

## ЗАКЛЮЧЕНИЕ

экспертной комиссии разового диссертационного совета 6D.KOA-049 при Таджикском техническом университете имени М.С. Осими по докторской диссертации **Худойбердиева Хуршеда Атохоновича** на тему **«Проектирование и реализация автоматических систем обработки информации на таджикском языке»**, представленной на соискание ученой степени доктора технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Экспертная комиссия разового диссертационного совета 6D.KOA-049 при Таджикском техническом университете имени М.С. Осими в составе его членов: доктора технических наук, доцента, председателя экспертной комиссии Мирзоева Сайъало Хабибулоевича и ее членов доктора технических наук, доцента Каландарбекова Имомёрбека, доктора технических наук, доцента Рахимова Нодира Одиловича, сформированная по решению совета, протокол № 1/ХЯд, от "15" марта 2024 года, рассмотрев диссертационную работу **Худойбердиева Хуршеда Атохоновича** на тему **«Проектирование и реализация автоматических систем обработки информации на таджикском языке»**, представленную на соискание ученой степени доктора технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей», в рамках ее предварительной экспертизы приняла следующее решение:

**Актуальность исследования.** Исследование направлено на разработку автоматических систем обработки информации на таджикском языке для применения в компьютерной лингвистике, включая проверку орфографии, машинный перевод на русский, синтез речи и голосовое управление. Эти системы работают на основе математических моделей и программных приложений. Создание системы проверки орфографии, синтеза речи, и голосового управления, а также системы машинного перевода, являются ключевыми задачами в области компьютерной лингвистики.

Использование информационно-коммуникационных технологий в Таджикистане привлекло внимание исследователей в области математики, информационных технологий и лингвистики. Под руководством академика НАНТ З.Д. Усманова была инициирована новая область исследований – компьютерная лингвистика, решение проблем которой ставит перед учеными ряд важных задач. Одной из основных задач для каждой страны является определение своего места в мировом сообществе в условиях глобализации. Странам необходимо принять решение: оставаться пассивными потребителями культурных и научно-технических достижений других народов или активно представлять свои национальные ценности и мировоззрение мировому сообществу. Это особенно важно для стран, находящихся на стадии развития современных технологий.

Стратегия изучения и развития математических, точных и естественных

наук в сфере образования и науки на период до 2030 года, Государственная стратегия «Информационно-коммуникационные технологии для развития Республики Таджикистан», Государственная программа развития государственного языка на 2020-2030 годы, составляя базовую основу актуальности исследования и данной диссертационной работы, имеют ключевое значение для подготовки специалистов в области информационных технологий и компьютерной лингвистики, обеспечивая стране активное участие в глобальном научно-техническом развитии.

**Цель исследования** заключается в разработке моделей, методов и алгоритмов для создания информационных систем автоматической обработки данных на таджикском языке с последующим использованием в человеко-машинных системах управления, основанных на естественном языковом взаимодействии.

**Научная специальность** диссертационной работы соответствует паспорту научной специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» по следующим пунктам:

П.1. Модели, методы и алгоритмы проектирования и анализа программ и программных систем, их эквивалентных преобразований, верификации и тестирования;

П.3. Модели, методы, алгоритмы, языки и программные инструменты для организации взаимодействия программ и программных систем;

П.4. Системы управления базами данных и знаний;

П.5. Программные системы символьных вычислений;

П.7. Человеко-машинные интерфейсы; модели, методы, алгоритмы и программные средства машинной графики, визуализации, обработки изображений, систем виртуальной реальности, мультимедийного общения.

**Научная новизна состоит в:**

- предложении новых научно-технических концепций, математических моделей, методов и структуры данных, которые образуют теоретическую базу для системного анализа и изучения текстовой информации;

- разработке методов и алгоритмов для практического, структурного и объектно-ориентированного проектирования систем автоматической обработки данных;

- представлении новых подходов к созданию программных средств для автоматического синтеза речи на таджикском языке, включая систему автоматической проверки правописания TajSpell в пакете программ Microsoft Office; программные модули для автоматического перевода текста с таджикского на русский и английский языки в виде интернет-приложения, доступного по адресу [tarjumon.tajlingvo.tj](http://tarjumon.tajlingvo.tj);

- предложении новых алгоритмов машинного перевода, создании компьютерных параллельных корпусов Tajik-Russian-Parallel Corpus и Tajik-English-Parallel Corpus в форме веб-приложений, а также программных модулей для автоматического перевода текста с таджикского на русский и английский языки;

- разработке новых моделей, методов синтеза речи и компьютерных программ Computer Tajik Text Narrator, Tajik Text-to-Speech, которые увеличивают эффективность практического применения информационно-коммуникационных технологий для решения актуальных лингвистических задач и задач речевых технологий на таджикском языке.

