

# ОБРАБОТКА ИНФОРМАЦИИ В ХРАНИЛИЩАХ ДАННЫХ

Рассматриваются существующие подходы к построению хранилищ данных и пятиуровневая архитектура хранилища данных.

## ВВЕДЕНИЕ

В настоящее время в результате непрерывного увеличения объемов обрабатываемой информации возникает необходимость в использовании новых эффективных технологий хранения и обработки данных. Одной из них является интеллектуальная обработка данных, позволяющая производить анализ данных с целью выработки наиболее эффективных бизнес-стратегий. Для проведения интеллектуального анализа реляционные базы данных должны быть преобразованы в, так называемые, хранилища данных.

### I. ПОДХОДЫ К ПОСТРОЕНИЮ ХРАНИЛИЩА ДАННЫХ

Хранилище данных (Data Warehouse) – это технология, которая объединяет структурированные данные из одного или нескольких источников для их дальнейшего анализа с целью повышения эффективности бизнес-аналитики и принятия стратегически важных бизнес-решений.

На сегодняшний день существует два основных подхода к построению Хранилища данных:

- Хранилище данных с архитектурой шины, или подход Ральфа Кимболла;
- Хранилище данных с корпоративной информационной системой, или подход Билла Инмона.

У каждого из подходов существует ряд преимуществ и недостатков, вытекающих из их архитектурных решений. Для совмещения преимуществ рассматриваемых подходов можно комбинировать данные методы для построения гибридного хранилища данных.

### II. ПЯТИУРОВНЕВАЯ АРХИТЕКТУРА ХРАНИЛИЩА ДАННЫХ

Различные системы хранилищ данных имеют разную структуру. Например, одни системы могут иметь операционные склады данных (ODS), а другие - витрины данных. Рассмотрим

*Геврасёва Ирина Петровна*, магистрант кафедры информационных технологий автоматизированных систем БГУИР, [irina\\_gvraseva@mail.ru](mailto:irina_gvraseva@mail.ru).

*Научный руководитель: Навроцкий Анатолий Александрович*, заведующий кафедрой информационных технологий автоматизированных систем БГУИР, кандидат физико-математических наук, доцент, [navrotsky@bsuir.by](mailto:navrotsky@bsuir.by).

основные уровни архитектуры хранилища данных (см. рис.1).

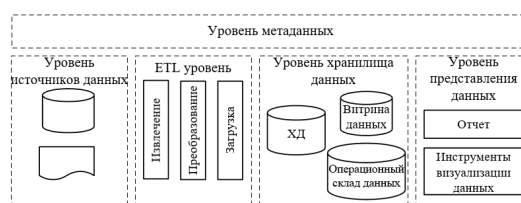


Рис. 1 – Уровни хранилища данных

Самый нижний уровень - это уровень источников данных. Этот уровень представляет собой различные источники информации, из которых данные поступают в хранилище данных. Второй уровень - это уровень извлечения, преобразования и загрузки данных (ETL уровень). Данные извлекаются из источников информации на уровень хранилища данных посредством ETL процесса. На уровне хранилища данных есть три компонента: операционный склад, хранилище и витрины данных. Уровень метаданных хранит информацию о данных (метаданные). Самый верхний уровень - это уровень представления данных, состоящий из инструментов, которые отображают информацию в разных форматах для разных пользователей.

### III. ВЫВОДЫ

Технология хранилища данных предоставляет гибкие способы доступа, агрегирования и визуализации информации. В статье представлены основные уровни для создания хранилища данных.

1. Kimball, R. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition / R. Kimball, M. Ross // Wiley Publishing, Inc. – 2013. – С. 564.
2. Inmon, W. H. Building the Data Warehouse, Fourth Edition / W. H. Inmon // Wiley Publishing, Inc. – 2005. – С. 576.