в отличие от уровня АХД данные в витринах денормализованы в целях ускорения аналитических запросов.

Реализованная в компании гибридная архитектура КХД сочетает в себе

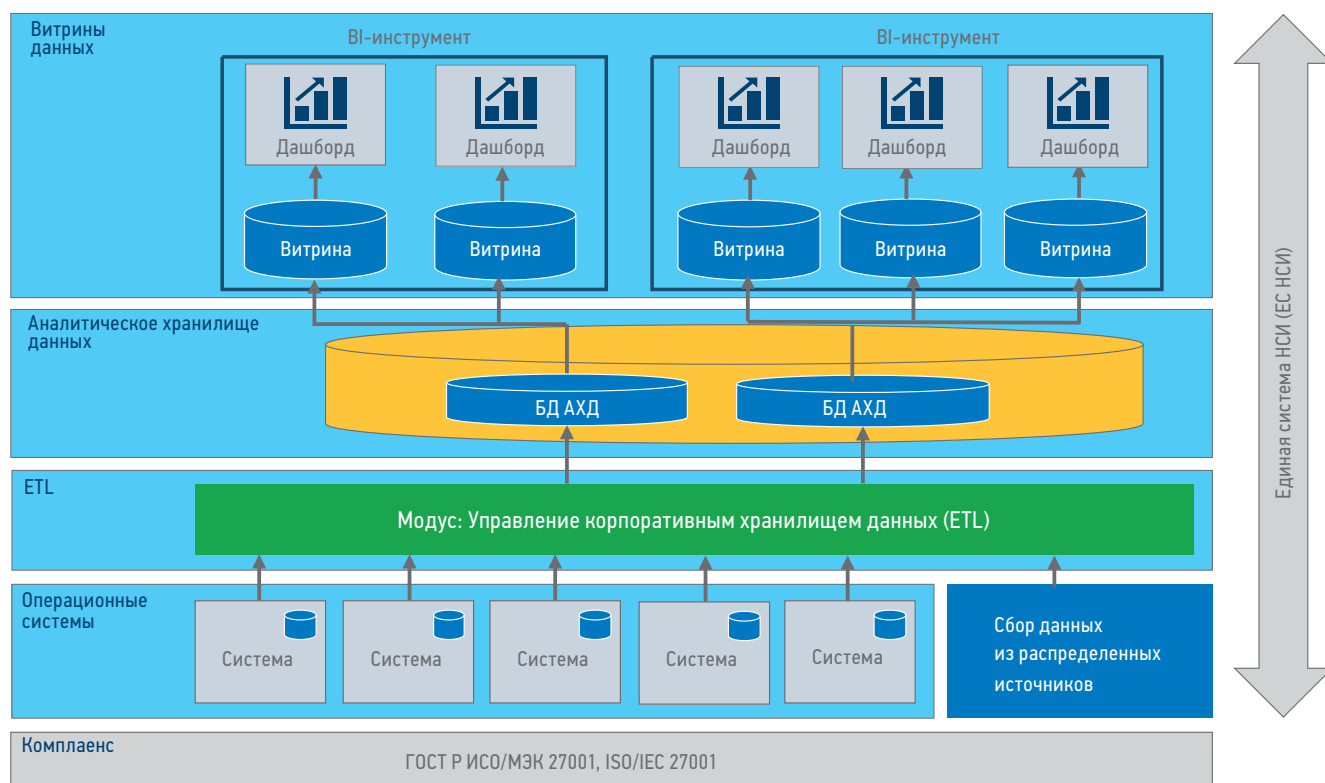


Рис. 3. Место корпоративного хранилища данных в ИТ-архитектуре компании. Составлен автором
Fig. 3. The place of a corporate data warehouse in a company's IT architecture. Prepared by the author

многоуровневость архитектуры от подхода Уильяма Инмона и принцип размерного моделирования от подхода Ральфа Кимбола, извлекая следующие преимущества:

- достижение баланса между целостностью и производительностью;
- возможность быстрого отклика на изменения в бизнес-требованиях без масштабных перестроек всей системы;
- адаптация под специфические требования различных BI-инструментов;
- единый способ хранения данных;
- возможность постепенного внедрения;
- высокая масштабируемость архитектуры.

