

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра дискретной математики и информационных технологий

**РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОДДЕРЖКИ
ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 271 группы
направления 09.04.01 — Информатика и вычислительная техника
факультета КНиИТ
Репенинга Павла Алексеевича

Научный руководитель
доцент, к. э. н.

Г. Ю. Чернышова

Заведующий кафедрой
доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2022

ВВЕДЕНИЕ

Онтологические модели используются в различных областях, где основным способом извлечения информации являются структурированные запросы. Такой способ доступа к данным позволяет формулировать сложные запросы к базе данных. В настоящее время исследования сосредоточены на разработке визуально-интерактивных интерфейсов базы данных с поддержкой генерации запросов.

Онтологии предметных областей активно адаптируются для моделирования данных и поиска информации. Поиск информации на основе онтологий нацелен на улучшение интерфейса между данными и поисковыми запросами, чтобы приблизить результаты к требованиям пользователей. Одним из основных преимуществ использования онтологии предметной области является ее способность определять модель данных в сочетании с соответствующими знаниями предметной области. Онтологии могут быть применены для повышения оперативности доступа к данным, предоставляя интерфейс для сложных запросов.

Методы анализа устойчивости развития экономических систем должны быть подкреплены соответствующими информационными технологиями и математическими моделями. Разработка и применение онтологических моделей позволит облегчить использование систем поддержки принятия решений широким кругом пользователей. Онтологические модели могут быть использованы в открытых распределенных информационных системах, основанных на знаниях. Значительные объемы как структурированных, так и неструктурированных данных могут быть исследованы на основе подобных моделей как на уровне выработки региональной стратегии развития, так и в оперативных задачах управления регионами.

Целью данной работы является разработка информационной системы поддержки принятия решений на основе онтологической модели.

Для достижения указанной цели были поставлены следующие задачи:

- анализ возможностей применения онтологических моделей для решения задач оценки уровня регионального развития;
- выбор программного обеспечения для построения онтологий, разработка онтологической модели, проектирование информационной системы;
- разработка функционала и интерфейса информационной системы под-

- держки принятия решений на основе онтологической модели;
- реализация запросов к базе знаний;
 - применение информационной системы поддержки принятия решений на основе онтологической модели для анализа уровня социально-экономического развития российских регионов.

Объектом исследования данной работы является онтологическое моделирование. Предметом исследования является информационная система поддержки принятия решений на основе онтологической модели.

В первой главе рассматриваются возможности, области применения и свойства онтологических моделей. Вторая глава посвящена проектированию онтологической модели для анализа социально-экономического развития регионов, определению множества концептов и отношений онтологии и разработке онтологической модели. В третьей главе рассматривается разработка информационной системы, реализация этапа обработки данных и применение информационной системы поддержки принятия решений на основе онтологической модели для оценки социально-экономического развития регионов России.

Магистерская работа состоит из введения, трех глав, заключения и списка используемых источников. Общий объем работы — 56 страниц, из них 50 страниц — основное содержание, включая 15 рисунков, список использованных источников информации — 27 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе представлен анализ возможностей применений онтологических моделей.

Разработка систем, основанных на знаниях — это дисциплина, тесно связанная с программной инженерией. Термин «инженерия знаний» часто ассоциируется с разработкой экспертных систем, включая методологии, а также методы представления знаний. С самого начала понятие онтологии в компьютерных науках возникло из этой дисциплины. Несмотря на общие корни, онтологии делают упор на такие аспекты, как межагентное взаимодействие и функциональная совместимость. В компьютерных науках понятие онтология интерпретируется по-разному, и конкретные онтологии могут различаться по нескольким параметрам, таким как: степень формальности; авторитетность; качество. Различные виды онтологий можно классифицировать по назначению, специфичности и выразительности. В связи с появлением «семантической паутины» онтологии, в последнее время, привлекают большое внимание. Были разработаны новые технологии и инструменты для представления онтологий. Это значительно упрощает их внедрение в реальных приложениях. Способы применения онтологий в программной инженерии разнообразны, что увеличивает терминологическую путаницу. Поэтому необходимо предоставить основу для категоризации потенциального использования онтологий в программной инженерии.

