

2

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.073.01 НА БАЗЕ
ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО УЧРЕЖДЕНИЯ
НАУКИ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ РОССИЙСКОЙ
АКАДЕМИИ НАУК, ФАНО РОССИИ, ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ
УЧЕНОЙ СТЕПЕНИ КАНДИДАТА ТЕХНИЧЕСКИХ НАУК

аттестационное дело № _____

решение диссертационного совета от 28 января 2015 года № 1

О присуждении Толкунову Александру Александровичу, гражданину Российской Федерации, ученой степени кандидата технических наук.

Диссертация «Модель оперативной аналитической обработки текстовых комментариев к законопроектам» по специальности 05.13.17 – Теоретические основы информатики принята к защите 7 ноября 2014 года, протокол № 5 диссертационным советом Д 002.073.01 на базе Федерального государственного бюджетного учреждения науки Институт проблем информатики Российской академии наук (ИПИ РАН), ФАНО России, 119333, Россия, г. Москва, ул. Вавилова, д. 44, корпус 2, приказ о создании диссертационного совета от 02.11.2012 № 714/нк.

Соискатель, Толкунов Александр Александрович, 1982 года рождения. В 2004 году соискатель окончил с отличием Государственное образовательное учреждение высшего профессионального образования Академию Службы специальной связи и информации при Федеральной службе охраны Российской Федерации (Академию Спецсвязи России). В 2013 году соискатель окончил адъюнктуру по очной форме обучения в Государственном казенном образовательном учреждении высшего профессионального образования Академии Федеральной службы охраны Российской Федерации (Академии ФСО России). Проходит военную службу по контракту в Академии ФСО России, г. Орёл, в должности преподавателя, майор.

Диссертация выполнена на кафедре «Автоматизированные информационные системы» Академии ФСО России. Научный руководитель – кандидат технических наук, доцент Кукушкин Александр Антонович, Академия ФСО России, начальник кафедры «Автоматизированные информационные системы».

Официальные оппоненты:

Хорошилов Александр Алексеевич, доктор технических наук, ИПИ РАН, отдел № 17, ведущий научный сотрудник;

Золотарёв Олег Васильевич, кандидат технических наук, доцент, Негосударственное образовательное учреждение высшего профессионального образования «Российский новый университет», г. Москва, доцент кафедры «Информационные системы в экономике и управлении», – дали положительные отзывы на диссертацию.

Ведущая организация – Федеральное государственное бюджетное учреждение науки Институт системного анализа Российской академии наук (ИСА РАН) в своем положительном заключении, подписанном Акимовой Галиной Павловной, кандидатом технических наук, ведущим научным сотрудником лаборатории № 9-7, и Арлазаровым Владимиром Львовичем, членом-корреспондентом РАН, доктором технических наук, заведующим лабораторией № 9-1, указала, что «диссертация представляет собой завершённую научно-квалификационную работу, в которой автором получено решение актуальной научной задачи по разработке модели обработки текстовых комментариев к законопроектам с использованием оперативной аналитической обработки текстов. Новые научные результаты, полученные диссертантом, имеют существенное значение для развития моделей представления знаний, содержащихся в слабоформализованных массивах мнений, и для совершенствования методов поддержки принятия управленческих решений в области законотворческой деятельности. Выводы и рекомендации достаточно обоснованы. Диссертационная работа соответствует требованиям пунктов 9, 10 Положения о присуждении ученых степеней, а ее автор заслуживает присуждения ему ученой степени кан-

дидата технических наук по специальности 05.13.17 – Теоретические основы информатики».

Соискатель имеет 8 опубликованных работ, все по теме диссертации, в том числе 3 работы в рецензируемых научных изданиях (без соавторства, общим объемом 1,2 авт. листа), 1 работа в научном журнале (без соавторства, 0,5 авт. листа), 4 работы общим объёмом 0,8 авт. листа – в сборниках трудов (из них без соавторства – 1), по теме работы получено 2 свидетельства о регистрации программ для ЭВМ в Роспатенте (в соавторстве). Наиболее значительные работы соискателя:

1. **Толкунов, А.А.** Модель оперативной аналитической обработки комментариев с операциями детализации на основе выявления типичных и полезных текстов / А.А. Толкунов // Научное обозрение. – 2013. – № 9. – С. 368–371 (*осуществлена постановка задачи и предложена модель обработки комментариев, полученных в ходе общественного обсуждения законопроектов. Приведены формальные выражения, описывающие преобразования, необходимые для эффективной детализации на основе типичных и полезных комментариев*).

