

## 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)

### 1. Актуальность темы исследования

Лавинообразное нарастание массы разнообразной информации в современном мире неизбежно приводит к информационной перегрузке, как человека, так и общества в целом. Фильтрация информационного шума без потери семантики информационных потоков является одной из наиболее актуальных тем для человечества. Эффективная организация извлечения релевантной информации из глобальных или локальных хранилищ позволит облегчить ее восприятие и преобразование в знание.

Стремительное развитие методов искусственного интеллекта и машинного обучения позволяет значительно улучшить качество информационного поиска и классификации информации. Однако использование и развитие методов машинного обучения требует огромных вычислительных ресурсов, недоступных среднестатистической бюджетной организации, что обуславливает необходимость разработки альтернативных подходов к решению задач поиска и классификации информации.

Таким образом, совершенствование моделей и алгоритмов обработки информации для реализации в системах организационного сопровождения деятельности научно-образовательных учреждений, безусловно, является актуальной научно-технической задачей, имеющей важное значение практически во всех отраслях.

### 2. Обоснованность научных положений, выводов и рекомендаций

Обоснованность научных положений, выводов и рекомендаций достаточно высокая и подтверждается:

- полнотой анализа теоретических и практических исследований отечественных и зарубежных ученых в области информационных технологий и тщательным изучением предмета исследования посредством изучения большого числа литературных источников;

- корректным применением, развитием известных методов и подходов при реализации информационных процессов и концепции информационного поиска в системах организационного сопровождения деятельности научно-образовательных учреждений;

- применением апробированных методов онтологического проектирования, математического моделирования и системного анализа;

- результатами экспериментальных исследований разработанных методов, моделей и алгоритмов в процессе их программной реализации, тестирования.

Достоверность научных положений, выводов и практических результатов подтверждается разработкой и тестированием программного модуля системы, о

чем свидетельствуют справки о внедрении, выполненными публикациями и положительной оценкой на научно-технических конференциях.

С учетом вышесказанного, научные положения и выводы диссертационной работы следует считать научно-обоснованными.

### **3. Характеристика работы**

Диссертационная работа содержит 199 страниц машинописного текста и состоит из введения, четырех разделов, заключения, списка литературы из 141 источника и 4 приложений. Основной текст диссертации иллюстрируется 82 рисунками и содержит 47 таблиц.

Во введении обоснована актуальность, сформулирована научная новизна, практическая ценность работы, дана ее общая характеристика.

**В первом разделе работы** «Анализ состояния вопроса, цель и задачи исследования» описывается необходимость совершенствования существующих и разработка новых подходов к сбору, хранению и обработке информационных ресурсов:

- в подразделе 1.1. (с 12-15) – описаны особенности информационных ресурсов как объектов исследований;

- в подразделе 1.2 (с. 15-18) - выполнен обзор и проведен анализ систем управления знаниями научно-образовательных организаций;

- в подразделе 1.3 (с. 18-24) - проанализированы основные модели, позволяющие представлять как текст в общем, так и знания;

- в подразделе 1.4 (с. 24-43) – осуществлен обзор методов оценки качества поиска и классификации информации, определены их узкие места и проблемы, требующие совершенствования.

- в подразделе 1.5 (с. 43-45) – на основе анализа поставлена цель диссертационной работы, сформулированы основные задачи и выбраны основные направления для достижения поставленной цели

- в подразделе 1.6 (с.45) - сформулированы выводы по разделу.

**Во втором разделе** разработана онтологическая модель, учитывающая специфику профессиональной деятельности сотрудников научно-образовательных организаций:

- в подразделе 2.1 (с. 46-48) приведено формальное описание онтологии;

- в подразделе 2.2 (с. 48-64) разработана «базовая» онтологическая модель системы;

- в подразделе 2.3 (с. 64-84) разработаны алгоритмы наполнения онтологической модели из разнородных источников информации, при формировании онтологии использовались как словарные способы пополнения, так и автоматическое извлечение информации из корпуса документов;

- в подразделе 2.4 (с. 84-75) описан онтологический подход к разработке информационной системы;

- в подразделе 2.5 (с. 88) сформулированы выводы по разделу.

