

ОТЗЫВ

официального оппонента на диссертацию
Андриевской Наталии Климовны на тему: «Совершенствование моделей и алгоритмов обработки информации в системах организационного сопровождения деятельности научно-образовательных учреждений»,
представленной на соискание ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

Актуальность темы исследования.

С развитием информационных технологий и распространением информационно-коммуникационной сети «Интернет» существенно возросли объемы информации в свободном доступе. Извлечение и накопление новых знаний является актуальной проблемой для любой организации, занимающейся научной или образовательной деятельностью. Поэтому разработка качественных моделей и алгоритмов обработки информации в системах организационного сопровождения деятельности научно-образовательных учреждений, формирование стандартизованного информационного каркаса, позволяющего ориентироваться на общие концепции научно-исследовательской и образовательной предметной области, который может быть использован в качестве выразительной и доказуемой модели для поддержки и извлечения научной и системной информации об образовательных и учебных ресурсах, а также для проверки, увязки и согласования данных имеет большую научно-практическую ценность.

Степень научной новизны результатов, полученных автором.

Новизна разработанных в диссертации научных положений, выводов и рекомендаций состоит в следующем.

1. Впервые разработана онтологическая модель научно-образовательной деятельности сотрудников вуза.

2. Усовершенствована гибридная мера определения семантической близости на базе модифицированной N-мерной модели представления знаний RDF-графа, использование которой повысило качество поиска, выраженное F-мерой, на 10,7% по сравнению с мерой «косинусного сходства».

3. Получила дальнейшее развитие векторная модель представления текстов на базе известных подходов bag-of-words и bag-of-concepts, улучшенная за счет применения онтологии и тематической редукции векторного пространства, что позволило при уменьшении размерности пространства с 2250 терминов до 30 терминов повысить скорость выполнения тестируемых алгоритмов более чем на порядок при незначительном снижении меры семантической близости на 6,2%.

4. Усовершенствована модель классификации данных, основанная на применении гибридной меры определения семантической близости, что привело к повышению качества классификации, выраженного F-мерой, по

сравнению с алгоритмами, использующими меру, вычисленную только по онтологии на 45,4%, «косинусную» меру – на 5,3% и «мягкую косинусную» меру – на 9,5%.

С формулировками пунктов научной новизны, полученных автором результатов, можно согласиться.

Теоретическая значимость работы.

Тщательный анализ проблем, возникающих при работе с информационными ресурсами (ИР), позволил автору выделить основные направления научных исследований и разработать математический аппарат, включающий модели и алгоритмы обработки, хранения и доступа к информации о научных и учебно-методических разработках, представленных слабоструктурированными фондами.

В работе автором предложена онтологическая модель научно-образовательной деятельности сотрудников. Разнообразие предложенных способов пополнения разработанной онтологии из хорошо зарекомендовавших себя источников и возможность участия в работе над онтологией высококвалифицированного эксперта гарантирует качество структуры и содержания полученной модели.

Теоретическая часть работы также содержит усовершенствованные алгоритмы и модели обработки текстов, применение которых дает лучшие результаты при поиске и классификации информации по сравнению с традиционно используемыми для этих операций. Следует отметить оригинальный подход автора к использованию трехмерного тензора семантических связей со значениями в диапазоне $[0;1]$ в качестве графовой модели представления знаний. Такой подход позволяет получать эффективную оценку семантической близости между двумя концептами через факторизацию срезов тензора. Тщательно и обоснованно подобраны параметры, используемые при расчете меры семантической близости. Для снижения размерности исходного признакового пространства и снижения вычислительной сложности алгоритмов обработки автором были исследованы две редуцированные модели представления текстовых информационных ресурсов, экспериментальными исследованиями обоснована целесообразность применения одной из них – модели без трансформации признакового пространства. Модель формирования семантически обогащенного поискового запроса, предложенная автором, повышает качество извлечения релевантных данных, а разработанный алгоритм классификации с применением гибридной меры семантической близости ИР позволяет лучше их структурировать для последующего доступа и управления.

