

# ПРИНЦИПЫ ИНТЕГРАЦИИ ИНФОРМАЦИОННЫХ РЕСУРСОВ

Выполнил Останин Степан

Группа 14204

01.12.2014

# Интеграция данных

- включает объединение данных, находящихся в различных источниках, и предоставление данных пользователям в унифицированном виде.



# Уровни интеграции данных

- Системы интеграции данных могут обеспечивать интеграцию данных на физическом, логическом и семантическом уровне.



# Задачи при создании системы интеграции

- Разработка архитектуры системы интеграции данных.
- Создание интегрирующей модели данных, являющейся основой единого пользовательского интерфейса в системе интеграции.
- Разработка методов отображения моделей данных и построение отображений в интегрирующую модель
- Интеграция метаданных, используемых в системе источников данных.
- Преодоление неоднородности источников данных.
- Разработка механизмов семантической интеграции источников данных.

# Консолидация

- В случае консолидации данные извлекаются из источников, и помещаются в Хранилище данных. Процесс заполнения Хранилища состоит из трех фаз — извлечение, преобразование, загрузка



# Федерализация

- В федеративных БД физического перемещения данных не происходит: данные остаются у владельцев, доступ к ним осуществляется при необходимости
- При использовании медиатора создается общее представление (модель) данных. Медиатор — посредник, поддерживающий единый пользовательский интерфейс

# Распространение данных

- Приложения распространения данных осуществляют копирование данных из одного места в другое



# Проблемы интеграции данных

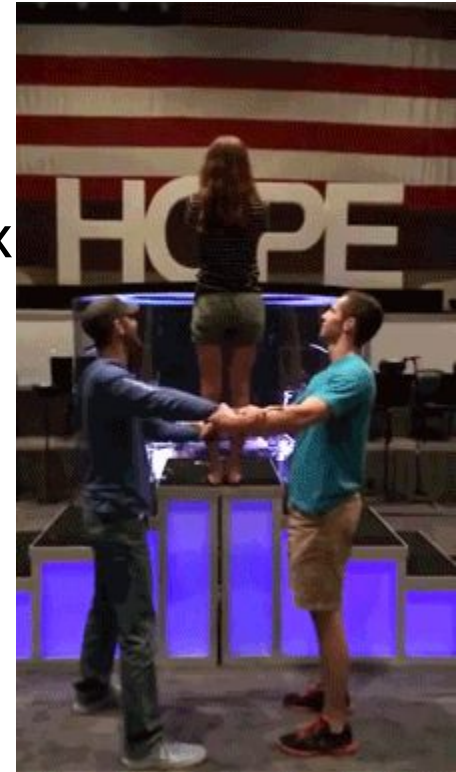
- Вне зависимости от выбранных технологии и метода интеграции данных, остаются вопросы, связанные с их смысловой интерпретацией и различиями в представлении одних и тех же вещей. Именно, приходится разрешать несоответствие несоответствие





# Типы несоответствия схем данных

- Конфликты неоднородности (используются различные модели данных для различных источников)
- Конфликты именованя (в различных схемах используется различная терминология, что приводит к омонимии и синонимии в именовании)
- Семантические конфликты (выбраны различные уровни абстракции для моделирования подобных сущностей реального мира)
- Структурные конфликты (одни и те же сущности представляются в разных источниках разными структурами данных).



# Типы несоответствия собственно данных

- Различие формата данных.
- Различие в представлении значений.
- Потеря актуальности данных одним из источников.
- Наличие ошибок операторского ввода (или ошибок распознавания бланков) в отдельных источниках данных.
- Намеренное внесение искажений с целью затруднить идентификацию сущностей.

# ИСТОЧНИКИ

- [https://ru.wikipedia.org/wiki/интеграция\\_данных](https://ru.wikipedia.org/wiki/интеграция_данных)
- МЕТОДЫ ИНТЕГРАЦИИ ДАННЫХ В ИНФОРМАЦИОННЫХ СИСТЕМАХ М.Р. Когаловский

# Конец