Онтологические модели используются в различных областях, например, в системах управления информацией, где одним из способов получения информации являются структурированные запросы. Онтологические модели используются для ситуационного управления большими системами, концептуального анализа и проектирования систем организационного управления. Считается, что онтологии являются лучшим способом представления многократного и совместного использования знаний. Также онтологии востребованы в моделировании мультиагентной системы, а для моделирования понятий предметной области используются онтологические категории верхнего уровня. При наличии декларативной спецификации терминов, онтологии могут применяться для анализа знаний в предметной области.

Выделяют следующие области применения онтологических моделей:

- отделение знаний предметной области от оперативных знаний;

- повторное использование знаний;
- совместное использование пользователями;
- анализ знаний в предметной области;
- создание явных допущений.

Совместное использование людьми и программными агентами общего понимания структуры информации – одна из наиболее общих целей разработки онтологии. Например, если несколько web-сайтов обладают информацией о платных медицинских услугах и публикуют одну и ту же базовую онтологию терминов, то программные агенты могут накапливать знания и применять их для ответов на запросы пользователей.

Зачастую построение онтологии само по себе не является конечной целью. Обычно онтологии используются другими программами для решения практических задач. Выделяют следующие сферы использования онтологий:

- машинный перевод;
- вопросно-ответные системы;
- информационный поиск;
- системы извлечения знаний;
- системы ведения диалога между компьютером и человеком;
- системы понимания языка.

Таким образом, перспективы развития онтологий в настоящее время связаны с представлением знаний, искусственным интеллектом и компьютерной обработкой данных. В частности, в сфере лингвистики онтологии могут быть использованы для семантического аннотирования текстовых корпусов, машинного перевода, автоматического разрешения многозначности и снятия омонимии на основе контекста, а также для создания ресурсов онтологического типа, находящихся на более низком уровне: глоссариев и тезаурусов.

Поскольку моделирование онтологий является утомительной и дорогостоящей задачей, всегда важно продемонстрировать преимущества, которые можно получить, применяя онтологии в программной инженерии. Это подчеркивается тем фактом, что большинство формальных основ онтологий существовали в течение длительного времени, не получив широкого распространения среди разработчиков программного обеспечения. Таким образом, очевидно, что нынешнее появление формализмов, основанных на логике, в контексте семантической паутины, является важным фактором.

Во второй главе рассматривается разработка онтологической модели для анализа социально-экономического развития регионов.

При создании онтологий целесообразно пользоваться подходящими инструментами. В настоящее время редакторы онтологий предоставляют средства кодирования формальной модели в том или ином виде. Функциональность редактора является важной характеристикой. Для работы с одним или более проектами выделяют базовый набор функций: сохранение проекта в нужном формализме и формате; открытие проекта; импорт из внешнего формата; редактирование метаданных проекта; редактирование онтологии. В набор возможных действий включают следующие функции: создание, редактирование, удаление понятий, отношений, аксиом и прочих структурных элементов онтологии, редактирование таксономии.

В данной работе используется редактор онтологий Protege. Он был выбран как современный редактор онтологий OS OWL. Protege включает в себя большой набор подключаемых модулей, таких как редактор запросов SPARQL.

Для построения онтологической модели были созданы основные концепты с последующим наращиванием нужных концептов. Для работы онтологии в тестовом режиме с социально-экономическими показателями была разработана система классов с возможностью их расширения:

- организации(company);
- реквизиты организации(company_details);
- показатели(indicators);
- регионы(regions).

Свойства классов:

- включают действия(include_actions);
- код и название организации(organization_code_and_name);
- код организации(organization_code);
- показатели организации(exponentsorganizations);
- идентификационный номер показателя(id_indicator);
- идентификационный номер организации(id_organizations);
- идентификационный номер региона(id_region);
- идентификационный номер реквизита(id_request);
- включает период(turn_on_period).

