

УДК 004.89

АНАЛИЗ АРГУМЕНТАЦИИ В ТЕКСТАХ НА РУССКОМ ЯЗЫКЕ

Осадчий Д. (университет ИТМО)

Научный руководитель – д.т.н., старший научный сотрудник Котельников Е. В.

(Национальный центр когнитивных разработок Университета ИТМО)

В предлагаемой работе приведен обзор исследований в сфере анализа аргументации и анализ существующего русскоязычного корпуса. Рассмотрены методы получения и аннотации корпусов и методы трансфера аннотации с корпусов на других языках.

Введение. Стремительное увеличение объема и повышение важности цифровой информации при принятии решений на разных уровнях ставит новые технологические вызовы к ее структурированию и оценке. Перспективным направлением для решения подобных задач является интеллектуальный анализ аргументов. На сегодняшний день проведено большое количество исследований и создано множество корпусов для задач анализа аргументации на английском языке, однако обработка русскоязычных текстов практически не представлена.

Основная часть. Интеллектуальный анализ аргументации – это междисциплинарное направление научных исследований, включающее в себя лингвистику, машинное обучение, технологии представления знаний и обработки естественного языка. Разметка корпусов является нетривиальной задачей, даже при ручной аннотации. Однако существует большое количество размеченных корпусов на английском и немецком языках, что позволяет произвести перенос аннотации на переведенные тексты. При этом различные алгоритмы, примененные к полученным таким образом корпусам, показывают сравнимые с оригинальными корпусами результаты.

Для экспериментов в настоящей работе был использован единственный доступный на сегодняшний день русскоязычный корпус, который является переводом размеченного корпуса ArgMicro, содержащего короткие тексты на английском и немецком языках. Была проведена классификация предложений корпуса по позиции («за»/«против») различными алгоритмами.

Выводы. В ходе обзора было установлено, что анализ аргументации в текстах на русском языке предоставляет большое количество направлений исследований, связанных как с созданием новых корпусов и методов, так и с трансфером технологий и аннотации с других языков.

Осадчий Д. (автор)

Котельников Е. В. (научный руководитель)