

Аннотация

К отчету по ГК П1045 от 20.08.2009

Проблема: «Обработка и доступ к информации, хранящейся в базах данных, средствами работы с онтологиями».

Направление: «Обработка, хранение, передача и защита информации».

Процесс стремительного и неконтролируемого роста глобальных и корпоративных информационных пространств делает все более трудной задачу обработки накапливаемой в них информации. В этом контексте важную роль играет развитие таких средств обработки данных, которые способны передать автоматизированным сетевым сервисам/агентам часть интеллектуальных функций по обработке информации - тех функций, которые выполняются сегодня человеком. Основным инструментом для достижения данных целей являются онтологии.

Концепция онтологий сегодня является одним из самых активно развивающихся формализмов представления знаний в интернете. Онтологии – эффективный вариант накопления и представления знаний, являющийся, с одной стороны, достаточно выразительным, с другой стороны, естественным и легким для понимания. Онтология определяется как иерархия классов и объектов плюс логическое описание свойств и механизмов взаимодействия этих объектов. Онтология – это формальный способ представления знаний, который состоит из таксономии, набора логических свойств предметной области и правил вывода. Онтология универсальна и способна описывать многие системы знаний. Понятие онтологии было формализовано и стандартизировано в рамках языка OWL консорциумом W3C.

Таким образом, онтологии представляют удобный инструмент способный взять часть интеллектуальных функций по обработке информации. Но в настоящее время подавляющая масса информации во всем мире представлена в базах данных, которые предназначены исключительно для хранения данных и никак не подходят для описания и представления знаний, а значит не подходят для реализации интеллектуальных функций обработки информации. Из чего следует две проблемы:

- Во-первых, невозможность построения над уже имеющимися базами данных интеллектуализированных сервисов «с нуля»
- Во-вторых, невозможность вовлечение баз данных хотя бы как источники данных в уже существующие интеллектуализированные сервисы по той причине, что формат онтологий и баз данных принципиально являются несовместимыми.

Как выход из сложившейся ситуации, может быть предложено три варианта:

- Копирование информации из баз данных в онтологии. Т.е. данный вариант предусматривает создание онтологий параллельно с базами данных. Данный подход имеет сложности связанные непосредственно с самим переносом информации из баз данных в онтологии и, что немаловажно, возникает сложность с синхронизацией между базами данных и онтологиями, т.е. когда «старый» сервис изменяет какую-то информацию в базе данных автоматически должна поменяться информация в онтологии. И наоборот, когда у нас меняется онтологии, должны произойти соответствующие изменения и в базе данных.
- Как вариант, чтобы избежать проблемы синхронизации, можно полностью заменить базы данных на онтологии. Т.е. данный вариант предусматривает полный перенос информации из баз данных в онтологии. Данный подход полностью решает проблему синхронизации, но наряду с имеющейся проблемой переноса информации порождает новую проблему адаптации имеющихся сервисов. Имеющиеся сервисы, ориентированные на работу с базами данных, будет крайне сложно перестроить на работу с онтологиями. Да и по затратам данная проблема намного серьезнее, чем проблема переноса информации
- Как альтернативу вышесказанному может быть предложен следующий вариант. Для работы старых «сервисов» нет никакой необходимости переносит информацию из баз данных в онтологии. А для создания сервисов «нового поколения» уже используется онтологии. Встает только проблема, как адаптировать информацию хранящуюся в старых базах данных необходимую для новых сервисов, работающих с онтологиями. Данная работа посвящена решению данной проблемы. Предлагается создание так называемого «переходника», позволяющего работать с базами данных как с онтологиями.

Такой подход позволяет работать с базами данных, как традиционным образом, так и средствами работы с онтологиями. Поскольку непосредственный перенос информации из баз данных в онтологии представляется крайне затруднительным, то такой подход является самым оптимальным способом адаптации баз данных к сервисам нового поколения.

Данная работа посвящена созданию такого переходника, позволяющего адаптировать имеющиеся базы данных к новым сервисам, т.е. разрабатываемый переходник обеспечивает онтологический программный интерфейс для доступа к онтологиям.

На первом этапе исследования производится теоретическое исследование возможности и способа представления базы данных как онтологии.

Второй этап исследования целиком посвящен программной реализации программных модулей, реализующих заявленный потенциал.

Руководитель проекта

И.А. Казаков