Применение предложенных моделей и алгоритмов даст возможность эффективнее использовать накопленные информационные ресурсы, связанные с научными разработками и образовательной деятельностью.

Практическая значимость работы.

Разработанные архитектурные модели системы управления информационными ресурсами и реализованный на их базе программный комплекс, который представляет из себя полноценный программный продукт, готовый к применению в условиях научно-образовательной организации, подтверждают работоспособность предложенных моделей и алгоритмов и позволяют решать следующие задачи:

- учет документации, связанной с научными разработками, причем документация может быть представлена файлами в разных форматах в структурированном и неструктурированном виде;
- автоматический разбор текста;
- тематическая классификация документации и семантический анализ текстов;
- семантический поиск информации.

Полученные результаты могут быть использованы любыми образовательными, научно-исследовательскими, проектными организациями, предприятиями, имеющими необходимость в эффективной организации работы с информационными ресурсами.

Достоверность и обоснованность научных результатов, выводов и рекомендаций, сформулированных в диссертации.

Достоверность и обоснованность результатов проведенного научного исследования обеспечивается целостным и комплексным подходом к научному исследованию, корректным применением используемых методов и результатами математического моделирования и программного тестирования, апробацией основных результатов в 6 рецензируемых сборниках ВАК ДНР, в 4 других изданиях и 4 научных конференциях, а также включенностью результатов в учебный процесс ГОУВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ», внедрением в ГУ «Автоматгормаш им. В.А. Антипова», о чем свидетельствует справка о внедрении № 12-319 от 9.06.2021 г.

Обзор работы.

Представленная на рецензию диссертационная работа содержит 199 страниц машинописного текста и состоит из введения, четырех разделов, заключения, списка литературы из 141 источника и 4 приложений.

Первый раздел посвящен описанию объекта исследования, результатам проведенного анализа состояния вопроса, на базе чего и сформулированы цель и задачи исследования, обоснована необходимость совершенствования существующих и разработка новых подходов к сбору, хранению и обработке информационных ресурсов.

Во втором разделе работы представлена реализация онтологического подхода в информационных системах сопровождения деятельности научно-

образовательных учреждений с помощью разработанного гибридного подхода к формированию онтологии.

В третьем разделе работы представлено описание предложенной математической модели информационного поиска, которая состоит из нескольких моделей: представления текстов; формирования поисковых запросов; извлечения релевантных данных по запросу.

Четвёртый раздел посвящён разработке системы учета информационных ресурсов и описывает архитектурные модели фреймворка системы управления информационных ресурсов.

Структурно-логическое построение и качество оформления диссертации, стиль изложения в ней результатов работы замечаний не вызывают. Текст сопровождается большим объёмом иллюстрированного материала в виде схем, графиков, таблиц, а также рисунков с экранными скринами результатов работы алгоритмов.

Диссертационная работа в целом по структуре и оформлению соответствует требованиям, предъявляемым ВАК ДНР к кандидатским диссертациям.

Автореферат отражает основную суть диссертационной работы. В нем изложены основные идеи, выводы диссертации, степень новизны, практическая значимость приведенных результатов исследований.

Замечания.

Вместе с тем есть несколько замечаний к тексту диссертации и автореферату:

1. Диссертационная работа имеет не совсем стандартную структуру, как правило, второй раздел должен быть посвящен описанию моделей, но не их реализации.
2. Отсутствует список условных обозначений и сокращений.
3. В разделе 2.3 не описан подход, основанный на обработке произвольных источников с помощью NLP-технологий с целью обнаружения контекстно-зависимых слов, относящихся к одной предметной области.
4. Не описаны модели уровня «извлечения и производства знаний» – одного из заявленных 6 уровней (стр. 91) при описании модели управления знаниями.
5. В формулах 3.6 и 3.7 диссертации (стр. 122) нет знака операции и ее описания, что нарушает требования к оформлению формул по ГОСТ.
6. В подразделе 4.1.5 не приведено обоснование выбора веб-браузера (Opera) и серверов данных и приложений.
7. Информация по разделам в автореферате представлена неравномерно, первый раздел изложен слишком сжато.

