

«УТВЕРЖДАЮ»
Директор Федерального
государственного учреждения
«Федеральный исследовательский центр
«Информатика и управление»
Российской академии наук»



ЗАКЛЮЧЕНИЕ

Федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук»

Диссертация Смирнова Ивана Валентиновича на тему «Методы разноуровневого анализа текстов на естественном языке и их приложения в системах информационного поиска и психолингвистических исследованиях» выполнена в Отделе №72 «Интеллектуальный анализ информации» отделения №7 Федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН).

Смирнов Иван Валентинович, 1978 года рождения, гражданин России, в 2003 году закончил Российский университет дружбы народов по специальности «Прикладная математика и информатика». В 2008 году в Институте системного анализа Российской академии наук защитил диссертацию на соискание степени кандидата физико-математических наук на тему «Исследование методов установления значений синтаксических единиц естественных языков на основе интеллектуального анализа данных». В 2015 году И.В. Смирнову присвоено звание доцента. С 2005 по 2018 год Смирнов И.В. работал в ИСА РАН, с 2018 по настоящее время работает в ФИЦ ИУ РАН.

В период подготовки диссертации соискатель Смирнов Иван Валентинович работал в отделе №72 ФИЦ ИУ РАН в должности заведующего отделом с 2018 по 2023 гг.

По итогам рассмотрения диссертации принято следующее заключение.

1. Разработка методов разноуровневого анализа текстов на естественном языке и их применение в информационном поиске и психолингвистических

исследованиях является актуальной и практически востребованной научной задачей, требующей для своего решения новых научных подходов.

2. Все результаты диссертационной работы получены автором самостоятельно либо под его непосредственным руководством.

3. Научная новизна результатов диссертации заключается в следующем:

- Разработаны новые методы семантического анализа текстов на русском языке, которые позволяют представить текст в виде множества слов с семантическими ролями и отношениями. Разработанные методы позволяют повысить качество интеллектуального анализа текстов за счет учета семантической структуры высказываний и ситуативных значений слов в предложениях.

- Впервые разработаны методы автоматического дискурсивного анализа текстов на русском языке, которые позволяют представить структуру текста в виде дерева риторического разбора. Такие методы расширяют возможности автоматической обработки текстов за счет учета связности высказываний.

- Разработаны новые принципы и алгоритмы информационного и вопросно-ответного поиска текстов, основанные на сопоставлении семантических структур текстов запроса и документов и обладающие за счет этого более высоким качеством по сравнению с подходами, не учитывающими семантику.

- Предложены новые разноуровневые психолингвистические текстовые показатели, отражающие взаимосвязь письменной речи и психологических особенностей авторов на уровне морфологии, синтаксиса, семантики и дискурса. Предложенные показатели интерпретируемы и позволяют повысить качество автоматической классификации текстов, написанных людьми с различными психологическими особенностями.

- Разработаны новые программные средства лингво-статистического анализа больших массивов текстов, позволяющие учитывать разноуровневую структуру текста при корпусных исследованиях. Разработаны новые программные средства психоэмоционального анализа текстов, позволяющие на основе разноуровневых структур выявлять более широкий, по сравнению с аналогами, набор психоэмоциональных характеристик.

4. Практическая значимость результатов диссертационной работы подтверждается их применением в ряде промышленных систем поиска и анализа текстов, используемых для решения прикладных задач текстовой аналитики в различных организациях, о чем имеются акты об использовании.

5. Достоверность результатов диссертационной работы подтверждается использованием надежных и апробированных методов исследования, согласованностью результатов проведенных экспериментальных исследований с известными теоретическими положениями в области интеллектуального анализа текстов. Результаты прошли апробацию на

международных и российских конференциях, представлены многочисленными публикациями с высокой цитируемостью.

6. Результаты диссертационного исследования соответствуют паспорту специальности 1.2.1 «Искусственный интеллект и машинное обучение», а именно:

- пункту 4 «Разработка методов, алгоритмов и создание систем искусственного интеллекта и машинного обучения для обработки и анализа текстов на естественном языке, для изображений, речи, биомедицины и других специальных видов данных»;
- пункту 5 в части «Методы и технологии поиска, приобретения и использования знаний и закономерностей, в том числе – эмпирических, в системах искусственного интеллекта»;
- пункту 7 в части «Разработка специализированного математического, алгоритмического и программного обеспечения систем искусственного интеллекта и машинного обучения».

