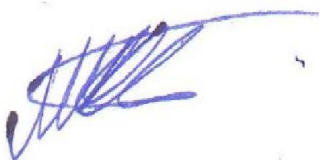
НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК ТАДЖИКИСТАНА

На правах рукописи

УДК 51:681.3+004.8
ББК 22.1+73
Р-27



Рахимов Манучехр Фарходович

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ДОКУМЕНТООБОРОТА НА ОСНОВЕ МИКРОСЕРВИСНОЙ АРХИТЕКТУРЫ И НЕЙРОСЕТЕВЫХ МЕХАНИЗМОВ BERT

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени доктора философии (PhD) по специальности 6D060200 – Информатика (6D060204 – Теоретические основы информатики)

Душанбе-2026

Диссертация выполнена в отделе математического моделирования Института математики имени академика А. Джураева Национальной академии наук Таджикистана

Научный руководитель:	**Комилиён Файзали Саъдулло** – доктор физико-математических наук, профессор кафедры информатики Таджикского национального университета
Официальные оппоненты:	**Худойбердиев Хуршед Атохонович** – доктор технических наук, доцент кафедры программирования и информационных систем Политехнического института Таджикского технического университета имени академика М.С. Осими в городе Худжанде,
	Гуломсафдаров Абдулназар Гуломназарович – кандидат технических наук, заведующий кафедрой компьютерных наук и цифровых технологий Душанбинского инновационного института
Ведущая организация:	Филиал Московского государственного университета имени М.В. Ломоносова в г. Душанбе

Защита диссертации состоится «17» декабря 2026 г. в 14^{00} часов на заседании диссертационного совета 6D.KOA-105 при Таджикском национальном университете по адресу: 734027, г. Душанбе, улица Буни-Хисорак, корпус 17, аудитория 203.

E-mail: rohatullo1956@mail.ru; номер мобильного телефона ученого секретаря: +992981097033.

С диссертацией и авторефератом можно ознакомиться на сайте http//:www.tnu.tj и научной библиотеке Таджикского национального университета.

Автореферат разослан «___» «_____» 2026 г.

Ученый секретарь диссертационного совета 6D.KOA-105, кандидат физико-математических наук, доцент — Садуллоев Р.И.

ВВЕДЕНИЕ

Актуальность темы исследования. Рост объёмов данных, усложнение внутренних коммуникаций и постоянное давление на скорость управленческих решений привели к тому, что традиционные модели документооборота всё чаще упираются в ограничения по производительности и по устойчивости обработки информационных потоков. В организациях с сильным документооборотом крупные объёмы информации, различающейся по структуре и по назначению, накапливаются постоянно. Указанные проблемы в этих условиях заметны сильнее. Большую часть процедур всё ещё либо выполняют вручную, либо регулярно привлекают к ним ручное участие. Вероятность ошибок из-за этого растёт, контроль усложняется. Дополнительное время уходит на принятие решений.

Под влиянием распределённых вычислительных технологий изменились и требования к современным системам. Приоритетными считают масштабируемость и устойчивость к сбоям. От систем также ждут способности держать производительность при меняющейся нагрузке. Архитектуры, разделяющие систему на независимые сервисы, выбирают всё чаще. Обработка данных при микросервисной архитектуре становится гибкой. Развитие программной системы упрощается за счёт декомпозиции на сервисы. Стабильность работы сохраняется даже при значительных нагрузках.

Цифровизацию государственного управления в Республике Таджикистан относят к приоритетным направлениям развития. Стратегические и нормативные акты фиксируют этот приоритет; среди них Национальная стратегия развития Республики Таджикистан на период до 2030 года, Концепция формирования электронного правительства в Республике Таджикистан, Программа реализации концепции государственной информационной политики Республики Таджикистан. В названных документах ставят задачи развития цифровой инфраструктуры и более широкого применения интеллектуальных методов обработки информации.

Указанные приоритеты увеличивают спрос на интеллектуальные системы документооборота. Им нужно хранить документы и передавать их между подразделениями. Параллельно требуется автоматический анализ содержания с извлечением важных данных, определением категории документа и поддержкой принятия решений. Чтобы такие системы появились, методы математического моделирования объединяют с технологиями вычислений и с современными алгоритмами обработки текста.

В последние годы быстро продвинулись методы обработки естественного языка. Трансформерные модели BERT (Bidirectional Encoder Representations from Transformers - Двунаправленные кодировщики-представления из трансформеров) расширили средства интеллектуального анализа текстов. С их помощью решают семантическую классификацию, извлечение сущностей, контекстную интерпретацию документов и автоматизированный анализ содержания точнее, чем при традиционных подходах. Однако внедрение этих моделей в системы документооборота сталкивается с практическими ограничениями. Особенно остро ограничения проявляются для низкоресурсных языков, где нет крупных размеченных корпусов и не хватает адаптированных языковых моделей.

Отдельная трудность состоит в том, чтобы встроить интеллектуальные алгоритмы в распределённую вычислительную среду. Во многих работах архитектурное проектирование и интеллектуальный анализ документов изучают разобщенно. Одни авторы сосредоточены на микросервисной организации программных систем, другие на отдельных алгоритмах NLP или на нейросетевых моделях обработки текста. При этом проблемы комплексного объединения микросервисной архитектуры и математических моделей интеллектуального анализа документов изучены ещё недостаточно.