Свойства данных:

- выручка, тыс.руб.(evenue);
- себестоимость продаж, тыс.руб. (cost_sales);
- валовая прибыль, тыс. руб. (gross_profit_margin);
- уровень валовой прибыли к выручке, % (level_gross_profit_to_revenue);
- среднегодовая стоимость основных средств, тыс.руб. (average_annual_cost_of_fixed_assets);
- фондоотдача, тыс. руб. (return_funds);
- рентабельность основных фондов, % (return_fixed_assets);
- среднегодовая стоимость оборотных средств, тыс.руб. (average_annual_cost_of_working_capital);
- время обращения оборотных средств, дни (working_capital_circulation_time);
- скорость обращения оборотных средств, число оборотов (working_capital_turnover_rate);
- среднесписочная численность работников, чел. (average_number_of_employees);
- расходы на оплату труда, тыс.руб. (labor_costs);
- среднемесячная заработная плата 1 работника, руб. (average_monthly_salary_one_employee);
- производительность труда, тыс.руб./чел. (labor_productivity);
- издержки обращения, тыс.руб. (handling_costs);
- уровень издержек обращения, % (level_circulation_costs);
- прибыль от продаж, тыс.руб. (profit_sales);
- рентабельность продаж, % (return_sales);
- прочие доходы, тыс.руб. (other_income);
- прочие расходы, тыс.руб. (other_expenses);
- прибыль до налогообложения тыс.руб. (profit_before_tax);
- текущий налог на прибыль, тыс.руб. (current_income_tax);
- чистая прибыль, тыс.руб. (net_profit);
- рентабельность деятельности, % (profitability_operations).

В результате разработанная онтология, на основе выше перечисленных концептов, ограничена областью интереса для конкретной задачи и позволит избежать несоответствия в получаемых данных.

В третьей главе представлена разработка информационной системы.

Функционал информационной системы был разработан с помощью средств языка программирования Java. Он может быть использован как инструмент для создания сложных web-приложений. Благодаря тому, что язык Java является объектно-ориентированным, становится возможным повторного использования программного кода для ускорения процесса разработки масштабирования разрабатываемого приложения. Чтобы упростить разработку web-приложения использовался Spring Framework, который представляет собой контейнер внедрения зависимостей с несколькими удобными слоями (например: доступ к базе данных, прокси, аспектно-ориентированное программирование, RPC, web-инфраструктура MVC), что и позволило быстрее создавать Java-приложения. Для сокращения временных затрат на разработку серверной части приложения применялись модули Spring MVC. Упрощение разработки становится ощутимо, когда приложение увеличится в размерах.

Созданный проект собирается с помощью Maven. Для сборки из шаблонов проекта в Maven существует функция сборки из архетипа, где архетип – это шаблон проекта, который используется для построения проекта. Архитектура серверной части приложения представлена на рисунке 1.

