

ОТЗЫВ

на автореферат диссертации Смирнова Ивана Валентиновича
«МЕТОДЫ РАЗНОУРОВНЕВОГО АНАЛИЗА ТЕКСТОВ НА
ЕСТЕСТВЕННОМ ЯЗЫКЕ И ИХ ПРИЛОЖЕНИЯ В СИСТЕМАХ
ИНФОРМАЦИОННОГО ПОИСКА И ПСИХОЛИНГВИСТИЧЕСКИХ
ИССЛЕДОВАНИЯХ»,

представленной на соискание ученой степени доктора технических наук по специальности 1.2.1 «Искусственный интеллект и машинное обучение»

Диссертация И.В. Смирнова посвящена разработке новых методов анализа текстов на естественном языке для повышения качества информационного поиска и психолингвистических исследований. Актуальность темы диссертации не вызывает сомнений, поскольку тексты сегодня являются важнейшим источником информации, а возможности инструментов по их эффективному анализу и поиску, особенно для русского языка, ограничены.

В своей работе И. В. Смирнов проводит тщательное исследование существующих методов анализа текстов, предлагает ряд новых методов для получения из русскоязычного текста разноуровневых структур, несущих информацию о смысле и связности текста. Далее он показывает, как эти структуры могут быть использованы для повышения качества поиска по запросу, вопросно-ответного поиска, извлечения информации из текстов, определения психологических особенностей автора письменной речи. Созданные в рамках диссертации оригинальные программные средства поиска и анализа текстов нашли применение в системах текстовой аналитики и научных исследованиях, что подтверждает не только теоретическую, но и практическую значимость работы. Особо стоит выделить инструмент психоэмоционального анализа текстов, основанный на предложенных в диссертации разноуровневых психолингвистических показателях, который применим к анализу текстов социальных сетей и исследованию сетевых коммуникаций.

Результаты диссертации представлены в 51 научной публикации, включая монографию и статьи в изданиях из Перечня ВАК, и успешно апробированы на ведущих профильных конференциях.

Автореферат имеет понятную структуру и написан ясным языком.

Хочется обратить внимание на следующие замечания к автореферату:

1. Неясно, какие методы помимо опорных векторов применялись в главе 5 для автоматического определения депрессивности по текстам социальных сетей и применялись ли современные языковые модели.
2. В тексте упоминается разработанная библиотека ISANLP, включающая методы семантического и дискурсивного анализа текстов на русском языке. Однако описание этой библиотеки (в частности, её состав и функции) отсутствует.

Эти замечания не снижают достоинств представленной работы, которая полностью отвечает критериям, предъявляемым к диссертациям на соискание степени доктора наук согласно Положению о присуждении ученых степеней, а ее автор, Смирнов Иван Валентинович заслуживает присуждения ученой степени доктора технических наук по специальности 1.2.1 – «Искусственный интеллект и машинное обучение».

доктор технических наук,

ведущий научный сотрудник лаборатории № 25

Института проблем управления РАН

117997, г. Москва, Профсоюзная, 65

Телефон: +7(905) 750 70 42, почта: gubanov@ipu.ru

Дмитрий Алексеевич Губанов

15.04.2024



Подпись Д.А. Губанова заверяю

Подпись

ЗАВЕД.

ВЕД. ИНЖЕНЕР

ГОРДЕЦВА Ю.Ю.