Важность темы темы связана также с адаптацией интеллектуальных систем к условиям реального государственного управления. В качестве среды в настоящем исследовании рассматривается Комитет по начальному и среднему профессиональному образованию при

Правительстве Республики Таджикистан. Такой выбор позволяет учесть действующие регламенты обработки документов, как устроено учреждение и информационные потоки, которые возникают в ходе его работы. Реальные условия делает разработанные решения более подходящими для дальнейшего внедрения.

Актуальность диссертации связана с некоторыми обстоятельствами. Требуется повысить эффективность документооборота, также возрастает потребность в масштабируемых и отказоустойчивых архитектурах, рассчитанных на интенсивную обработку данных. Инструментарий интеллектуального анализа текстовой информации тоже расширяется. Существующие NLP-решения слабо приспособлены к низкоресурсным языковым средам, и это ограничивает их практическое применение. Сохраняется задача построения математически обоснованных моделей интеллектуальной обработки документов.

Степень исследования научной темы. Научная тема «Математическое моделирование интеллектуальной системы документооборота на основе микросервисной архитектуры и нейросетевых механизмов BERT» исследована на уровне фундаментальных и прикладных аспектов, включая разработку математических моделей, проектирование архитектуры системы, методы адаптации трансформерных моделей к низкоресурсным языкам, а также их практическую реализацию и апробацию в условиях реальной информационной инфраструктуры.

При разработке комплекса компьютерных моделей использованы результаты исследований в области обработки естественного языка. Существенный вклад внесли Ashish Vaswani [23], Jacob Devlin [15] и Thomas Wolf [24], заложившие основы трансформерной архитектуры и модели BERT, обеспечивающей высокую точность анализа текстовых данных.

Архитектурные аспекты опираются на работы Sam Newman [19], Martin Fowler [17] и Thomas Erl [16], в которых микросервисная архитектура рассматривается как эффективный подход к построению масштабируемых и отказоустойчивых систем. Развитие данных идей отражено в исследованиях Chris Richardson [22], формализовавшего ключевые паттерны распределённых приложений.

Особое значение в контексте интеллектуальных систем приобретают технологии контейнеризации и оркестрации. Использование платформ Docker (платформа контейнеризации приложений и сервисов) и Kubernetes (платформа автоматического управления контейнеризированными приложениями), как показано в работах Chris Richardson [22], а также Brendan Burns и Joe Beda [13] обеспечивает изоляцию сервисов, автоматизацию развёртывания и эффективное управление масштабированием, что позволяет интегрировать ресурсоёмкие компоненты машинного обучения в распределённую инфраструктуру. Это позволяет интегрировать вычислительно сложные компоненты, включая модели машинного обучения, в распределённую инфраструктуру с возможностью динамического перераспределения ресурсов.

В отечественных исследованиях рассматривались вопросы проектирования распределённых систем и оптимизации обработки больших объёмов текстовой информации в условиях ограниченных вычислительных ресурсов [2-3; 5-6; 12; 1-А–12-А].

Проведённый анализ показал, что интеграция микросервисной архитектуры с моделями BERT (модели семейства BERT – совокупность языковых моделей, основанных на архитектуре BERT и адаптированных для различных задач анализа текста) формирует основу интеллектуальных систем документооборота, обеспечивающих автоматическое извлечение и структурирование данных. Вместе с тем недостаточно исследованы вопросы адаптации моделей к низкоресурсным языкам и их эффективной интеграции в распределённые среды.

Методологическую основу исследования составили работы отечественных и зарубежных учёных в области математического моделирования, распределённых систем и машинного обучения: Ф.С. Комилиён, З.Дж. Усманов, М.К. Юнуси, М.Р. Ёров, Х.А. Худойбердиев, В.В. Меншуткин, M. Fowler, S. Newman, L. Bass, I. Weber, L. Zhu, B. Burns,

J. Beda, K. Hightower, N. Dragoni, C. Pahl, Q. Zhang, M. Chen, J. Devlin, M.W. Chang, K. Lee, K. Toutanova, A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, A. Rogers, T. Wolf, C. Richardson, T. Brown, P. Jamshidi, D. Taibi, T. Erl, J. Thönes, N. Alshuqayran, H. Zhang, S. Giallorenzo, A.L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, L. Safina, J. Lewis и другие.

В рамках диссертации получили развитие следующие направления: формализация процессов обработки документов, разработка микросервисной архитектуры с выделенным интеллектуальным компонентом, адаптация моделей BERT для низкоресурсного языка (язык, для которого доступно ограниченное количество цифровых текстовых ресурсов и размеченных данных), а также создание практического алгоритма реализации системы, ориентированного на условия функционирования государственных структур. Особое внимание в исследовании уделено задачам масштабируемости, отказоустойчивости и эффективного использования вычислительных ресурсов.

Несмотря на достигнутые результаты, остаются открытыми задачи повышения производительности при больших объёмах данных, расширения мультиязычной поддержки и интеграции с государственными системами электронного документооборота.

Таким образом, проведённое исследование занимает промежуточное положение между фундаментальными разработками трансформерных моделей и их прикладным внедрением в интеллектуальные системы документооборота, закрывая ряд ранее недостаточно изученных аспектов, связанных с адаптацией моделей BERT к низкоресурсным языкам и комбинированным масштабированием микросервисных архитектур.

Связь исследований с программами и научными темами. Диссертационное исследование выполнено в соответствии с приоритетными направлениями научно-технического и цифрового развития Республики Таджикистан и направлено на развитие методов математического моделирования, интеллектуальных информационных систем и технологий обработки данных.