Этап 3 — Стандартизация и нормализация данных

На этапе стандартизации и нормализации данных создана единая система нормативно-справочной информации (ЕС НСИ), которая служит основой для управления данными и обеспечивает распространение единственно верной и актуальной справочной информации из мастер-систем в системы-потребители.

В рамках реализации в компании ЕС НСИ на основе решения «1С:MDM Управление нормативно-справочной информацией» выполнены следующие задачи:

1. составлен список необходимых справочников, использующихся в бизнес-процессах;
2. назначены основные источники данных (мастер-системы) для каждого справочника;
3. внедрены глобальные уникальные идентификаторы (GUID) в операционных системах в качестве ссылок на записи в ЕС НСИ, обеспечивающих однозначную идентификацию элементов справочников;
4. настроены механизмы распространения данных ЕС НСИ из мастер-систем в системы-потребители через ETL-процессы;
5. обеспечено управление изменениями и версиями для своевременных обновлений в системах-потребителях.

Этап 4 — Реализация процессов по сбору отсутствующих данных

Для сбора отсутствующих в операционных системах данных разработан модуль, который поддерживает загрузку файлов в формате CSV (рис. 5).

CSV выбран ввиду простоты его структуры и скорости обработки скриптами. Значения вносимых данных проверяются на соответствие бизнес-логике и справочникам ЕС НСИ, что помогает выявлять и устранять несоответствия в данных до их загрузки в БД.