Все полученные результаты были реализованы в программном комплексе TajLINGVO, который позволяет:

- существенно сократить время изучения таджикского языка как для пользователей в Таджикистане, так и за его пределами;

- повысить уровень обоснованности принимаемых решений в области компьютерной лингвистики и задач таджикского языка;

- обеспечить создание и использование корректного контента на таджикском языке в интернете.

**Достоверность и обоснованность** результатов, полученных в диссертационной работе, обоснована предложенными математическими моделями элементов текстовой информации с целью последующей обработки. В свою очередь достоверность спроектированных автоматических систем и компьютерных модулей подтверждена корректным выбором исходных данных и репрезентативной выборки текстовой информации в формировании постановки задач разработки математических и компьютерных средств и их реализации в информационных системах для автоматической обработки информации на таджикском языке. В диссертационной работе использованы результаты, полученные ранее другими авторами и отмеченные ссылками.

**Методы исследования.** Методы исследования включали в себя систематический анализ, математическую статистику, обработку данных, теорию алгоритмов, математическую и компьютерную лингвистику, а также компьютерное моделирование информационных систем, технологии программирования и обработки данных.

**Положения, выносимые на защиту:** концепция автоматической обработки текстовой информации на таджикском языке, как объект научного исследования и программные средства для систематического анализа, на основе которых определены понятия и теоретические аспекты проектирования системы автоматической проверки правописания, машинного перевода и синтеза речи на таджикском языке.

**Личный вклад автора.** Постановка задач, их реализация, все результаты и выводы, представленные в диссертации, были получены автором лично или под его руководством, являются оригинальными, новыми и опубликованы в открытой литературе. Программный комплекс TajLINGVO зарегистрирован в Национальном патентном центре Министерства экономического развития и торговли Республики Таджикистан, а созданные автором проекты не имеют аналогов. Полный список проектов доступен на веб-сайте [www.tajlingvo.tj](http://www.tajlingvo.tj).

**Теоретическое значение работы** заключается в представлении методов и алгоритмов обработки элементов текста и акустических сигналов на естественном языке, способствующих изучению таджикского языка. За

последние годы были протестированы, улучшены и внедрены автоматические системы и новые приложения в составе программного комплекса TajLINGVO.

**Практическое значение** результатов исследования подтверждается опытом создания и апробирования программных средств для разработки электронных словарей, электронного тезауруса, автоматического синтеза речи, проверки правописания и автоматического перевода, которые можно использовать в научно-исследовательских институтах для проведения исследований в сфере компьютерной лингвистики, а также для эффективного делопроизводства на таджикском языке на предприятиях не зависимо от их отраслевой принадлежности.

Основные результаты исследований прошли опытную эксплуатацию и подтверждены соответствующими актами о практическом внедрении в следующих организациях:

1. Худжандский научный центр Национальной академии наук Таджикистана;
2. Управление по инвестициям и управлению государственным имуществом Согдийской области;
3. ГОУ Худжандский государственный университет имени академика Б. Гафурова;
4. Таджикский государственный университет права, бизнеса и политики;
5. Политехнический институт Таджикского технического университета имени академика М.С. Осими в городе Худжанде;
6. ЗАО «Душанбе Сити Банк».

**Степень достоверности полученных результатов.** Надежность спроектированных автоматических систем и программного комплекса подтверждена соответствующими актами о практическом внедрении, документами о выдаче государственного регистрационного номера интеллектуальной продукции и информационных ресурсов в Национальном патентно-информационном центре Министерства экономического развития и торговли Республики Таджикистан. Достоверность результатов подтверждается также признанием заслуг автора в данной области науки со стороны различных организаций и учреждений республики. В частности, это премия имени академика С.У. Умарова в области физико-математических, химических, геологических и технических наук Национальной академии наук РТ, 2015 г.; Госпремия для учёных и преподавателей естественных, точных и математических дисциплин, 2021 г.; диплом третьей степени республиканского конкурса «Наука – цвет процветания», номинация инновация и нововведение, 2021 г.; Почетная грамота и медаль «100 НОВЫХ ЛИЦ» стран Содружества Независимых Государств, 2022 г.

