

ПОДХОДЫ К РЕШЕНИЮ ПРОБЛЕМ РАЗРАБОТКИ БАЗЫ ЗНАНИЙ ПО КУРСОВЫМ ПРОЕКТАМ СТУДЕНТОВ ВУЗОВ

Рассматриваются проблемы онтологического моделирования и семантической формализации знаний при создании интеллектуальной базы курсовых проектов с использованием OWL, LLM и алгоритмов WSD.

Введение

В условиях цифровизации вузов традиционные файловые репозитории перестают отвечать задачам управления знаниями. Современная парадигма требует перехода к интеллектуальным базам знаний, где курсовой проект представлен как структурированный граф понятий и связей. Основная сложность заключается в глубокой формализации и автоматизации извлечения знаний из текстов.

I. Архитектурный переход к онтологиям

Онтологическое моделирование на базе OWL предлагает подход «открытого мира» [1]. Знания здесь представлены триплетами, что позволяет расширять структуру БЗ без перепроектирования схемы и использовать логический вывод для ответов на сложные запросы о связях технологий и методик исследования [2].

II. Формализация и структура проекта

Для унификации структуры курсовых работ параметры форматирования используются как семантические маркеры. В спроектированной доменной онтологии ключевой класс «Course Project» связан с:

- **Субъектами:** Student, Supervisor.
- **Контентом:** Topic, Technologies, Method.
- **Процессом:** Relevance, Goal, Task.
- **Результатом:** Project Result, Grade.

Это позволяет трансформировать ER-модели в OWL-классы, сохраняя иерархию сущностей [3].

III. Использование LLM и проблема WSD

LLM (например, Gemini 1.5 Pro) позволяют автоматизировать наполнение онтологий через many-shot

prompting, достигая точности F1-score до 0.99 [4]. Главной проблемой остается полисемия. Для её разрешения в системе применяются:

1. **Методы на знаниях:** анализ контекста алгоритмом Леска с опорой на WordNet [5].

2. **Дистрибутивная семантика:** использование эмбедингов, где вектор слова зависит от окружения, что позволяет точно определить смысл термина.

IV. Заключение

Онтологическое моделирование в связке с LLM и методами WSD создает гибкое ядро для анализа академических знаний и эффективной поддержки педагогических решений.

Список литературы

1. Hay, D. C. Data Modeling, RDF, & OWL – Part One: An Introduction To Ontologies / D. C. Hay // MIT Course Materials. – 2013. – URL: <http://web.mit.edu/smadnick/www/Courses/2013BIGData/07a%20Introduction%20to%20Ontologies.pdf>.
2. Web Ontology Language (OWL) / CORE Scholar. – Wright State University, 2013. – URL: https://corescholar.libraries.wright.edu/context/cse/article/1183/viewcontent/Web_Ontology_Language.pdf.
3. On Transforming the ER Model to Ontology Using Protégé OWL Tool // International Journal of Computer Theory and Engineering (IJCTE). – 2010. – Vol. 2, № 5. – P. 722-726. – URL: <https://www.ijcte.org/papers/914-I023.pdf>.
4. Automated Extraction of Key Entities from Non-English Mammography Reports Using Named Entity Recognition with Prompt Engineering / Y. J. Kim, H. J. Lee, J. H. Park [et al.] // PubMed Central (PMC). – 2024. – URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11852152/>.
5. Lesk's Algorithm: A Method for Word Sense Disambiguation in Text Analytics // Medium (Data Science). – 2019. – URL: <https://medium.com/data-science/lesks-algorithm-a-method-for-word-sense-disambiguation-in-text-analytics-52c157a2fdff>.

Поплавский Кирилл Дмитриевич, студент каф. ИИТ БГУИР, kirillpoplavskiy8@gmail.com.

Башура Артем Олегович, студент каф. ИИТ БГУИР, artembasural@gmail.com.

Кулик Евгений Дмитриевич, студент каф. ИИТ БГУИР, yevgeniy_kulik@list.ru.

Научный руководитель: Зотов Никита Владимирович, ассистент каф. ИИТ БГУИР, n.zotov@bsuir.by.