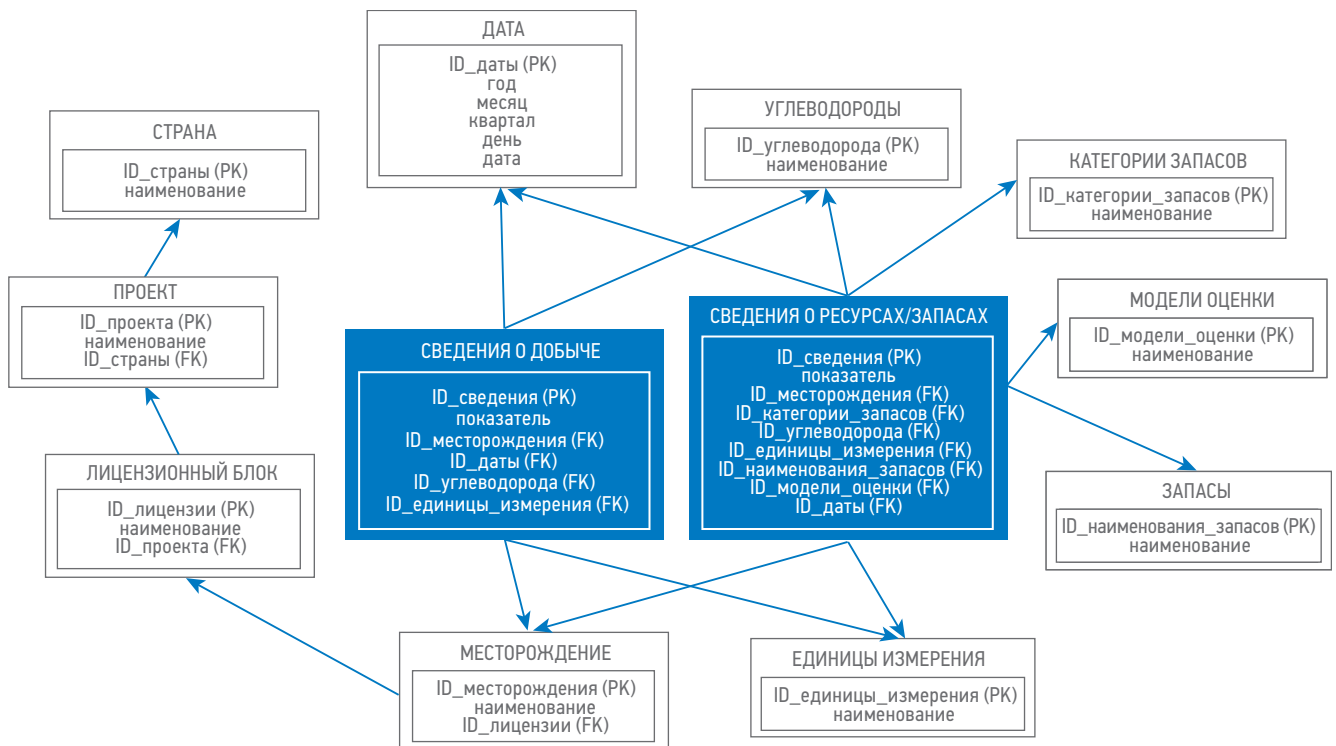


Рис. 4. Структура базы данных на PostgreSQL. Составлен автором
 Fig. 4. Database structure in PostgreSQL. Prepared by the author

Разграничение прав доступа к модулю между загрузчиками и администраторами данных усиливает контроль над вводимой информацией. Запись данных в БД производится только в случае согласования администратором данных.

Этап 5 — Выбор стека ETL-инструментов

В настройке ETL-процессов часто используется специализированное программное обеспечение в роли посредника между операционными системами и хранилищем [6]. В компании отдано предпочтение существующим решениям по сравнению с разработкой собственного, при выборе учтены следующие факторы:

- поддержка реализации сложных бизнес-правил и преобразований данных;

- поддержка различных форматов данных, реализуемых в компании;
- возможность интеграции механизма управления мастер-данными;
- поддержка планирования, логирования и мониторинга ETL-процессов;
- безопасность и управление доступом;
- интеграция с существующей ИТ-инфраструктурой компании;
- наличие понятной документации;
- доступность лицензий, стоимость владения, поддержка и развитие продукта.

С учетом всех аспектов выбран программный продукт «Модус: Управление корпоративным хранилищем данных (ETL)» (Modus ETL) на базе 1С, который позволяет выгружать данные из разных источников (учетных систем, баз данных, web-сервисов, файлов),

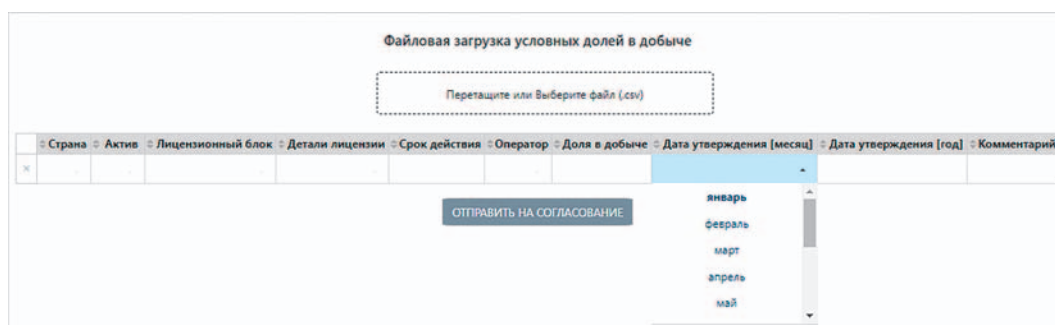


Рис. 5. Интерфейс модуля загрузки данных. Скриншот из программы ИАС «Отчетность и BI». Составлен автором
 Fig. 5. Data loading module interface. Screenshot from the program IAS «Reporting and BI». Prepared by the author

обрабатывать полученные данные по предварительно настроенным правилам и сохранять результаты в хранилище.

Этап 6 — Разработка ETL-процессов по загрузке данных в хранилище

В рамках единого подхода к организации ETL-процессов по загрузке данных в хранилище разработана концепция взаимодействия операционных систем с Modus ETL. Согласно ей, системы-источники сосредотачивают бизнес-правила и алгоритмы верификации на своей стороне, обеспечивая таким образом соответствие поставляемых данных определенным требованиям до передачи ETL-процессам. Извлечение данных осуществляется через представления (views) с доступом на чтение, что предотвращает искажение в операционных системах. На полученной выборке Modus ETL выполняет преобразования по приведению данных к формату, необходимому для АХД, в том числе связывает их с соответствующими записями в ЕС НСИ.

Концепция реализована в Modus ETL через установку правил выгрузки и активацию составов выгрузок. В системе определены источники данных, структуры выгружаемых таблиц, сопоставлены поля источников с целевыми полями АХД и реализованы преобразования типов данных для корректной загрузки. Пример

настройки правила выгрузки приведен на **рис. 6**.

