**Семантически-ориентированный естественно-языковой интерфейс для взаимодействия с Системой взаимосвязанных открытых данных (Linked Open Data)**

В современном мире всё большее распространение получают связные данные в связи с внедрением соответствующих принципов в Интернете. Связность данных обеспечивается в системе взаимосвязанных открытых данных (Linked Open Data). В качестве внутреннего формата хранения и представления используется RDF.

RDF (Resource Description Framework) – формат описания ресурсов и их взаимосвязей. Записи в формате RDF представляют собой тройки вида {Субъект, Отношение, Объект}. Субъект и объект являются элементами множества ресурсов, а отношение – множества свойств (подмножества ресурсов, описывающих свойства и отношения между ресурсами). Для обращения к взаимосвязанным данным в формате RDF используется язык SPARQL.

SPARQL – язык запросов к взаимосвязанным данным в формате RDF. Имеет SQL-подобную структуру запроса, а именно «SELECT FROM WHERE». Иными словами, запрос на языке SPARQL определяет какой ресурс является искомым, в каком источнике проводить поиск и каким параметрам должен отвечать искомый ресурс.

Учитывая тот факт, что для обращения к LOD необходимо владение языком SPARQL и хотя бы иметь представление о формате RDF, использование системы становится невозможным или сильно затруднённым для людей, не имеющих опыта использования специальных языков запросов, таких как SQL или SPARQL. В свою очередь LOD была бы полезна в разнообразных отраслях, таких как медицина, здравоохранение, научные исследования и бизнес, т.к. обеспечивает семантическую связь между разными ресурсами, в качестве которых могут выступать научные публикации, разнообразные отчеты компаний и законодательные акты. В данной ситуации очевидным решение является разработка интеллектуального естественно-языкового интерфейса, обеспечивающего обращение к LOD без специальных технических навыков, что позволит расширить использование LOD в отраслях, где система уже используется, и внедрить, где не используется. Но таких естественно-языковых интерфейсов, использующих в качестве языка для запросов русский нет, поэтому темой данной работы ставится разработка такого интерфейса.

Для использования естественно-языкового интерфейса при обращении к LOD необходимо реализовать перевод запроса с естественного языка на язык запросов SPARQL. Поскольку данные в LOD связываются семантически, то и при переводе необходимо учитывать семантику запроса. Следовательно, при переводе требуется промежуточный язык, позволяющий описать семантическое представление запроса на естественном языке, которое далее будет переведено в конструкции языка SPARQL.

Существует несколько подходов описания семантического представления, а именно: AMR, грамматика Монтегю и К-представление Фомичева В.А. (концептуальное представление). AMR и грамматика Монтегю специализированы на семантическом представлении английского языка и с их помощью возможно описать только семантическое представление простых предложений. В свою очередь К-представление представляет формализованный аппарат для описания семантического представления текстов (дискурсов), а не только отдельных предложения, на русском языке. Учитывая все выше написанное, для описания семантического представления запроса на естественном языке выбран подход К-представление.