2. **Толкунов, А.А.** Алгоритмы формирования выборок типичных и полезных комментариев в ходе обработки результатов общественного обсуждения законопроектов / А.А. Толкунов // Информационные системы и технологии. – 2014. – № 1(81). – С. 47–58 (*рассмотрены подходы к решению задачи повышения оперативности выделения высказываний при обработке массива комментариев интернет-пользователей, полученного в результате общественного обсуждения законопроекта. Предложены алгоритмы формирования выборок комментариев при заданном бюджете времени, учитывающие неравномерность распределения высказываний в комментариях. Приведены результаты экспериментов, показывающие эффективность предложенных алгоритмов в сравнении с существующим подходом к обработке массива комментариев*).

3. **Толкунов, А.А.** Математические основы и программная архитектура системы оперативной аналитической обработки комментариев к обсуждаемым законопроектам / А.А. Толкунов // Системы управления и информационные технологии. – 2013. – № 3.1(53). – С. 173–177 (*рассмотрен подход к построе-*

нию системы оперативной аналитической обработки комментариев, полученных в ходе общественного обсуждения законопроектов. Предложена методика применения системы. Приведены результаты экспериментов, доказывающие эффективность предлагаемой системы).

На диссертацию и автореферат прислали положительные отзывы:

1. Малахов И.Ю., к.т.н., научный консультант Управления обеспечения правовой информатизации Спецсвязи ФСО России. Отмечено, что результаты работы развивают представление о закономерностях в массивах мнений в отношении правовых актов и совершенствуют методы многомерного анализа текстовых массивов. Замечания: 1) из автореферата не ясно, требуется ли изменить существующие нормативно закреплённые процедуры общественного обсуждения законопроектов для применения предложенной в диссертации модели; 2) автор не отметил, изучалась ли возможность реализации разработанной модели на программно-технических средствах информационно-аналитических и ситуационных центров государственных органов.

2. Лазарев В.М., д.т.н., профессор, главный научный сотрудник ЗАО НТЦ «Поиск-ИТ». Он отметил, что предлагаемые решения не противоречат известным, основаны на адекватном математическом аппарате. Замечания: 1) в автореферате автор не привел определения понятий «выборка типичных комментариев», «выборка полезных комментариев», «типичное высказывание»; 2) на странице 8 указана операция сводного реферирования, однако не ясно, как она осуществляется на практике; 3) в структуре системы на рисунке 10 введено два блока оценки подобия, но не ясно, в чем различие механизмов их функционирования.

3. Пичкур А.Б., к.ф.-м.н., доцент, начальник Института криптографии, связи и информатики Академии ФСБ России. Он отметил, что к достоинствам работы можно отнести оригинальность предложенного подхода к обработке комментариев и внедрение ее результатов в практику деятельности аналитических подразделений ФСО России и в учебный процесс Академии

ФСО России. Замечания по автореферату: 1) показатели качества, используемые для оценки эффективности разработанной технологии обработки комментариев, в автореферате недостаточно подробно раскрыты с содержательной точки зрения. Так, показатель $r_v = \frac{q_v - k_v}{k_v}$ адекватности модели (стр. 15) принимает неположительные значения (так как всегда верно $0 \leq q_v \leq k_v$), тогда как по смыслу термина «показатель адекватности» его значения ожидаются положительными (неотрицательными); 2) при описании экспериментов по оценке эффективности разработанных алгоритмов (стр. 16-17) не приводятся сведения о количестве экспертов и размерах тестовых выборок; 3) предлагаемые автором алгоритмы обработки комментариев основываются на использовании алгоритмов кластерного анализа текстов. Однако в автореферате не описываются используемые меры близости текстов, используемые алгоритмы кластерного анализа и выбранные параметры их работы; 4) при описании модели оперативной аналитической обработки комментариев, а также алгоритмов обработки текстов в автореферате не описываются результаты сравнительного анализа различных вариантов их реализации.