Каждый день в установленное время Modus ETL автоматически инициирует процессы выгрузки из источников в соответствии с заданными параметрами, обеспечивая поддержку актуальности данных в АХД. Система отслеживает процессы загрузки данных, идентифицирует и регистрирует возникающие ошибки и предоставляет инструменты для их своевременного устранения, благодаря чему риск нарушения целостности данных в процессе их загрузки в АХД минимален. Стоит отметить, что более безопасный и стандартизированный способ интеграции данных между системами может быть реализован передачей данных в ETL-инструмент через программные интерфейсы (API, от англ. «application programming interface» — дословно «программный интерфейс приложения»). Однако в реальной практике, особенно в компаниях, где многие системы разработаны подрядчиками, создание и поддержка API может требовать значительных финансовых и временных ресурсов. В таких случаях целесообразно рассматривать альтернативные методы интеграции, которые могут быть менее затратными, но при этом удовлетворять текущим бизнес-потребностям и по мере развития КХД переходить к более стандартизированным и гибким решениям, таким как API.

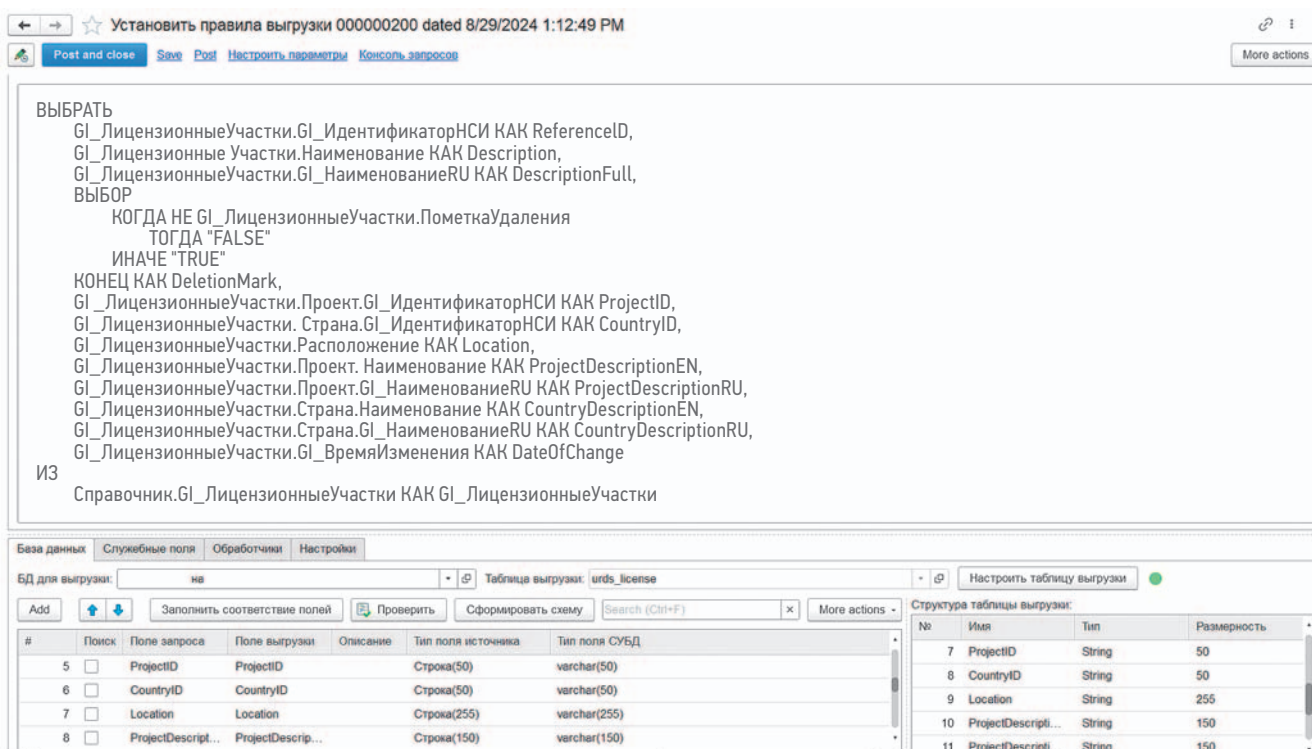


Рис. 6. Пример установки правила выгрузки в Modus ETL. Скриншот из программы Modus ETL. Составлен автором
Fig. 6. Example of setting an unloading rule in Modus ETL. Screenshot from the program Modus ETL. Prepared by the author

