

0.1. Дородных Н.О., Видья А., Юрин А. Автоматизированное создание онтологий на основе электронных таблиц

На сегодняшний день в рамках исследований по искусственному интеллекту накоплен большой опыт и широкий спектр различных методов и средств представления знаний. В последнее время широкую популярность приобрели онтологические модели представления знаний [1], которые рассматриваются в качестве основного элемента в концепции «Семантической паутины» (Semantic Web). Как правило, онтологии используются системными аналитиками и экспертами предметной области для концептуализации своих знаний и их представления в формате, пригодном для машинной обработки. Онтологии являются ядром современных предметно-ориентированных систем, основанных на знаниях. Однако, несмотря на значительные успехи в данной области и наличие мощных инструментальных средств в виде различных онтологических редакторов (например, Protege, ONTOedit, OntoStudio, WebOnto, Fluent Editor и др.), до сих пор разработка онтологий является одним из самых узких мест при построении подобного рода систем [2]. В данном контексте актуальным является использование различных источников информации (например, тексты, базы данных, таблицы, диаграммы, веб-ресурсы и т.п.) для автоматизированного формирования онтологий.

В данной работе предлагается использовать электронные таблицы для поддержки формирования онтологий предметной области. Таким образом, целью данной работы является разработка метода автоматизированного создания онтологий на основе анализа и преобразования электронных таблиц в формате Excel. Метод основан на модельно-ориентированном подходе (Model-Driven Engineering) и принципах модельных трансформаций [3]. Основной особенностью разработанного метода, определяющими его новизну, является использование канонической (реляционной) формы [4] для представления электронных таблиц, обладающих произвольной компоновкой и стилями оформления.

Метод состоит из четырех основных шагов:

1. Трансформация исходных электронных таблиц с произвольной компоновкой в каноническую форму.
2. Получение фрагментов онтологии на основе анализа и трансформации канонических электронных таблиц.
3. Агрегация отдельных фрагментов онтологии в единую полную онтологическую модель.
4. Генерация кода онтологической модели в формате OWL2 DL [5].

Разработанный метод реализован в форме прототипа программного средства и использовался при

формировании онтологии деградационных процессов в нефтехимии.

Работа выполнена при финансовой поддержке Совета по грантам Президента России (проект: МК-1647.2020.9).

Список литературы

- [1] GUARINO N. Formal Ontology in Information Systems // Proc. the First Interna. Conf. on Formal Ontology in Information Systems (FOIS'98). 1998. Vol. 46. P. 3–15.
- [2] ГАВРИЛОВА Т.А., КУДРЯВЦЕВ Д.В., МУРОМЦЕВ Д.И. Инженерия знаний. Модели и методы // СПб.: Лань, 2016. 324 с.
- [3] CRETU L.G., FLORIN D. Model-Driven Engineering of Information Systems: Principles, Techniques, and Practice // Apple Academic Press, 2014.
- [4] DORODNYKH N.O., YURIN A.YU., SHIGAROV A.O. Conceptual Model Engineering for Industrial Safety Inspection Based on Spreadsheet Data Analysis // Communications in Computer and Information Science. Modelling and Development of Intelligent Systems (MDIS 2019). 2020. Vol. 1126. P. 51–65.
- [5] GRAU B.C., HORROCKS I., MOTIK B., PARSIA B., PATEL-SCHNEIDER P., SATTLER U. OWL 2: The next step for OWL // Web Semantics: Science, Services and Agents on the World Wide Web. 2008. Vol. 6. No. 4. P. 309–322.