Тематика диссертации непосредственно связана с реализацией государственных стратегий и программ в области цифровизации, включая Национальной стратегии развития Республики Таджикистан до 2030 года [10], Концепции формирования электронного правительства в Республике Таджикистан [9] и Программы цифровой экономики Республики Таджикистан [11], в рамках которых предусмотрено внедрение современных информационных технологий, систем электронного документооборота и методов искусственного интеллекта в процессы государственного управления.

Диссертация выполнена в рамках научного направления кафедры информатики Таджикского национального университета «Разработка и исследование математико-компьютерных моделей экосистем водоёмов, заповедников и системы образования», зарегистрированного в Национальном патентно-информационным центром Республики Таджикистан (№ 0116TJ00657), и расширяет его за счёт применения методов математического моделирования и искусственного интеллекта к задачам автоматизации документооборота.

Результаты исследования ориентированы на разработку и внедрение интеллектуальных систем документооборота, основанных на микросервисной архитектуре и трансформерных моделях типа BERT, что соответствует современным тенденциям цифровой трансформации и повышению эффективности управления информационными ресурсами.

ОБЩАЯ ХАРАКТЕРИСТИКА ИССЛЕДОВАНИЯ

Диссертационное исследование посвящено разработке и математическому обоснованию интеллектуальной системы документооборота, реализованной на основе микросервисной архитектуры и трансформерных моделей типа BERT, обеспечивающей повышение эффективности обработки документов, а также обладающей свойствами масштабируемости и адаптивности к изменяющимся условиям эксплуатации.

Цель исследования заключается в разработке и теоретико-прикладном обосновании математической модели интеллектуальной системы документооборота, построенной с использованием микросервисной архитектуры и нейросетевых механизмов семейства BERT, обеспечивающей повышение точности обработки документов, устойчивость функционирования системы и возможность её масштабирования (способность системы сохранять эффективность при увеличении нагрузки) в условиях распределённой вычислительной среды.

Поставленная цель обусловила необходимость решения комплекса взаимосвязанных исследовательских и прикладных задач, охватывающих как архитектурные, так и алгоритмические аспекты построения интеллектуальной системы документооборота.

Задачи исследования:

1. Провести анализ современных подходов к построению распределённых информационных систем, включая исследование особенностей микросервисной архитектуры, её преимуществ, ограничений и областей практического применения.
2. Исследовать существующие методы масштабирования микросервисных приложений и механизмы обеспечения отказоустойчивости в распределённых программных средах с высокой интенсивностью обработки запросов.
3. Разработать архитектурную модель интеллектуальной системы документооборота с учётом требований к модульности, расширяемости, устойчивости и взаимодействию отдельных сервисов.
4. Выполнить адаптацию трансформерных моделей типа BERT для решения задач извлечения сущностей, семантического анализа и классификации документов в условиях ограниченности языковых ресурсов.
5. Разработать методы формирования и генерации обучающих данных, ориентированные на повышение эффективности обучения нейросетевых моделей обработки текстовой информации.
6. Реализовать программный прототип интеллектуальной системы документооборота с последующей экспериментальной оценкой эффективности предложенных моделей, алгоритмов и архитектурных решений на основе совокупности количественных показателей.

Объект исследования – интеллектуальные системы документооборота, функционирующие в условиях цифровой трансформации и построенные на основе распределённых микросервисных архитектур.

Предмет исследования – методы и модели математического обеспечения, алгоритмы и архитектурные решения построения и масштабирования интеллектуальной системы документооборота с использованием трансформерных моделей типа BERT в микросервисной среде.

Теоретическую основу исследования составляют концепции математического и компьютерного моделирования информационных систем, микросервисная архитектура и интеллектуальный подход к процессам автоматизации документооборота. Связь темы с существующими теориями проявляется в критическом анализе трансформерных моделей типа BERT и решении проблем адаптации многоязычных моделей в условиях ограниченности ресурсов малоресурсных языков. Разработка комплекса математических моделей, гибридных методов обработки текста и архитектурных решений обеспечила теоретическую базу для проведения компьютерных экспериментов и научного управления документооборотом.

Научная новизна диссертационной работы связана с разработкой теоретических и прикладных подходов к построению интеллектуальных систем документооборота, основанных на совместном использовании микросервисной архитектуры, методов математического моделирования и трансформерных моделей обработки естественного языка. В процессе исследования получены результаты, обладающие как элементами научной новизны, так и практической значимостью.

1. Разработана математическая модель интеллектуальной системы документооборота, в которой объединены принципы микросервисной организации вычислительной среды и трансформерные механизмы обработки текстовых данных. Предложенная модель формализует процессы извлечения, структурирования и интерпретации информации в условиях распределённой обработки и высокой интенсивности документных потоков [6-А; 7-А].

2. Предложено архитектурное решение распределённой системы документооборота с выделением специализированного компонента интеллектуального анализа – AI Service (сервис искусственного интеллекта). В отличие от традиционных монолитных систем, предложенный подход обеспечивает независимое масштабирование ресурсоёмких вычислительных модулей, что повышает производительность системы и упрощает её дальнейшее развитие [3-А].

3. Разработан метод адаптации многоязычных моделей BERT к условиям ограниченности размеченных языковых данных. Метод основан на использовании синтетически сформированных обучающих выборок и последующем дообучении модели; это позволяет повысить качество извлечения именованных сущностей в низкоресурсных языках [5-А].