Этап 7 — Разработка ETL-процессов по насыщению витрин данных

По мере готовности БД, составляющих АХД, в Modus ETL активированы сценарии для насыщения витрин данных с автоматическим запуском по расписанию в виде пакетов. Конфигурация сценариев обработки данных осуществлена в интерфейсе WorkFlow посредством размещения последовательных и параллельных веток сценария на холсте, передачи данных и потока управления между шагами, настройки правил обработки данных с использованием готовых шаблонов для ETL-операций (выборки,

группировки, соединения, отбора данных и др.). На рис. 7 представлен пример сценария обработки данных для насыщения витрины ресурсной базы активов и лог выполнения каждого шага сценария. Полученные данные визуализированы на дашборде, реализованном на Python, для упрощения их восприятия и ускорения анализа (рис. 8).

Этап 8 — Управление доступом на уровнях архитектуры хранилища

Согласно утвержденной матрицы доступа в Modus ETL разграничен доступ к базам данным, источникам данных, составам

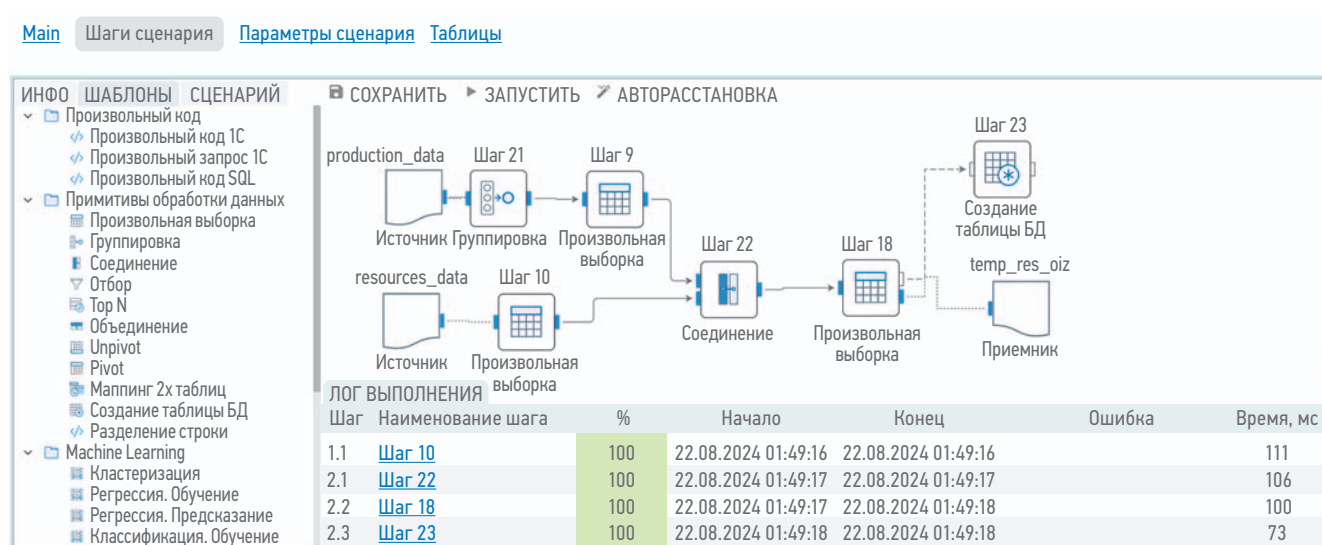


Рис. 7. Пример настройки сценария обработки данных в Modus ETL. Скриншот из программы Modus ETL. Составлен автором
 Fig. 7. Example of configuring a data processing scenario in Modus ETL. Screenshot from the program Modus ETL. Prepared by the author