7. Основное содержание и результаты диссертации опубликованы автором в 51 научной публикации в рецензируемых изданиях, из них 1 рецензируемая монография, 18 публикаций в рецензируемых научных изданиях, отнесенных к категориям К-1 или К-2 из Перечня ВАК, 21 публикация в изданиях, индексируемых в БД Scopus и Web of Science. В рамках работы по диссертации получены 1 патент на изобретение РФ, 1 патент на полезную модель и 8 свидетельств о регистрации программы для ЭВМ. Наиболее значимые публикации:

- Смирнов И.В. Интеллектуальный анализ текстов на основе методов разноуровневой обработки естественного языка. – М.: ФИЦ ИУ РАН, 2023. – 356 с. (РИНЦ)

- Смирнов И.В. Разноуровневая обработка естественного языка для интеллектуального поиска и анализа текстов // Искусственный интеллект и принятие решений – 2023. – №1. – С.90-99. (К-1, RSCI)

- Смирнов И.В. Программные средства психоэмоционального анализа текстов // Информационные технологии и вычислительные системы. – 2023. – №1. – С.27-38. (К-1, RSCI)

- Чистова Е. В., Ларионов Д. С., Шелманов А. О., Латыпова Е. А., Смирнов И. В. Открытое извлечение информации из текстов. Часть III. Система вопросно-ответного поиска // Искусственный интеллект и принятие решений. – 2021. – №4. – С. 35-49. (К-1, RSCI)

- Шелманов А.О., Девяткин Д.А., Исаков В.А., Смирнов И.В. Открытое извлечение информации из текстов. Часть II. Извлечение семантических отношений с помощью машинного обучения без учителя // Искусственный интеллект и принятие решений. – 2019. – №2. – С. 39-49. (К-1, RSCI)

- Станкевич, М. А., Игнатъев, Н. А., Смирнов, И. В., Кисельникова, Н. В. Выявление личностных черт у пользователей социальной сети Вконтакте // Вопросы кибербезопасности. – 2019. – №. 4. – С. 80-87. (К-1, RSCI)
- Кузнецова, Ю. М., Смирнов, И. В., Станкевич, М. А., Чудова, Н. В. Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований. Часть 2. Машина РСА и опыт ее использования // Искусственный интеллект и принятие решений. – 2019. – №. 3. – С. 40-51. (К-1, RSCI)
- Ениколопов, С. Н., Кузнецова, Ю. М., Смирнов, И. В., Станкевич, М. А., Чудова, Н. В. Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований. Часть 1. Методические и методологические аспекты // Искусственный интеллект и принятие решений. – 2019. – №. 2. – С. 28-38. (К-1, RSCI)
- Шелманов А.О., Исаков В.А., Станкевич М.А., Смирнов И.В. Открытое извлечение информации из текстов. Часть I. Постановка задачи и обзор методов // Искусственный интеллект и принятие решений. – 2018. – №2. – С. 47-67. (К-1, RSCI)
- Суворов Р.Е., Шелманов А.О., Каменская М.А., Смирнов И.В. Активное машинное обучение в задаче извлечения информации из научных текстов. // Искусственный интеллект и принятие решений. – 2017. – №4. – С. 40-52. (К-1, RSCI)
- Осипов Г.С., Смирнов И.В. Семантический анализ научных текстов и их больших массивов // Системы высокой доступности. – 2016. – №1. – С.41-44. (К-2)
- Шелманов А.О., Каменская М.А., Ананьева М.И., Смирнов И.В. Семантико-синтаксический анализ текстов в задачах вопросно-ответного поиска и извлечения определений // Искусственный интеллект и принятие решений. – 2016. – №4. – С. 47-61. (К-1, RSCI)
- И.В. Смирнов, А.О. Шелманов, Е.С. Кузнецова, И.В. Храмоин Семантико-синтаксический анализ естественных языков. Часть II. Метод семантико-синтаксического анализа текстов // Искусственный интеллект и принятие решений – 2014. – №1 – С. 11-24. (К-1, RSCI)
- Осипов Г.С., Смирнов И.В., Тихомиров И.А. Реляционно-ситуационный метод поиска и анализа текстов и его приложения // Искусственный интеллект и принятие решений. – 2008. – №2. – С. 3-10. (К-1, RSCI)
- Chistova E., Shelmanov A., Pisarevskaya D., Kobozeva M., Isakov V., Panchenko A., Toldova S., Smirnov I. RST Discourse Parser for Russian: an Experimental Study of Deep Learning Models // International Conference on