4. Предложен гибридный метод обработки документов, сочетающий трансформерные модели машинного обучения с алгоритмами, основанными на правилах. Такое сочетание рассматривается как способ повысить устойчивость системы при неполных и неоднородных данных, а также сделать результаты автоматизированного анализа более интерпретируемыми [3-А].

5. Систематизированы методы масштабирования микросервисных систем, включая горизонтальное (увеличение производительности системы путем добавления новых вычислительных узлов) и вертикальное (увеличение вычислительных ресурсов отдельного узла путем наращивания мощности процессора, памяти или дискового пространства) распределение вычислительной нагрузки. В рамках исследования показана целесообразность их комбинированного применения для повышения производительности и отказоустойчивости распределённых приложений, включающих компоненты интеллектуальной обработки данных [8-А; 10-А].

6. Теоретически обоснована роль контейнеризации и механизмов оркестрации в построении распределённых интеллектуальных систем [3-А]. Показано, что данные технологии создают основу для автоматического масштабирования сервисов, балансировки нагрузки, повышения отказоустойчивости и более рационального распределения вычислительных ресурсов [11-А].

7. На основе разработанной математической модели интеллектуальной системы документооборота, ориентированной на применение микросервисной архитектуры и моделей BERT, предложен практический алгоритм реализации системы с учётом особенностей организационной структуры Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан. Разработанный подход адаптирован к условиям существующей информационной инфраструктуры и ориентирован на возможность практического внедрения [1-А–5-А].

8. Проведена экспериментальная апробация предложенных моделей и алгоритмов, результаты которой подтвердили повышение точности обработки документов и сокращение времени выполнения операций по сравнению с традиционными подходами к документообороту. Полученные результаты свидетельствуют о прикладной эффективности разработанных решений и возможности их использования в распределённых информационных системах различного назначения [1-А].

Совокупность представленных результатов формирует научно обоснованную базу для дальнейшего развития интеллектуальных систем документооборота нового поколения, ориентированных на функционирование в условиях цифровой экономики и распределённой вычислительной инфраструктуры.

Положения, выносимые на защиту:

1. Разработана и обоснована математическая модель и архитектурное решение интеллектуальной системы документооборота, которые построены на базе микросервисной архитектуры с использованием трансформерных моделей типа BERT и обеспечивают формализацию процессов извлечения, структурирования и обработки данных в условиях реальной информационной инфраструктуры. При этом за счет выделения специального компонента интеллектуального анализа и применения совместной горизонтальной и вертикальной масштабируемости микросервисов гарантируются высокая производительность и устойчивость системы к изменяющимся нагрузкам.
2. Доказано, что метод адаптации мультиязычных трансформерных моделей BERT на основе синтетической генерации обучающих данных и переобучения является эффективным для условий малоресурсных языков. На этой основе разработан гибридный метод обработки документов, объединяющий алгоритмы машинного обучения и подходы на основе правил, позволяющий существенно повысить качество и точность обработки текстов в условиях ограниченности языковых ресурсов.
3. Установлено, что внедрение практического алгоритма реализации интеллектуальной системы документооборота в структуру Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан обеспечивает сокращение времени обработки документов и повышение точности извлечения информации, что подтверждает высокую практическую значимость результатов исследования и их экспериментальную устойчивость.

Теоретическое и практическое значения исследования. Теоретическая значимость диссертационной работы заключается в развитии научных основ математического моделирования интеллектуальных систем документооборота в условиях распределённых вычислительных сред. Полученные результаты расширяют существующие представления о принципах построения микросервисных архитектур за счёт их интеграции с современными методами обработки естественного языка на основе трансформерных моделей.

Разработанная математическая модель интеллектуальной системы документооборота формализует процессы извлечения и структурирования данных, что вносит вклад в развитие методов анализа текстовой информации и теории распределённых информационных систем. Существенным является теоретическое обоснование гибридного подхода, объединяющего методы машинного обучения и алгоритмы на основе правил, что позволяет повысить устойчивость и интерпретируемость интеллектуальных систем.

Важным теоретическим результатом является систематизация и обоснование методов горизонтального и вертикального масштабирования микросервисных систем, а также доказательство эффективности их комбинированного применения в задачах обработки данных. Кроме того, разработан метод адаптации трансформерных моделей к условиям низкоресурсных языков, что расширяет возможности применения современных NLP-технологий (совокупность методов и инструментов обработки текстовых данных) в многоязычных средах.

В совокупности полученные результаты формируют теоретическую базу для дальнейших исследований в области интеллектуальных распределённых систем, цифрового документооборота и прикладного использования методов искусственного интеллекта.

Практическая значимость исследования определяется возможностью непосредственного применения разработанных моделей, методов и архитектурных решений при создании и внедрении интеллектуальных систем документооборота в различных отраслях.

На основе предложенной математической модели и архитектурного подхода разработана интеллектуальная система, обеспечивающая автоматизацию процессов обработки документов, включая извлечение данных, классификацию и формирование структурированной информации. Внедрение системы позволяет существенно сократить

время обработки документов, снизить количество ошибок ручного ввода и повысить эффективность управленческих процессов.

Особую практическую ценность представляет разработанный инженерно-прикладной алгоритм реализации системы, адаптированный под структуру Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан, что подтверждает применимость результатов исследования в реальных условиях государственной информационной инфраструктуры.