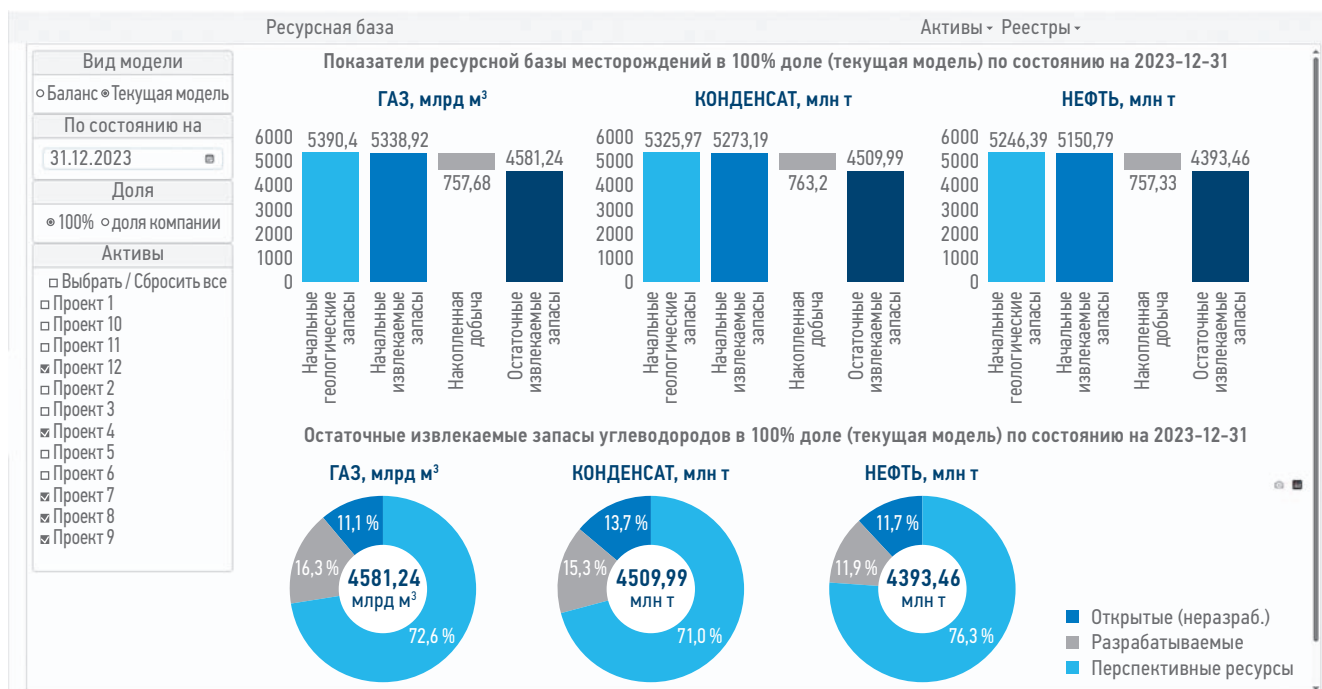


Рис. 8. Пример дашборда ресурсной базы. Скриншот из программы ИАС «Отчетность и BI». Составлен автором
 Fig. 8. Example of a resource base dashboard. Screenshot from the program IAS «Reporting and BI». Prepared by the author

Таблица 1. Мониторинг и анализ рисков построенной архитектуры. Составлена автором
Table 1. Monitoring and risks analysis in the constructed architecture. Prepared by the author

Тип угроз	Примеры угроз	Риск			Критерий риска	Решение	Остаточный риск			Решение
		вероятность	последствия	значение			вероятность	последствия	значение	
Физические	Воздействие огня	Крайне низкая	Критические	Низкое	Средний	Принять	-	-	-	Принять
Окружающая среда	Землетрясение Ураганы Стихийные бедствия	Крайне низкая	Критические	Низкое	Средний	Принять	-	-	-	Принять
Инфраструктура	Электроснабжение Телекоммуникация Нарушение кабельных линий	Средняя	Средние	Среднее	Средний	Передать (ЦОД)	-	-	-	Принять
Технические сбои	Угроза исчерпания вычислительных ресурсов Несовместимость систем Сбои программного обеспечения	Низкая	Средние	Среднее	Средний	Передать (ЦОД)	-	-	-	Принять
Человеческий фактор	Недостаточная квалификация Удаление данных Кража данных	Средняя	Средние	Среднее	Средний	Обработка	Низкая	Низкие	Низкое	Принять
Организационные	Несо согласованность бизнес-процессов Недостаточная вовлеченность бизнеса Сопrotивление пользователей и руководства Нарушение сроков и бюджета	Средняя	Средние	Среднее	Средний	Обработка	Низкая	Низкие	Низкое	Принять
Третья сторона (подрядчики)	Удаление данных Распространение данных	Средняя	Средние	Среднее	Средний	Обработка	Низкая	Низкие	Низкое	Принять