4. Балунин Е., к.т.н., доцент, начальник кафедры вычислительной техники, Столяров М., к.т.н., заместитель начальника кафедры вычислительной техники, Новиков А., к.т.н., доцент, проф. кафедры вычислительной техники Череповецкого высшего военного инженерного училища радиоэлектроники. Они отметили, что рецензируемая работа содержит новые научно обоснованные технические решения, имеющие существенное значение для их использования в задачах, требующих оперативного анализа текстовых массивов. Замечания по автореферату: 1) в силу не использования в работе моделей семантической интерпретации текстов утверждений и комментариев применение разработанного подхода, как представляется, сохраняет значительную степень участия экспертов-аналитиков в контуре обработки мнений граждан, что, в свою очередь, не решает в полной мере проблему ручного труда и волюнтаризма в принятии окончательных решений по корректировке обсуждаемых законопроектов; кро-

ме того, замещающие семантику и традиционную логику обработки текстовых документов аналитические модели, как представляется, могут не быть столь же адекватным для решаемых задач инструментом, как упомянутые подходы; 2) не приведены оценки трудоемкости и затрат на внедрение полученных результатов; 3) не ясно, при каких условиях время ознакомления со всеми классами комментариев будет минимальным.

5. Константинов И.С., д.т.н., профессор, проректор по научной и инновационной деятельности Федерального государственного автономного образовательного учреждения высшего профессионального образования «Белгородский государственный национальный исследовательский университет». Он отметил, что диссертация вносит вклад в развитие моделей информационных процессов, а ее результаты будут полезны при проектировании систем и комплексов поддержки принятия решений в информационно-аналитических и ситуационных центрах государственных органов. Замечания по автореферату: 1) объем автореферата не позволяет понять, существует ли взаимно однозначное отображение классов комментариев и ячеек применяемой для их анализа модели гиперкуба; 2) из автореферата не ясны граничные объемы массива комментариев, для которых применение разработанной модели является целесообразным и эффективным.

6. Огарок А.Л., к.т.н., с.н.с., начальник отдела исследований и разработок в области наукоёмких технологий обработки данных, Управление перспективных технологий и поисковых разработок ФГАНУ ЦИТиС, Богданов Ю.М., к.т.н., с.н.с., директор ФГАНУ ЦИТиС по науке и информационным технологиям. Они отметили, что к достоинствам работы относится подробное описание экспериментов, достаточное количество графического и табличного материала, в полном объёме позволяющее оценить их выполнение и результаты. Замечания и рекомендации: при извлечении учитываемых автором в работе объектов, во-первых, следовало бы большее внимание уделить лингвистическому подходу, а во-вторых – отображению извлекаемых структур данных на онтологии.

Выбор официальных оппонентов и ведущей организации обосновывается их компетенцией и достижениями в соответствующей отрасли науки, наличием публикаций в сфере исследования.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

– разработаны научно обоснованные технические решения, имеющие существенное значение для развития страны;

– предложена и обоснована оригинальная гипотеза, состоящая в том, что результатом применения модели оперативной аналитической обработки текстовых комментариев к законопроектам, учитывающей многомерность представления мнения и неравномерность распределения высказываний по комментариям с использованием методов классификации, обработки текстов и оптимизации, является повышение темпа изучения совокупности высказываний при выполнении требований к качественным показателям процедуры обработки массива комментариев;

– разработана новая модель оперативной аналитической обработки комментариев, отличающаяся от известных способом формирования гиперкуба на основе многомерного представления мнения, а также набором операций погружения в текстовые данные, учитывающих неравномерность распределения высказываний в комментариях интернет-пользователей;

– разработан комплекс алгоритмов, включающий алгоритм формирования выборки типичных комментариев, учитывающий содержательное подобие комментариев и сходство с типичным представителем мнения для приоритетного отбора комментариев, и алгоритм формирования выборки полезных комментариев, учитывающий атрибуты конструктивных предложений и предпочтения интернет-пользователей для приоритетного отбора комментариев.