Предложенные решения также могут быть использованы:

– в государственном секторе при автоматизации документооборота и оказании электронных услуг;

– в коммерческих организациях для обработки договорной и финансовой документации;

– в финансовых учреждениях при реализации процедур контроля и анализа данных.

Дополнительно практическая значимость проявляется в возможности масштабирования системы, её интеграции с существующими информационными платформами и адаптации к различным предметным областям.

Степень достоверности результатов диссертации обоснована реализацией комплексного исследования научно обоснованных запланированных задач, использованием современных технологий математического и компьютерного моделирования, а также трансформерных моделей BERT. Достоверность результатов также подтверждается проведением компьютерных экспериментов на основе реальных данных информационной инфраструктуры, синтетической генерацией обучающих данных и экспериментальным внедрением разработанных алгоритмов в структуру Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан.

Соответствие диссертации паспорту научной специальности. Диссертация и ее содержание охватывают области научных исследований паспорта специальности доктора философии (PhD) 6D060204 – Теоретические основы информатики и соответствует следующим направлениям (пунктам) специальности:

Пункт 4. Комплексные исследования научно-технических и фундаментально-прикладных проблем с применением современной технологии математического моделирования и вычислительного эксперимента.

Пункт 11. Машинное обучение и искусственный интеллект: теоретические основы обучающихся систем, методы обнаружения новых знаний (Data Mining), логический вывод.

Пункт 15. Модели информационных процессов и структур.

Пункт 18. Новые средства приобретения знаний и создания онтологии, средства интеллектуализации бизнес-процессов.

Пункт 25. Распределённые многопользовательские системы.

Пункт 26. Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента.

Личный вклад соискателя ученой степени. Соискатель участвовал в формировании темы диссертации, самостоятельно провел анализ литературы, разработал математическую модель интеллектуальной системы документооборота на основе микросервисной архитектуры и BERT, реализовал практический алгоритм, провел эксперименты, обработку и интерпретацию результатов, а также подготовил методические рекомендации и текст диссертации.

Апробация и применение. Результаты диссертационного исследования обсуждались на научных семинарах Института математики имени академика А. Джураева Национальной академии наук Таджикистана (НАНТ), на семинаре кафедры информатики Таджикского национального университета (ТНУ), а также были представлены на республиканских и международных конференциях 2022-2026 гг. в НАНТ, ТНУ, Академии государственного управления при Президенте Республики Таджикистан, Филиале Московского государственного университета им. М.В. Ломоносова в г. Душанбе, Бохтарском

государственном университете имени Носира Хусрава, Таджикском международном университете туризма и предпринимательства, Технологическом университете Таджикистана и других профильных учреждениях.

Публикации по теме диссертации. Основные результаты диссертационного исследования опубликованы в 12 работах [1-А–12-А], включая 7 статей в рецензируемых журналах Республики Таджикистан, 4 тезиса докладов, представленных на международных и республиканских конференциях, и 1 статью в зарубежном журнале.

Структура и объем диссертации. Диссертация включает «Введение», «Общую характеристику исследования», четыре главы, раздел «Выводы» с подразделами «Основные научные результаты диссертации» и «Рекомендации по практическому использованию результатов», раздел «Литература» с подразделами «Список использованных источников» и «Список публикаций соискателя», а также «Приложения». Работа выполнена на 197 страницах в Microsoft Word 2016, содержит 9 таблиц и 15 рисунков, нумерация которых осуществляется в соответствии с главами диссертации. Список литературы с перечнем публикаций соискателя включает 182 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснованы актуальность, объект и предмет, цель и задачи, степень разработанности, связь с госпрограммами, новизна, положения на защиту, теоретическая и практическая значимость. Указана связь работы с государственными программами. Сформулированы научная новизна, положения, выносимые на защиту, теоретическая и практическая значимость. Приведены сведения о соответствии работы паспорту специальности, апробации результатов в публикациях, а также описаны структура и объём диссертации.

В первой главе «проведён анализ научной литературы и современных технологических решений, позволивший сформировать системное представление о развитии микросервисной архитектуры и интеллектуальных систем документооборота в условиях цифровой трансформации» [16; 18; 20; 22].

Установлено, что микросервисная архитектура является закономерным этапом эволюции программных систем, обеспечивающим переход к распределённым и адаптивным решениям (рисунок 1.1).

Доказано, что её преимущества проявляются в модульности, автономности и независимости развертывания сервисов. При этом выявлены ограничения, обусловленные усложнением управления, необходимостью внедрения DevOps-практик и обеспечением согласованности данных.

Обосновано, что эффективное функционирование микросервисных систем достигается при соблюдении принципов слабой связанности, изоляции данных и автоматизации жизненного цикла. Показано, что применение контейнеризации, оркестрации, асинхронных механизмов взаимодействия и облачных технологий обеспечивает требуемый уровень масштабируемости и отказоустойчивости.



Рисунок 1.1. Эволюция архитектур программных систем

Установлено, что повышение эффективности достигается при комплексном использовании горизонтального масштабирования, событийно-ориентированной архитектуры, кэширования и балансировки нагрузки. Доказана необходимость применения дополнительных методов оптимизации для вычислительно сложных компонентов, включая батчирование и использование аппаратных ускорителей.

Разработана интеллектуальная система документооборота на основе микросервисной архитектуры с выделенным компонентом интеллектуального анализа, обеспечивающая независимую масштабируемость и гибкость развития.