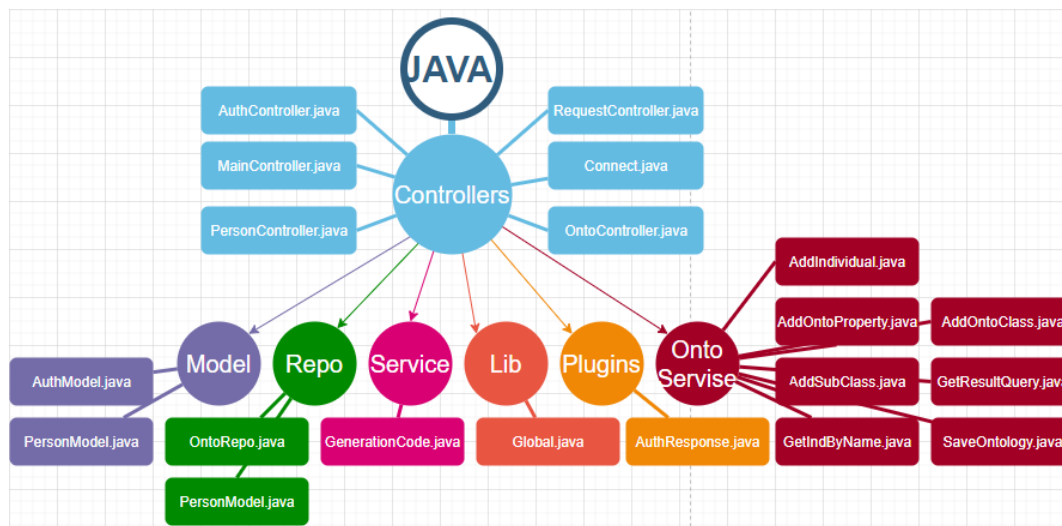


Рисунок 1 – Архитектура серверной части приложения

Архитектура проекта построена таким образом, что при выполнении каких-либо действий, все запросы обрабатывает Controllers и запускает соответствующие действия. Для взаимодействия с онтологической моделью используется OntoController.java. При разработке инструментов взаимодействия с онтологической базой знаний было решено множество проблем, так как в наборе инструментов Spring Framework нет инструментов для работы с онто-

логическими базами знаний, то для этих целей необходимо было использовать другой API.

Для процедурной генерации онтологической базы знаний и другой работы с онтологическими моделями используется Jena API.

Был разработан контролер `OntoContoller.java` который позволит взаимодействовать с файлом онтологической базы знаний на жестком диске. Для этого можно использовать основной функционал Jena API, который позволяет создавать экземпляры классов и связывать классы с заготовленными отношениями. Jena API включает в себя большой набор инструментов для работы с онтологическими базами знаний.

Разработанная система позволит взаимодействовать с онтологической базой знаний для добавления, чтения различных фрагментов базы знаний. Информационная система позволит выполнять SPARQL-запросы для анализа данных и построить поисковые запрос таким образом, чтобы можно было выявить какие-либо аналитические сведения исходя из полученных результатов.

Для разработки пользовательского интерфейса в информационных системах существует несколько подходящих инструментов: язык гипертекстовой разметки HTML, мультипарадигменный язык программирования JavaScript; jQuery – набор функций JavaScript.

При разработке HTML-документа используются стандартные подходы к верстке страницы. Используются CSS-стили и стандартные теги HTML.

Для реализации сложных манипуляций с пользовательским интерфейсом используется jQuery, который позволит обрабатывать множество событий во время работы с интерфейсом и отправлять запросы к необходимым контроллерам.

В результате разработанная архитектура обладает возможностью оперативной модификации приложения с расширением базы знаний и позволит добавлять новый функционал, не изменяя заложенную логику в информационную систему.

ЗАКЛЮЧЕНИЕ

Онтологическое моделирование может сыграть важную роль в анализе региональной статистики. В настоящее время онтологическое моделирование и онтологические базы знаний применяются в управлении информационными системами в образовательных и научных организациях, в медицинских учреждениях и предприятиях.

Разработанная информационная система обладает следующим функционалом:

- формирование пользовательских запросов;
- визуализация результатов запросов;
- вывод описательной статистики.

Особенностью приложения является возможность многократного использования данных без дополнительного обращения к базе знаний. Приложение обеспечивает возможность отображения данных в разных форматах в зависимости от запросов пользователей.

Предлагается использование онтологической модели, которая была спроектирована методом построения по наполнению. Такая онтология строится на основе реальных данных, что позволяет избежать затрат времени на дальнейшие доработки в логической составляющей онтологии.

На этапе проектирования информационной системы была выбрана трехслойная архитектура, а именно пользовательский интерфейс, серверная часть и онтологическая модель.

Для реализации пользовательского интерфейса был использован язык гипертекстовой разметки HTML, язык программирования JavaScript и jQuery. При проектировании интерфейса для упрощения разработки все компоненты были разделены на модули, такой подход позволит определить местонахождение ошибок и осуществлять расширение проекта, не вызывая изменений в архитектуре. Графики и таблицы были разработаны с использованием шаблона, он генерирует программный код интерфейса в зависимости от передаваемых данных и параметров. Такой шаблон избавляет от повторяющегося программного кода, и при расширении проекта позволит избежать ошибок в интерфейсе.