Основные результаты диссертации обсуждены на многочисленных конференциях и научных семинарах, таких как: международная конференция «Современные вопросы математики», посвященная 50-летию Института математики имени А. Джуроева НАНТ, (26-27 мая 2023г.), г. Душанбе; всероссийская научно-практическая конференция с участием международных представителей по теме «Обмен информацией в междисциплинарных

исследованиях II», (14 апреля 2023г.), Академия права и управления ФСИН России, г. Рязань, Российская Федерация; международная научно-практическая конференция «Новые достижения в области естественных наук и информационных технологий», Российско-таджикский славянский университет, (30 мая 2023г.), г. Душанбе; научно-практическая конференция «Практические информационные системы: проблемы моделирования, использование в развивающихся странах», ХПИ ТТУ им. академика М.С. Осими (2012г., 2017г., 2023г.), а также на ежегодных научно-практических семинарах «Новые информационные технологии в автоматических системах», Институт прикладной математики им. М.В. Келдыша РАН (с 2013 по 2019).

**Соответствие диссертации критериям, установленным Положением о порядке присуждения учёных степеней ВАК при Президенте РТ**

**Публикации по теме диссертации.** По материалам диссертационного исследования опубликовано 68 работ, в том числе 25 (11 без соавторства) из которых опубликованы в журналах, рекомендованных ВАК при Президенте РТ и ВАК РФ, 27 статей в международных сборниках статей и журналах. В патентно-информационном центре при Министерстве экономического развития и торговли Республики Таджикистан получено 18 свидетельств о государственной регистрации информационных ресурсов и интеллектуальной продукции.

**Основное содержание диссертации.** Диссертационное исследование состоит из 328 компьютерных страниц, введения, 6 глав, 19 таблиц, 15 рисунков, библиографии с 322 названиями и 2 приложений.

**Оригинальность содержания диссертации** составляет 88,44%. Цитирования оформлены корректно. Заимствованного материала, использованного в диссертации без ссылки на автора либо на источник заимствования не обнаружено. Научных работ, выполненных соискателем ученой степени в соавторстве без ссылок на соавторов не выявлено.

**Комиссия рекомендует:**

1. Тематика и содержание диссертации Худойбердиева Х.А. соответствуют специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей», что позволяет диссертационному совету принять данную диссертацию к защите.

2. Автореферат и опубликованные работы полностью отражают основное содержание диссертации. Рекомендуется разрешить публикацию автореферата.

3. Количество публикаций в рецензируемых изданиях соответствует установленным требованиям и стандартам Типового положения о диссертационном совете, порядка присвоения учёных степеней и ученых званий.

4. В качестве официальных оппонентов по диссертации комиссия разового диссертационного совета 6D.KOA-049 рекомендует назначить следующих специалистов:

- академика НАНТ, доктора физико-математических наук, профессора Илолова Мамадшо Илоловича, заведующего отделом математического моделирования динамических процессов Центра инновационного развития науки и новых технологий НАНТ;

- доктора технических наук, доцента Пруцкого Александра Викторовича, профессора кафедры «Вычислительная и прикладная математика» Федерального государственного бюджетного образовательного учреждения высшего образования «Рязанский государственный радиотехнический университет», Российская Федерация;

- доктора технических наук, профессора Бекназарову Саиду Сафибуллаевну, профессора кафедры «Телевизионные и медиа технологии» Ташкентского университета информационных технологий имени Мухаммада ал-Хоразмий Республики Узбекистан.

5. В качестве ведущей организации назначить Национальный университет Таджикистана.

6. Разрешить размещение информации о предстоящей защите, текста диссертации и автореферата на официальных сайтах ВАК при Президенте Республики Таджикистан и Таджикского технического университета имени М.С. Осими.

Председатель экспертной комиссии,

д.т.н., доцент



Подпись д.т.н., доцента Мирзоева С.Х. заверяю

(подпись, печать)

С.Х. Мирзоев

Член экспертной комиссии,

д.т.н., доцент



Подпись д.т.н., доцента Каландарбекова И. заверяю

(подпись, печать)

И. Каландарбеков

Член экспертной комиссии,  
д.т.н., доцент

Подпись д.т.н., доцента Рахимова Н.О. заверяю

(подпись, печать)



Н.О. Рахимов