Теоретическая значимость исследования обоснована тем, что:

– изучены закономерности в массивах текстовых комментариев к законопроектам, расширяющие возможности моделей и методов представления и обработки знаний, содержащихся в слабоформализованных массивах мнений;

– применительно к проблематике диссертации результативно использован комплекс существующих методов исследования на основе математической статистики, теории оптимизации, моделей представления знаний, моделей семантической фильтрации текстов, методов классификации, кластеризации и реферирования текстов;

– проведена модернизация моделей оперативной аналитической обработки текстов введением операций, обеспечивающих качественную детализацию на основе выборок по нескольким критериям.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

– разработаны и внедрены в практику деятельности ведомственных аналитических подразделений предложения по построению и применению системы оперативной аналитической обработки комментариев (основные результаты диссертационной работы реализованы в в/ч 63166 и Академии ФСО России, что подтверждается соответствующими актами реализации и внедрения);

– определены пределы и перспективы практического использования разработанных модели и алгоритмов, создана система практических рекомендаций по их технической реализации;

– создана модель эффективной обработки слабоформализованных массивов мнений в отношении законопроектов с возможностью учета временных ограничений.

Оценка обоснованности и достоверности результатов исследования выявила:

– результаты получены с использованием лицензионного программного обеспечения, а именно: MS Windows XP, MS Excel 2007, STATISTICA 6,

Mathcad 14, Delphi 7, C++Builder 6; зарегистрированных в ФГУ ФИПС программ для ЭВМ: «StyleClassifier», «Тонус», «Программа семантической фильтрации текстов», «Comment Eliminator»; показана воспроизводимость результатов исследования в различных условиях: разные законопроекты, разные по числу обсуждаемых пунктов и мощности классы комментариев;

– идея работы базируется на анализе практики и обобщении передового опыта обработки комментариев комиссиями экспертов по итогам общественного обсуждения законопроектов «О полиции» (7 августа – 15 сентября 2010 г.) и «Об образовании» (1 декабря 2010 г. – 1 февраля 2011 г.);

– теоретическая часть построена на известных, проверяемых данных, размещенных на интернет-порталах zakonoproekt2012.ru и regulation.gov.ru и полностью согласуется с опубликованными экспериментальными данными по теме диссертации;

– использованы современные методики обработки исходной информации, включающие классификацию, кластерный анализ, семантическую фильтрацию, автоматическое реферирование, сортировку, дисперсионный анализ, аппроксимацию, оптимизацию методом множителей Лагранжа; представительные выборочные совокупности, сформированные на основе анализа массивов комментариев двух масштабных обсуждений общим объемом около 30 000 текстов с обоснованием подбора объектов наблюдения и измерения на основе анализа числа пунктов и подпунктов в статьях обсуждаемых законопроектов.

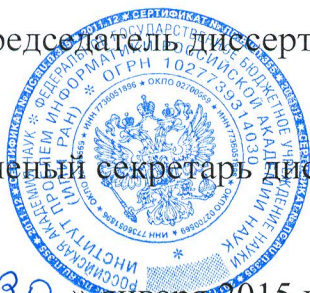
Личный вклад соискателя состоит в:

непосредственном участии на всех этапах диссертационного исследования; активном участии соискателя в получении исходных данных и проведении научных экспериментов; апробации и обосновании полученных результатов исследований; обработке и интерпретации экспериментальных данных, выполненных лично автором; подготовке основных публикаций по выполненной работе.

На заседании 28 января 2015 года диссертационный совет принял решение присудить Толкунову А.А. ученую степень кандидата технических наук.

При проведении тайного голосования диссертационный совет в количестве 20 человек, из них 7 докторов наук по специальности рассматриваемой диссертации, участвовавших в заседании, из 27 человек, входящих в состав совета, проголосовали: за - 20, против - нет, недействительных бюллетеней - нет.

Председатель диссертационного совета



[Handwritten signature in blue ink]

Соколов И.А.

Ученый секретарь диссертационного совета

[Handwritten signature in blue ink]

Гринченко С.Н.

« 30 » января 2015 года

М.П.