Экспериментально подтверждено, что внедрение нейросетевых алгоритмов семейства BERT позволяет достигать прецизионных результатов при верификации именованных сущностей (вплоть до 93% по интегральному критерию $\mathcal{F}_1$). В работе верифицирована реализуемость масштабирования полиязычных платформ на специфику таджикского языка, относящегося к категории малоресурсных, посредством генерации искусственных обучающих выборок и последующей селективной тонкой настройки (fine-tuning).

Научное оригинальное значение представленного раздела базируется на конвергенции сервисно-ориентированной архитектуры и инструментария интеллектуального анализа данных, проектировании комбинированного алгоритма структурирования информационных потоков, а также на валидации трансформерных конфигураций для лингвистических систем с дефицитом цифровых корпусов.

Утилитарная ценность предложенного решения аргументирована минимизацией временных затрат на верификацию документации и деградацией частоты возникновения операционных сбоев, что детерминирует высокую продуктивность внедрения разработанного комплекса в реальном секторе.

В ходе изысканий были эксплицированы критические барьеры, обусловленные аппаратной энергоемкостью нейросетей и спецификой десегментированных текстовых массивов. Сформулированы векторы последующих разработок, сфокусированные на компрессии параметров нейромоделей, масштабировании мультимодальных концептов и модернизации конвейеров управления жизненным циклом прикладных ML-систем.

Таким образом, доказано, что интеграция микросервисной архитектуры и методов искусственного интеллекта является обоснованным и перспективным направлением развития систем документооборота, соответствующим современным требованиям цифровой экономики.

Вторая глава диссертации посвящена исследованию архитектурных решений, направленных на обеспечение масштабируемости микросервисных систем. «Проведённый анализ позволил сформировать системное представление о современных подходах к

повышению производительности, отказоустойчивости и адаптивности распределённых информационных систем, ориентированных на обработку больших объёмов данных и применение методов искусственного интеллекта» [5; 7; 8; 11-А].

Установлено, что обеспечение масштабируемости в микросервисной архитектуре базируется на комплексном использовании горизонтального и вертикального масштабирования как взаимодополняющих стратегий. Доказано, что горизонтальное масштабирование обеспечивает эластичность и отказоустойчивость за счёт репликации сервисов и распределения нагрузки, тогда как вертикальное масштабирование эффективно при решении ресурсоёмких вычислительных задач, характерных для компонентов машинного обучения. Показано, что наибольшая эффективность достигается при их комбинированном применении с учётом неоднородности нагрузки и функциональной специфики сервисов (таблица 2.1).

Таблица 2.1. Сравнительный анализ подходов

Критерий	**Горизонтальное масштабирование**	**Вертикальное масштабирование**
Принцип	Увеличение числа узлов	Увеличение ресурсов узла
Отказоустойчивость	Высокая	Низкая
Масштабируемость	Практически неограниченная	Ограниченная
Сложность реализации	Высокая	Низкая
Стоимость	Оптимизируемая	Высокая при росте

Обосновано, что технологической основой масштабируемых микросервисных систем является контейнеризация, обеспечивающая изоляцию компонентов, устранение конфликтов зависимостей и переносимость приложений. Установлено, что управление контейнерной инфраструктурой реализуется средствами оркестрации, обеспечивающими автоматизацию развертывания, балансировку нагрузки и контроль состояния сервисов.

Доказано, что применение механизмов автоматического масштабирования (HPA, VPA, Cluster Autoscaling) позволяет динамически адаптировать систему к изменяющейся нагрузке и оптимизировать использование ресурсов. Показано, что эффективность данных механизмов определяется корректной настройкой метрик и политик управления.

Выявлено, что в интеллектуальных микросервисных системах задачи масштабирования имеют специфический характер: сервисы обработки запросов преимущественно масштабируются горизонтально, тогда как вычислительно сложные компоненты требуют вертикального масштабирования с использованием специализированных ресурсов. Обоснована целесообразность применения гибридных архитектурных решений, включающих асинхронные взаимодействия, кэширование и пакетную обработку данных.

Также обосновано фундаментальное значение механизмов оркестрации в качестве базового контура координации распределенных комплексов, гарантирующего автономную регенерацию узлов, бесшовную актуализацию программных компонентов и практическую реализацию декларативных моделей администрирования. Подобный подход существенно повышает детерминированность и общую контролируемость ИТ-ландшафта.

Вместе с тем эксплицирован комплекс деструктивных факторов, выраженных в росте топологической сложности окружения, императивности глубокой интеграции DevOps-методологий, ужесточении критериев сквозного аудита и кибербезопасности, а также в технологических барьерах репликации компонентов, сохраняющих состояние (stateful).

Прикладной потенциал сформулированных в данном разделе выводов сопряжен со спецификой их проецирования на процессы архитектурного синтеза высоконагруженных интеллектуальных платформ (в том числе цифрового документооборота). Это позволяет гарантировать высокие показатели пропускной способности, живучести и оптимальной утилизации аппаратных мощностей.

Аргументировано, что конвергенция подходов к разнонаправленному (горизонтальному и вертикальному) масштабированию, изолированной контейнеризации и автоматизированной диспетчеризации закладывает базис для проектирования современных MSA-архитектур. Оперирование данными паттернами выступает облигатным фактором создания гибких, отказоустойчивых и адаптивных ИТ-продуктов, релевантных вызовам современной цифровой экосистемы [2-А; 9-А].

В третьей главе диссертационного исследования представлен компаративный анализ монолитных и сервисно-ориентированных архитектурных паттернов сквозь призму модернизации цифровых платформ и рационализации использования инфраструктурных мощностей. Реализованный комплексный анализ послужил базисом для систематизации ключевых детерминант, определяющих достоинства и недостатки каждого из изучаемых концептов.

