

Реферат

Цель работы

Целью работы является исследование архитектур систем интеграции данных на базе технологий семантической паутины и прототипирование системы интеграции данных.

Актуальность проводимых исследований

Интеграция данных включает объединение данных, находящихся в различных источниках и предоставление данных пользователям в унифицированном виде. Этот процесс становится существенным как в коммерческих задачах (когда двум похожим компаниям необходимо объединить их базы данных), так и в научных (комбинирование результатов исследования из различных биоинформационных репозиториях, для примера). Роль интеграции данных возрастает, когда увеличивается объём и необходимость совместного использования данных.

Большие организации обычно имеют огромные базы данных, которые были спроектированы разными людьми, в разное время и хотя представляют семантически однородные данные, но представлены разными способами и разными моделями данных. Многие решения в этой области настаивают на замене этих устаревших унаследованных систем, другими более современными и более гибкими системами. Такой подход часто является не дешевым с точки зрения ресурсов и времени, что не всегда является приемлемым.

Актуальность проводимых исследований заключается в поиске возможностей интеграции данных с использованием технологий семантической паутины для решения основных проблем интеграции данных.

Решаемые в работе задачи

В работе представлены теоретические сведения об архитектурах систем интеграции данных, уровни и проблемы интеграции данных, и технологиях семантической паутины.

Достигнутые результаты

Результатом исследований является создание прототипа системы интеграции данных на базе технологий семантической паутины. Выполнено тестирование прототипа и проведен анализ результатов производительности работы системы.

Научная новизна

Научная новизна работы состоит в исследовании подходов к построению систем интеграции данных и архитектур систем интеграции данных.

Практическая ценность

Практическая ценность работы заключается в создании прототипа системы интеграции данных на базе технологий семантической паутины.

Выводы

Систематизировав сведения о подходах к построению систем интеграции данных, проанализировав основные проблемы, связанные с интеграцией данных, рассмотрев возможности использования технологий семантической паутины в интеграции данных, в ходе работы был создан прототип программной системы, позволяющий выполнять интеграцию данных разных реляционных баз данных.

Работа состоит из 113 с., 21 рисунка, 14 таблиц, 26 источников.

Перечень ключевых слов: интеграция данных, реляционные базы данных, семантическая паутина, RDF, RDFS, SPARQL, OWL, D2RQ, IRI.