

## РЕЦЕНЗИЯ

на статью Е.В.Третьяк

«Measuring Similarity of Fiction Texts Based on Distributional Semantic Models (Case Study of the Russian Original Text and English Translations of M.Bulgakov's Novel «The Master and Margarita»)», представленную к публикации в журнале «The International Journal of Open Information Technologies (INJOIT)»

Статья посвящена оценке близости англоязычных переводов романа М.А.Булгакова «Мастер и Маргарита» на основе дистрибутивно-семантических моделей. Автор развивает идею о том, что определение близости предложений можно осуществлять с использованием векторных представлений слов в многомерном признаковом пространстве корпуса текстов. Вектор предложения представляется как усредненный вектор слов, сравнение векторов пар предложений производится с учетом косинуса угла между ними. Параллельно с косинусной оценкой автор провел исследование методов перифразирования, с помощью которых можно определить близость предложений, при этом суммарные веса трансформаций позволили уточнить результаты работы с дистрибутивно-семантическими моделями. Автором были проведены серии экспериментов для трех типов предложений, различающихся длиной и степенью сложности. Полученные результаты были подвергнуты статистической оценке. Данные экспериментов были использованы для выделения лексических конструкций – синонимичных кандидатов в переводные эквиваленты, которые извлекались из пар предложений с высоким индексом близости.

Статья соответствует тематике издания «The International Journal of Open Information Technologies (INJOIT)», в ней описано проведенное практическое исследование, текст статьи подготовлен в соответствии с правилами издания. Статья рекомендуется к публикации в очередном выпуске журнала «The International Journal of Open Information Technologies (INJOIT)».

доцент кафедры математической лингвистики  
филологического факультета СПбГУ

О.А.Митрофанова