8. Из текста автореферата не совсем понятно, в чем заключается модификация модели Bag-of-concepts, так как не приведено ее отличие от существующей.

В целом отмеченные замечания не снижают общего положительного впечатления от работы, а также научной и практической ценности полученных результатов диссертационной работы.

Соответствие диссертации критериям «Положения о присуждении ученых степеней» и паспорту научной специальности.

Диссертация Андриевской Н.К. соответствует требованиям «Положения о присуждении ученых степеней», утвержденного постановлением Совета Министров Донецкой Народной Республики №2-13 от 27 февраля 2015 года:

- пункту 2.2 – диссертация является научно-квалификационной работой;
- пункту 2.3 – диссертация написана автором самостоятельно, обладает внутренним единством, содержит новые научные результаты и положения, представляемые для публичной защиты, и подтверждает личный вклад автора диссертации в науку;
- пункту 2.4 – диссертационная работа содержит две справки о внедрении, в которых подтверждается использование полученных научных результатов в учебном процессе и на предприятии;
- пункту 2.6 – основные научные результаты диссертации опубликованы в 10-ти работах в рецензируемых научных изданиях;
- пункту 2.11.2 – основные научные результаты диссертации опубликованы в 6-ти статьях в рецензируемых научных изданиях, рекомендованных Министерством образования и науки Донецкой Народной Республики;
- пункту 2.13 – в диссертационной работе соискатель ученой степени кандидата технических наук Андриевская Н.К. ссылается на авторов и источники заимствования материалов и отдельных результатов.

Тема и содержание диссертационной работы соискателя Андриевской Н.К. соответствует паспорту научной специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки) в частности:

- пункту 4 – «Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации»;
- пункту 5 – «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации»;
- пункту 8 – «Теоретико-множественный и теоретико-информационный анализ сложных систем».

Заключение по диссертационной работе.

Диссертационная работа Андриевской Наталии Климовны выполнена на актуальную тему, является законченной научно-квалификационной работой, которая содержит обоснованные и достоверные результаты и научные выводы. Выводы и рекомендации в работе являются достаточно обоснованными. Указанные в настоящем отзыве замечания не являются определяющими и не снижают общей положительной оценки диссертационной работы.

Профиль диссертации соответствует специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки), а сама работа соответствует требованиям ВАК Донецкой Народной Республики, предъявленным к кандидатским диссертациям и удовлетворяют требования п. 2.2 «Положения о присуждении ученых степеней», утвержденного Постановлением Совета Министров Донецкой Народной Республики от 27.02.2015 г. №2-13.

По моему мнению, Андриевская Наталия Климовна заслуживает присуждения степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

Официальный оппонент,

кандидат технических наук по специальности 05.13.23 – «Системы и средства искусственного интеллекта», доцент по специальности 05.13.01 – «Системный анализ, управление и обработка информации (по отраслям)», доцент кафедры компьютерных технологий ГОУ ВПО «Донецкий национальный университет» _____ Т.В. Ермоленко

283001 г. Донецк,
ул. Университетская 24,
тел. (062) 302 06 00
Факс (062) 302 07 49
эл. почта: t.ermolenko@donnu.ru
Сайт: www.donnu.ru

Я, Ермоленко Т.В., даю согласие на автоматизированную обработку персональных данных, приведенных в этом документе.

_____ Т.В. Ермоленко

Подпись Т.В. Ермоленко заверяю:



УЧЕНЫЙ СЕКРЕТАРЬ
М.Н. МИХАЛЬЧЕНКО