При разработке серверной части приложения предлагается использование языка программирования Java, контейнера внедрения зависимостей со

слоями Spring Framework, Spring MVC и сборщика проектов Maven. Серверная часть использует наследование классов контроллеров, что позволяет реализовать главный класс. Данный класс разбивает задачи на компоненты для выполнения которых используются инструменты, отвечающие за определенный функционал, и выполняет распределение ресурсов серверной части приложения. Для процедурной генерации онтологической базы знаний и работы с онтологическими моделями предлагается использование Jena API.

Для реализации запросов к базе знаний была разработана система, которая генерирует SPARQL-запросы в зависимости от выбранных параметров. Серверная часть позволяет обрабатывать полученные данные из базы знаний и отправлять их для отображения в пользовательский интерфейс.

При построении онтологии предлагается использование Protege. Protege поддерживает наиболее распространенные языки представления знаний и имеет легко расширяемую архитектуру.

Для анализа социально-экономических показателей региона были использованы данные, которые находятся в открытом доступе.

В результате разработанная информационная система позволяет анализировать социально-экономические показатели и фильтровать полученные данные для анализа развития региона. Онтологическая модель позволит использовать ресурсы информационной системы без значительных затрат и сбоев в работе системы, а получаемые данные представлены в легком для восприятия формате.

В качестве дальнейшего развития информационной системы предлагается расширение онтологической модели для использования в семантической сети, что позволит наполнять базу знаний пользователями для анализа большего количества данных.

Данная работа была представлена на студенческой научной конференции (г. Саратов, СГУ имени Н. Г. Чернышевского, 22 апреля 2022 г.).

Основные источники информации:

- 1 Ontology 101 - Protege Wiki [Электронный ресурс]. — URL: <https://protegewiki.stanford.edu/wiki/Ontology101> (Дата обращения 25.10.2021). Загл. с экр. Яз. рус.
- 2 Nirenburg, S. Ontological semantics / S. Nirenburg, V. Raskin // Computational Linguistics. — 2005. — Pp. 147–152.

- 3 Arafeh, M. Ontology based recommender system using social network data / M. Arafeh, P. Ceravolo, A. Mourad, E. Damiani // Future Generation Computer Systems. — 2021. — Pp. 769–779.
- 4 Ghahremanloo, L. Using ontology design patterns to represent sustainability indicator sets / L. Ghahremanloo, J. A. Thom, L. Magee // Lecture Notes in Computer Science. — 2017. — Pp. 70–81.
- 5 Kumazawa, T. Toward knowledge structuring of sustainability science based on ontology engineering / T. Kumazawa, O. Saito, K. Kozaki, T. Matsui, R. Mizoguchi // Sustainability Science. — 2009. — Vol. 4, no. 1. — Pp. 99–116.
- 6 Ефименко, И. В. Онтологическое моделирование экономики предприятий и отраслей современной России / И. В. Ефименко, В. Хорошевский. — Москва, 2011. — 68 с.
- 7 Ефименко, И. В. Онтологическое моделирование предприятий: методы и технологии / И. В. Ефименко, В. Ф. Хорошевский. — Екатеринбург, 2019. — 236 с.
- 8 Jena API Tutorial for AllegroGraph [Электронный ресурс]. — URL: <https://franz.com/agraph/support/documentation/current/java-tutorial/jena-tutorial.html> (Дата обращения 27.05.2022). Загл. с экр. Яз. англ.
- 9 Using Apache Jena for Aviation Data: Create, Import and Query a Dataset [Электронный ресурс]. — URL: https://www.bytefish.de/blog/apache_jena.html (Дата обращения 15.05.2022). Загл. с экр. Яз. англ.
- 10 Constructing SPARQL Queries [Электронный ресурс]. — URL: <https://medium.com/wallscope/constructing-sparql-queries-ca63b8b9ac02> (Дата обращения 16.05.2022). Загл. с экр. Яз. англ.