Analysis of Images, Social Networks and Texts. Lecture Notes in Computer Science. – 2021. – V. 12602. – pp. 105-119. (Scopus)

– Smirnov I., Stankevich M., Kuznetsova Y., Suvorova M., Larionov D., Nikitina E., Savelov M., Grigoriev O. TITANIS: A Tool for Intelligent Text Analysis in Social Media // In: Kovalev S.M., Kuznetsov S.O., Panov A.I. (eds) Artificial Intelligence. RCAI 2021. Lecture Notes in Computer Science. – 2021. – V. 12948. – pp 232-247. (Scopus, WoS)

– Stankevich M., Latyshev A., Kuminskaya E., Smirnov I., Grigoriev O. Depression Detection from Social Media Texts // Data Analytics and Management in Data Intensive Domains: XXI International Conference DAMDID/RCDL'2019: Conference Proceedings. – 2019. – pp. 352-362. (Scopus)

– Stankevich M., Smirnov I., Kuznetsova Y., Kiselnikova N., Enikolopov S. Predicting Depression from Essays in Russian // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2019”. – 2019. – pp. 637–647. (Scopus)

– Stankevich M., Isakov V., Devyatkin D. and Smirnov I. Feature Engineering for Depression Detection in Social Media // Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods. – 2018. – pp. 426-431. (Scopus)

– Osipov, G., Smirnov, I., Tikhomirov, I., Sochenkov, I., & Shelmanov, A. Exactus Expert – Search and Analytical Engine for Research and Development Support. // Novel Applications of Intelligent Systems. – 2016. – pp. 269-285. (Scopus)

– Shelmanov A. O., Smirnov I. V. Methods for semantic role labeling of Russian texts // Computational Linguistics and Intellectual Technologies. Proceedings of International Conference Dialog. – 2014. – V. 13. – №20. – pp. 607-620. (Scopus)

– Osipov G., Smirnov I., Tikhomirov I., Sochenkov I., Shelmanov A., and Shvets A. Information Retrieval for R&D Support // Paltoglou, Georgios, Loizides, Fernando, Hansen, Preben (Eds.) Professional Search in the Modern World. – Lecture Notes in Computer Science (LNCS). – 2014. – Vol. 8830. – pp. 45–69. (Scopus)

– Gennady Osipov, Ivan Smirnov, Ilya Tikhomirov, Olga Vybornova. Technologies for Semantic Analysis of Scientific Publications // Proceedings of 2012 IEEE 6th International Conference Intelligent Systems. – 2012. – Vol. II. – pp. 58-62. (Scopus)

Диссертация соответствует требованиям п.14 Положения о присуждении ученых степеней (Положение правительства РФ №842 от 24.09.2013 в

действующей редакции) и не содержит заимствованного материала без ссылки на авторов.

Диссертация Смирнова Ивана Валентиновича «Методы разноуровневого анализа текстов на естественном языке и их приложения в системах информационного поиска и психолингвистических исследованиях» является законченной научно-квалификационной работой, результаты которой можно квалифицировать как научное достижение в области интеллектуального анализа текстов на естественном языке, имеющее важное значение для создания эффективных систем информационного поиска и текстовой аналитики. Диссертация соответствует критериям Положения о присуждении ученых степеней, предъявляемым к докторским диссертациям, и рекомендуется к защите на соискание ученой степени доктора технических наук по специальности 1.2.1 «Искусственный интеллект и машинное обучение» в диссертационном совете 24.1.224.03 при Федеральном государственном учреждении «Федеральный исследовательский центр «Информатика и управление» Российской академии наук».

Заключение принято единогласно на заседании семинара Отделения №7 ФИЦ ИУ РАН 13 октября 2023г., протокол №2. Присутствовало на заседании 15 человек. Результаты голосования: «за» - 15 человек, «против» - нет, «воздержался» - нет.

Руководитель семинара
г.н.с., рук. Отделения 7
д.т.н.

О.Г. Григорьев
«13» октября 2023

Секретарь семинара
в.н.с. отдела 73
к.ф.-м.н.

И.В. Соченков
«13» октября 2023



И.Г. Григорьев и И.В. Соченков
заверяю
секретарь ФИЦ ИУ РАН
д.т.н. В.Н. Захаров
10 2023 г.