**Третий раздел** описывает разработку моделей, методов и алгоритмов выявления, приобретения и классификации знаний:

– подраздел 3.1 (с. 89–92) посвящен разработке концептуальной метамодели для учета связанности знаний и обеспечения однородности представления данных в рамках единой тематики проектируемой системы, ядром которой является онтология;

– в подразделе 3.2 (с. 92–95) модифицирована модель N-мерного представления знаний на базе RDF-графа в виде трехмерного тензора семантических связей;

– в подразделе 3.3 (с. 96–105) усовершенствована гибридная мера оценки семантической близости на базе модифицированной модели N-мерного представления RDF-графа знаний, что дало возможность определять сходство с учетом семантики, частотных характеристик текста, контекста и структуры онтологии и улучшить качество поиска разработанной модели;

– в подразделе 3.4 (с. 106-125) разработаны модели, методы и алгоритмы уровней выявления, поиска и извлечения знаний, разработана модель представления текстовых информационных ресурсов, которая использует редукцию по тематическим разделам онтологии и существенно снижает размерность задачи;

– в подразделе 3.5 (с. 125-132) разработаны модели, методы и алгоритмы уровней интеграции и хранения данных, в том числе и модель классификации, которая позволяет осуществлять эффективную классификацию документов, относящихся к предметной области;

– в подразделе 3.6 (с. 132-134) сформированы выводы по разделу, представленные результаты проведения экспериментальных исследований подтверждают повышение качества поиска при применении разработанных моделей.

**В четвертом разделе** разработана система учета информационных ресурсов сотрудников научно-образовательных организаций:

– в подразделе 4.1 (с. 134-143) приведена разработка структурной архитектурной модели фреймворка, комплекс полученных моделей системы в полной мере отражает ее архитектурные особенности на всех уровнях проектирования и развертывания;

– в подразделе 4.2 (с. 144-147) выполнена разработка на базе фреймворка системы учета информационных ресурсов научно-образовательной деятельности сотрудников вуза;

– в подразделе 4.3 (с. 147-165) описана разработка и тестирование программного модуля учета научно-исследовательской деятельности сотрудников кафедры вуза «Наука», тестирование подтвердило работоспособность и корректность разработанных моделей и алгоритмов обработки информации;

– в подразделе 4.4 (с. 165-166) сформированы выводы по разделу.

В заключении (с. 166-169) перечислены решения поставленных научных задач и выполненных программных разработок, содержащихся в диссертации, совокупность которых позволили автору достичь поставленных целей исследования.

#### **4. Новизна научных положений, выводов и рекомендаций**

Научная новизна полученных результатов заключается 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%.
5. 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%.

#### **5. Ценность результатов исследований**

Теоретическая значимость результатов исследований заключается в развитии моделей и алгоритмов обработки информационных ресурсов научно-образовательных учреждений и переходу к онтологическому и семантическому моделированию.

Практическое значение результатов исследований:

1. На основе проведенных ранее исследований, разработанных моделей и алгоритмов выполнена программная реализация системы управления ИР в рамках кафедры вуза.
2. Использование документированной прикладной онтологии дает возможность разработчикам систем повторно использовать и развивать данную онтологию, а различным ИС – интегрировать данные и обеспечивать обмен данными на основе онтологии.
3. Предложенные подходы и математические модели, разработанные программные модули фреймворка могут быть применены при создании различных СУЗ, а также систем обработки ИР любых учреждений.
4. Разработанные в ходе выполнения диссертационной работы модели и методы использованы в учебном процессе кафедры автоматизированных систем управления ГОУВПО «ДОННТУ» при выполнении курсовых работ и выпускных квалификационных работ студентов.
5. Разработанный программный модуль «Наука» успешно прошел тестирование в ГУ «Автоматгормаш им. В.А. Антипова» (г. Донецк) в условиях отдела систем управления.

## **6. Анализ публикаций**

По теме диссертационной работы всего было опубликовано 10 научных работ. Из них 6 работ в изданиях, рекомендованных ВАК ДНР. Результаты теоретических исследований опубликованы в научных журналах «Информатика и кибернетика» (ДонНТУ) и «Вестник ДонНУ Сер.Г:Технические науки», а также в Международном научно-теоретическом журнале «Проблемы искусственного интеллекта» (г. Донецк). Основные положения диссертационной работы апробированы на четырех научно-технических конференциях.

## **7. Замечания к диссертационной работе**

1. В разделе 1 не уделено внимание нейросетевым подходам несмотря на то, что они являются широко распространёнными и перспективными при решении задачи классификации текстов.