выгрузок, правилам выгрузок, сценариям и пакетам обработки данных. Регулярный пересмотр прав в системе помогает предотвратить несанкционированный доступ к конфиденциальной информации и снизить риски утечек данных.

Этап 9 — Мониторинг и анализ рисков архитектуры корпоративного хранилища данных

В процессе проработки архитектуры КХД идентифицированы уязвимости, которые могут быть использованы угрозами различной природы и привести к реализации риска, а также определен остаточный риск после обработки базового. Полученные результаты представлены в **таблице 1**.

Постоянный мониторинг и регулярный пересмотр мер безопасности обеспечивают устойчивость КХД к новым вызовам и угрозам,

поддерживая высокий уровень защиты данных.

ЗАКЛЮЧЕНИЕ

Необходимость обеспечения оперативного доступа к актуальной информации, сбора и обработки разнородных данных действительно представляет существенные вызовы для нефтегазовых компаний. Реальные примеры интеграции, в том числе приведенный в настоящей статье, служат ценным ориентиром для понимания реализованных стратегий, показывают практическое применение интеграционных фреймворков, технологий и методологий, а также предлагают информацию о приобретенных уроках и ключевых факторах, способствовавших успешным проектам интеграции.

Список литературы

1. Devlin B. Thirty Years of Data Warehousing // Business Intelligence Journal. — 2018. — Vol. 23. — №1. — P. 12–24.
2. Ahamed B., Ramkumar T. Data Integration — Challenges, Techniques and Future Directions: A Comprehensive Study // Indian Journal of Science and Technology. — 2016. — Vol. 9. — №44. — P. 1–9. <https://doi.org/10.17485/ijst/2016/v9i44/105314>
3. Katiyar N., Kulshreshtha A., Singh P.K. A Review of Integration Techniques of Multi-Geoscience Data-Sets in Mineral Prospectivity Mapping // Earth Sciences. — 2024. — №4. — P. 127–140. <https://doi.org/10.11648/j.earth.20241304.12>
4. Yong-Xin Z., Qing-Zhong L., Zhao-Hui P. A Novel Method for Data Conflict Resolution using Multiple Rules // Computer Science and Information Systems. — 2013. — Vol. 13. — №1. — P. 215–235. <https://doi.org/10.2298/CSIS110613005Y>

5. Alma'aitah W. Z., Quraan A., AL-Aswadi F.N., Alkhawaldeh R.S., Alazab M., Awaja A. Integration Approaches for Heterogeneous Big Data: A Survey // *Cybernetics and Information Technologies*. — 2024. — Vol. 24 — №1. — P. 3–20. <https://doi.org/10.2478/cait-2024-0001>
6. Chinta U., Chhapola A., Jain S. Integration of Salesforce with External Systems: Best Practices for Seamless Data Flow // *Journal of Quantum Science and Technology*. — 2024. — Vol. 1. — №3. — P. 25–41. <https://doi.org/10.36676/jqst.v1.i3.25>
7. Seenivasan D. Critical Security Enhancements for ETL Workflows: Addressing Emerging Threats and Ensuring Data Integrity // *International Journal of Innovative Research in Computer and Communication Engineering*. — 2024. — Vol. 12. — №3. — P. 1301–1313. <https://doi.org/10.15680/IJIRCCCE.2024.1203001>
8. Поршнев С.В., Пonomарева О.А. Проблемы разработки модели угроз для хранилища гетерогенных данных // *Вестник УрФО. Безопасность в информационной сфере*. — 2024. — №2(52). — С. 89–98.
9. Yessad L., Labiod A. Comparative study of data warehouses modeling approaches: Inmon, Kimball and Data Vault // *Proceedings of the 2016 International Conference on System Reliability and Science (ICSRS), Paris, France, 15–18 November 2016*. — 2016. — P. 95–99. <https://doi.org/10.1109/ICSRS.2016.7815845>

