

# Гармонизация данных (Data harmonization)

Разделы: [Бизнес-задачи](#)

Гармонизация данных в целом аналогична задаче их интеграции, поскольку также предполагает объединение информации из разрозненных источников в единую среду хранения (хранилище или озеро). Однако она идет еще дальше, изменяя сведения в соответствии с единой схемой. Гармонизация позволяет использовать всю корпоративную информацию для бизнес-процессов без каких-либо ограничений, возникающих из-за различий в типах, видах и форматах данных, а также их низкого качества.

Решения по гармонизации данных могут сильно различаться в зависимости от ряда факторов:

- Объем и разнообразие используемых источников.
- Структурные различия между источниками данных (структурированные, слабоструктурированные, неструктурированные).
- Уровень достоверности и качества информации из каждого источника.
- Скорость выполнения обновлений.
- Цели соответствующего бизнес-процесса.

В большинстве случаев организации используют платформу ETL для создания полностью гармонизированного источника в централизованном хранилище или витрине данных. Процесс гармонизации обычно содержит следующие шаги:

- **Анализ.** Бизнес-аналитики изучают доступные источники данных и определяют те, которые наиболее соответствуют бизнес-процессам компании.
- **Разработка схемы.** Создается схема, определяющая способ обработки данных в рамках их гармонизации, включая процессы трансформации и очистки.
- **Извлечение.** Информация загружается из источников в сыром, необработанном виде, «как есть».
- **Преобразование.** К сырым, необработанным данным применяется разработанная схема преобразования, которая включает их обновление и интеграцию.
- **Загрузка.** Информация из нескольких источников загружается в централизованную среду хранения.
- **Подтверждение.** Аналитики подтверждают, что гармонизированные данные соответствуют заданным критериям качества, а также удостоверяют, что полученный набор соответствует бизнес-требованиям. Например, если целью было

создание хранилища клиентских данных, то они должны содержать точную и исчерпывающую информацию о каждом клиенте.

- **Развертывание.** Гармонизированные данные предоставляются бизнес-пользователям для поддержки их деятельности.

Гармонизация данных может использоваться для решения многих, проблем связанных с данными:

- **Аналитика и бизнес-аналитика.** Гармонизированные данные, в отличие от использования источников по отдельности, позволяют получить более достоверное и расширенное представление о компании и ее бизнес-процессах.
- **Управление взаимоотношениями с клиентами.** Данные о клиентах могут находиться в различных частях информационной системы компании. Гармонизация позволяет консолидировать их в одном месте. При этом любые нерелевантные или неточные материалы будут исключены.
- **Логистика.** Гармонизированные данные позволяют улучшить решения, связанные с цепочками поставок, предоставляя максимально точную картину текущего состояния компании. Это помогает повысить своевременность заказов и доставок, и выявлять проблемы по мере их возникновения.
- **Аудит и финансовый контроль.** Гармонизированные данные могут помочь ускорить процессы, связанные с аудитом и внутренним финансовым контролем за счет сокращения времени, затрачиваемого на сопоставление и очистку информации.

Гармонизация также является ключевым шагом в разработке стратегии управления основными данными компании.