2. Не обосновано использование термина «весовые функции» для обозначения мер TF и TF-IDF (стр. 29).

3. В разделе 2 не рассмотрены вопросы оценки качества полученной онтологической модели.

4. В разделе 3.2 стр. 93 недостаточно проработаны методы предобработки текстов, включающие проверку орфографии и идентификации частей речи.

5. На рисунке стр. 85 указан модуль проверки на уникальность, но в дальнейшем о нем нет упоминаний.

6. В тексте диссертации и автореферата имеются незначительные стилистические и орфографические неточности, например на рисунке стр. 87 «описания» вместо «описание», а также встречаются отдельные недостаточно строгие формулировки.

7. Алгоритмы методов для наглядности следовало бы привести в виде блок-схем.

8. Модель размещения компонентов системы отражает размещение в условиях определенной инфраструктуры научно-образовательной организации. Из работы не ясно, какое ПО является обязательным при развертывании в пределах других организаций.

## **8. Заключение о соответствии диссертации паспорту научной специальности**

Тема и содержание диссертации Андриевской Н.К. соответствует паспорту специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

Полученные результаты, положения и выводы отвечают требованиям паспорта специальности в частности следующих пунктов: п.4 «Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации»; п.5 «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации»; п.8 «Теоретико-множественный и теоретико-информационный анализ сложных систем».

## **9. 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, а именно:

– пункту 2.2 – диссертация является научно-квалификационной работой, в которой решена актуальная научно-техническая задача повышения эффективности системы управления информационными ресурсами научно-образовательных учреждений за счет применения онтологического подхода, разработки новых и усовершенствования существующих моделей и алгоритмов поиска, хранения и классификации данных;

– 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 два Акта внедрения, в которых подтверждается использование полученных научных результатов в учебном процессе и на предприятии ГУ «Автоматгормаш им. В.А. Антипова» (г. Донецк);

– пункту 2.6 – основные научные результаты диссертации опубликованы в 6-ти работах в рецензируемых научных изданиях;

– 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.

## **10. Выводы**

В целом выявленные недостатки и сделанные замечания не умаляют достоинств представленной работы.

Диссертация Андриевской Наталии Климовны на тему «Совершенствование моделей и алгоритмов обработки информации в системах организационного сопровождения деятельности научно-образовательных учреждений является законченным научно-исследовательским трудом, выполненным автором на высоком научном уровне.

Задача актуальна, решение обладает научной новизной, представленные научные результаты получены на основе глубокого анализа по исследуемой тематике и являются достаточно достоверными, выводы по каждому разделу диссертации четкие, тесно связаны с содержанием и отражают суть проведенных исследований, заключения обоснованы.

Работа написана доходчиво, грамотно, хорошим научным языком и аккуратно оформлена, проиллюстрирована большим количеством рисунков. Иллюстрации достаточно информативны. Все обозначения, используемые в тексте диссертации и, в частности, в формулах имеют пояснения. Диссертация оформлена аккуратно. Содержание автореферата соответствует содержанию диссертации.

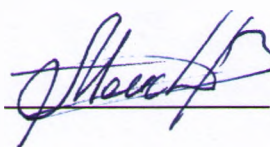
Все пункты новизны находят свое раскрытие в опубликованных автором работах, и представленная диссертация соответствует специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

Работа соответствует требованиям ВАК Донецкой Народной Республики, предъявленным к кандидатским диссертациям, и удовлетворяют требованиям п. 2.2 «Положения о присуждении ученых степеней», утвержденного Постановлением Совета Министров Донецкой Народной Республики от 27.02.2015 г. №2-13.

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,

доктор технических наук,  
профессор кафедры «Информационные  
технологии и компьютерные системы»  
ФГАОУ ВО «Севастопольский  
государственный университет»

  
Д.В. Моисеев

Адрес: 299053, Российская Федерация,  
г. Севастополь, ул. Университетская, 33,  
тел.: +7(8692) 435-002,  
факс: +7(8692) 243-590  
эл. почта: info@sevsuu.ru  
сайт: http://sevsu.ru

Я, Моисеев Дмитрий Владимирович, даю согласие на автоматизированную обработку моих персональных данных.

  
Д.В. Моисеев

Подпись д-ра наук Моисеева Д.В.  
удостоверяю:  
начальник отдела кадров



  
С.В. Бернацкая