

## ПОСТРОЕНИЕ СЕМАНТИЧЕСКОЙ СЕТИ С ПОМОЩЬЮ ГЕНЕРАТИВНЫХ МОДЕЛЕЙ

Сафонова О.Д. (ИТМО)

Научный руководитель – Клименков С.В. (ИТМО)

**Введение.** В условиях роста неформализованных текстовых данных на естественных языках повышается актуальность применения семантических сетей в системах интеллектуального извлечения информации. Формализация представлений данных позволяет обеспечивать контекстно-зависимый анализ запросов пользователя. При этом необходимым условием эффективного функционирования таких систем является предварительное наполнение семантической сети структурированными знаниями (концептами) в рамках заданной предметной области.

В виду необходимости постоянного расширения концептов (узлов) семантической сети возможно применение генеративных моделей для автоматизированного структурирования данных по предметной области. Такие модели позволят формировать новые концепты и релевантные связи на основе анализа текстовых данных предметной области. Интеграция семантической сети с генеративными моделями позволит повысить полноту и актуальность предоставляемых сетью данных.

Эффективность семантической сети предполагает непрерывное обновление новыми концептами и связями в фоновом режиме для поддержания актуальности данных. Генеративные модели позволят динамически добавлять новые концепты и уточнять существующие концепты и связи с помощью информации из открытых источников сети Интернет. Такая архитектура обеспечивает устойчивость системы к устареванию данных.

**Основная часть.** В основной части исследования рассматривается возможность применения генеративных моделей для предварительной обработки данных и интеграции с семантической сетью.

На первом этапе генеративные модели будут использоваться для обработки исходных текстовых данных на естественном языке, их семантического обогащения и создания узлов семантической сети. На финальном этапе – для уточнения связей между концептами и автоматического добавления новых узлов.

Актуальность применения генеративных моделей для этой задачи обусловлена возможностью обобщения и контекстного вывода. Такие модели позволяют сочетать генеративные механизмы и символические структуры-команды семантической сети. Такие команды имеют SQL-подобную структуру, что позволяет текстовой генеративной модели выводить концепты семантической сети в формате запросов.

С технической точки зрения генеративные модели будут связаны в отдельный модуль генерации данных, имеющий интерфейс для взаимодействия с пользователем. Кроме того, данный модуль будет интегрироваться с семантической сетью напрямую.

В качестве обучающей базы для генеративной модели будет использован набор текстовых данных на естественных языках, релевантный для заданной предметной области. При этом ключевой этап – предварительное обучение генеративных моделей на документации семантической сети, описывающей связи и виды отношений между концептами. Такой подход позволит обеспечить согласованность внутреннего представления модели с ограничениями семантической сети, что снизит вероятность семантических расхождений при генерации запросов для семантической сети. Для повышения точности применяется RAG архитектура, обеспечивающая обращение к актуальной информации на этапе генерации.

**Выводы.** Интеграция семантической сети с генеративными моделями формирует основу для построения адаптивных систем интеллектуального информационного извлечения. Такие системы будут способны динамически развиваться благодаря актуализации данных с применением RAG архитектуры. Предложенный подход позволит поддерживать данные в актуальном и согласованном состоянии.

#### **Список использованных источников**

1. Клименков С.В., Николаев В.В., Харитонов А.Е., Гаврилов А.В., Письмак А.Е., Покид А.В. Применение семантической сети для хранения слабоструктурированных данных // Инженерный вестник Дона. — 2020. — № 2. — URL: <http://ivdon.ru/ru/magazine/archive/n2y2020/6339>
2. Klimenkov S., Tsopa E., Pismak A., Yarkeev A. Reconstruction of Implied Semantic Relations in Russian Wiktionary // Proc. of the 8th Intern. Joint Conf. on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KDIR). 2016.
3. Osika V. P., Klimenkov S., Tsopa E., Pismak A., Nikolaev V., Yarkeev A. Method of reconstruction of semantic relations using translangual information // Proc. of the 9th Intern. Joint Conf. on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KDIR). 2017.

Автор \_\_\_\_\_ Сафонова О.Д.

Научный руководитель \_\_\_\_\_ Клименков С.В.