References

1. Devlin B. Thirty Years of Data Warehousing. *Business Intelligence Journal*, 2018, vol. 23, no. 1, pp. 12–24.
2. Ahamed B., Ramkumar T. Data Integration — Challenges, Techniques and Future Directions: A Comprehensive Study. *Indian Journal of Science and Technology*, 2016, vol. 9, no. 44, pp. 1–9. <https://doi.org/10.17485/ijst/2016/v9i44/105314>
3. Katiyar N., Kulshreshtha A., Singh P.K. A Review of Integration Techniques of Multi-Geoscience Data-Sets in Mineral Prospectivity Mapping. *Earth Sciences*, 2024, vol. 13, no. 4, pp. 127–140. <https://doi.org/10.11648/j.earth.20241304.12>
4. Yong-Xin Z., Qing-Zhong L., Zhao-Hui P. A Novel Method for Data Conflict Resolution using Multiple Rules. *Computer Science and Information Systems*, 2013, vol. 10, no. 1, pp. 215–235. <https://doi.org/10.2298/CSIS110613005Y>
5. Alma'aitah W.Z., Quraan A., AL-Aswadi F.N., Alkhawaldeh R.S., Alazab M., Awaja A. Integration Approaches for Heterogeneous Big Data: A Survey. *Cybernetics and Information Technologies*, 2024, vol. 24, no. 1, pp. 3–20. <https://doi.org/10.2478/cait-2024-0001>
6. Chinta U., Chhapola A., Jain S. Integration of Salesforce with External Systems: Best Practices for Seamless Data Flow. *Journal of Quantum Science and Technology*, 2024, vol. 1, no. 3, pp. 25–41. <https://doi.org/10.36676/jqst.v1.i3.25>
7. Seenivasan D. Critical Security Enhancements for ETL Workflows: Addressing Emerging Threats and Ensuring Data Integrity. *International Journal of Innovative Research in Computer and Communication Engineering*, 2024, vol. 12, no. 3, pp. 1301–1313. <https://doi.org/10.15680/IJIRCCCE.2024.1203001>
8. Porshnev S.V., Ponomareva O.A. [Challenges in Developing a Threat Model for a Heterogeneous Data Warehouse] Problemy razrabotki modeli ugroz dlia khranilishcha geterogennykh dannykh. *Vestnik UrFO. Bezopasnost v informatsionnoi sfere*. 2024, no. 2, pp. 89–98. (In Russ.)
9. Yessad L., Labiod A. Comparative study of data warehouses modeling approaches: Inmon, Kimball and Data Vault. *In Proceedings of the 2016 International Conference on System Reliability and Science (ICSRS), Paris, France, 15–18 November 2016*, 2016, pp. 95–99. <https://doi.org/10.1109/ICSRS.2016.7815845>

ВКЛАД АВТОРА / AUTHOR CONTRIBUTIONS

И.Р. Шеховцова — разработала общую концепцию статьи, подготовила текст публикации и рисунки к статье, окончательно утвердила публикуемую версию статьи и согласна принять на себя ответственность за все аспекты работы.

Irina R. Shekhovtsova — developed the article general concept, prepared the text of the article, prepared the figures for the article, approved the final version of the article and accepted the responsibility for all aspects of the work.

СВЕДЕНИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Ирина Романовна Шеховцова — ведущий специалист отдела автоматизированных систем управления МК ООО «Газпром Интернешнл Лимитед»
197022, Россия, г. Санкт-Петербург,
наб. реки Большой Невки, д. 24, стр. 1.
e-mail: I.Shekhovtsova@gazprom-international.com

Irina R. Shekhovtsova — Lead Specialist of Department of Automated Control Systems ILCC Gazprom International Limited
Building 1, 24 Bolshaia Nevka Embankment, 197022, Saint Petersburg, Russia.
e-mail: I.Shekhovtsova@gazprom-international.com