«Вектор форсированной эволюции индустрии программной инженерии диктует необходимость ужесточения критериев адаптивности, горизонтального расширения, живучести и латентности развёртывания апдейтов ПО. Классические монолитные структуры, инкапсулирующие совокупный функционал внутри неделимого программного контура, при долгосрочной эксплуатации обнаруживают стагнацию, вызванную топологической сложностью масштабирования, высокими рисками при рефакторинге и барьерами для трансфера инновационных стеков» [1; 4; 5].

В качестве релевантной альтернативы классическому монолиту позиционируется микросервисная организация ИТ-ландшафта, базирующаяся на атомарном расщеплении системы на автономные модули, жестко привязанные к конкретным изолированным бизнес-компетенциям. Подобная парадигма катализирует гибкость инженерных процессов, нивелирует риски при дистрибуции обновлений, локализует точки отказа и санкционирует внедрение принципов технологического полиморфизма в границах единой экосистемы.

Базисным технологическим драйвером миграции на рельсы сервис-ориентированных платформ выступает инструментарий изолированной контейнеризации и автоматизированной диспетчеризации (в частности, Docker и Kubernetes). Данные программные решения гарантируют унификацию процессов дистрибуции, динамического расширения и координации распределенных программных узлов. Параллельно актуализируется значимость методологического базиса DevOps, включая концепции непрерывной интеграции (CI) и непрерывного развертывания (CD), автоматизирующих инженерный цикл и форсирующих показатель Time-to-Market для наукоемких ИТ-продуктов.

Тем не менее результативность интеграции микросервисной парадигмы детерминирована не только сугубо техническими аспектами, но и императивом глубокой реконфигурации организационно-управленческих и корпоративных паттернов. Необходим дрейф в сторону аджайл-моделей администрирования, интенсификация кросс-функционального взаимодействия между инженерными группами и департаментами эксплуатации, а также укоренение синергетической среды и механизмов кросс-трансфера компетенций.

Системная конвергенция аппаратных, структурных и стратегических трансформаций делегирует предприятиям возможность интенсифицировать утилизацию инфраструктурного ИТ-потенциала, оперативно реагировать на волатильность рыночной конъюнктуры, катализировать инновационную деятельность и максимизировать качественные метрики клиентских сервисов [6-А; 8-А].

Таким образом, эволюция архитектурных концептов от монолита к микросервисным решениям выступает весомым предикатором роста конкурентоспособности высоконагруженных информационных комплексов и их полной валидности вызовам современной цифровой экономики.

Четвертая глава диссертации посвящена математическому моделированию микросервисной интеллектуальной системы документооборота с использованием моделей

BERT. Проведённое исследование позволило разработать и обосновать математическую и архитектурную модель интеллектуальной системы документооборота, основанной на микросервисном подходе и применении трансформерных моделей класса BERT. Полученные результаты имеют как теоретическую, так и значительную практическую ценность, особенно в контексте цифровой трансформации государственных и корпоративных структур Республики Таджикистан.

В работе используется синтез архитектурного проектирования распределённых систем и технологий NLP (Natural Language Processing – обработка естественного языка). Данный инструментарий выбран для обеспечения высокой производительности системы: способности выдерживать нагрузки, минимизации сбоев и достижения максимальной достоверности при анализе текстов. Технология NLP объединяет методы лингвистики и машинного обучения для автоматизации работы с текстами: перевода, анализа тональности, создания чат-ботов и умных помощников [14; 21].

Математическая постановка задачи NER: В рамках формализации процедуры распознавания именованных сущностей исходный текстовый документ представляется как упорядоченная последовательность дискретных токенов (token – базовая неделимая единица текстовой информации: слово, часть слова, символ или знак пунктуации):

$$\mathcal{X} = (x_1, x_2, \cdots, x_n),$$

где x_j – это j-ый токен, полученный в процессе токенизации.

Целевая задача извлечения структурированных данных сводится к классификации последовательностей (Sequence Labeling) – поиску наиболее вероятной цепочки меток:

$$\mathcal{Y} = (y_1, y_2, \cdots, y_n), y_j \in \mathcal{C},$$

где $\mathcal{C}$ – конечное множество классов BIO-разметки.

При этом каждый элемент y_j выбирается из фиксированного алфавита классов BIO-разметки (Beginning-Inside-Outside – стандартная схема кодирования токенов). В терминах вероятностного моделирования последовательностей итоговое уравнение оптимизации для вектора параметров трансформерной модели BERT (θ) формулируется как задача условной максимизации:

$$\hat{\mathcal{Y}} = \arg \underbrace{\max}_{\mathcal{Y}} \mathcal{P}(\mathcal{Y} \mid \mathcal{X};\ \theta).$$

Данный математический аппарат позволяет интерпретировать контекстное извлечение сущностей как максимизацию совместной условной вероятности всей цепочки меток при заданном векторе входных токенов.

Механизм самовнимание в трансформерной архитектуре: Для вычисления указанной условной вероятности элементы последовательности $\mathcal{X}$ проецируются через матрицу векторных представлений (embedding matrix) в плотные непрерыные эмбеддинги, формируя матрицу признаков $\mathcal{E} \in \mathbb{R}^{n \text{x} d}$, где d – размерность скрытого пространства модели. На основе полученной матрицы вычисляются линейные проекции – матрицы запросов ($\mathcal{Q}$), ключей ($\mathcal{K}$) и значений ($\mathcal{V}$):

$$\mathcal{Q} = \mathcal{E}\mathcal{W}^{\mathcal{Q}}, \mathcal{K} = \mathcal{E}\mathcal{W}^{\mathcal{K}}, \mathcal{V} = \mathcal{E}\mathcal{W}^{\mathcal{V}},$$

где $\mathcal{W}^{\mathcal{Q}}, \mathcal{W}^{\mathcal{K}}, \mathcal{W}^{\mathcal{V}} \in \mathbb{R}^{d \text{x} d_k}$ – матрицы обучаемых весов линейных слоев.

Математический базис концепции скалярного произведения с наложением маски самовнимания (Attention Mechanism) описывается слудующим выражением:

$$\mathcal{A}(\mathcal{Q}, \mathcal{K}, \mathcal{V}) = softmax\left(\frac{\mathcal{Q}\mathcal{K}^T}{\sqrt{d_k}}\right)\mathcal{V},$$

в котором символами $\mathcal{Q}, \mathcal{K}, \mathcal{V}$ обозначены матрицы запросов, ключей и значений соответственно. Параметр d_k отражает размерность пространства ключей, определяющая масштабный коэффициент для нормирования градиентов функции $softmax$ при оценке интерференции токенов [5-А].

Развитие базовой архитектуры в терминах многоголового внимания (Multi-Head Attention) формализуется с помощью системы уравнений, позволяющих агрегировать информацию из различных скрытых подпространств [5-А]:

$$\mathcal{M}(\mathcal{Q},\mathcal{K},\mathcal{V}) = Concat(\mathcal{H}_1,\mathcal{H}_2,\cdots,\mathcal{H}_h)\mathcal{W}^o, \text{ где } \mathcal{H}_i = \mathcal{A}\left(\mathcal{Q}\mathcal{W}_i^{\mathcal{Q}},\mathcal{K}\mathcal{W}_i^{\mathcal{K}},\mathcal{V}\mathcal{W}_i^{\mathcal{V}}\right).$$

Указанный алгоритм детерминирует способность нейросети улавливать долгосрочные семантические и синтаксические корреляции внутри текстовой последовательности.

Функциональный слой токенизации и классификации: На финальной фазе детекции и структурирования информационных единиц латентные эмбеддинги, сгенерированные предобученной моделью BERT, транслируются на полносвязный линейный слой с целью аппроксимации распределения вероятностей BIO-меток (Begin, Inside, Outside).

Предположим, что упорядоченное множество $\mathcal{H} = (\hbar_1,\hbar_2,\cdots,\hbar_n)$ репрезентирует последовательность скрытых представлений (векторов) на ивыходе последнего слоя модели BERT. В данной нотации каждый вектор $\hbar_j \in \mathbb{R}^d$ $(j \in \{1,\dots,n\})$ представляет собой контекстно-зависимый эмбеддинг j-го токена анализируемой цепочки, а константа d фиксирует размерность скрытого представления трансформерной конфигурации.

Расчёт вектора логитов $z_j \in \mathbb{R}^{|\mathcal{C}|}$ для каждого j-го токена осуществляется путем линейного преобразования его скрытого состояния $\hbar_j$:

$$z_j = \mathcal{W}\hbar_j + \mathcal{b},$$

где $\mathcal{W} \in \mathbb{R}^{|\mathcal{C}|\text{x}d}$ – матрица обучаемых весов классификационного слоя, $\mathcal{b} \in \mathbb{R}^{|\mathcal{C}|}$ – вектор смещения (bias), а $|\mathcal{C}|$ – мощность множества классов BIO-разметки (количество целевых меток).

Вероятностная модель классификации для определения принадлежности j-го токена к конкретному классу $c \in \mathcal{C}$ формализуется с помощью функции softmax над векторном логитов z_j:

$$\mathcal{P}\left(y_j = c \mid \hbar_j\right) = \frac{exp\left(z_{j,c}\right)}{\sum_{c' \in \mathcal{C}} exp\left(z_{j,c'}\right)},$$

где $z_{j,c}$ – компоненты вектора логитов, соответствующая классу c, а y_j – рекзультирующая классификационная метка для j-го токена послодовательности.

В данном контексте softmax – это математическая функция, которая принимает на вход вектор произвольных вещественных чисел (логитов) и преобразует его в вектор вероятностей.

Функция потерь и процедура оптимизации: Для обучения параметров трансформерной модели θ используется функция многоклассовой перекрестной энтропии (Cross-Entropy Loss). Специфика токенизации субсловоформ (WordPiece) в условиях ограниченного объёма таджикского корпуса данных требует модификации процедуры расчёта потерь. Чтобы исключить искажение градиентов, из процесса оптимизации исключаются все субтокеновые фрагменты слов (замаскированные индексом – 100), за исключением первого.

Математическое выражение для вычисления функции потерь на уровне последовательности имеет вид:

$$\mathcal{L}(\theta) = -\sum_{j \in \mathcal{J}} \sum_{c \in \mathcal{C}} y_{j,c} \log \mathcal{P}\left(y_j = c \mid \hbar_j; \theta\right),$$

где $\mathcal{J}$ – подмножество индексов токенов, являющихся первыми (или единственными) субтокенами исходных слов в последовательности, $\mathcal{C}$ – множество классов BIO-разметки, а $y_{j,c} \in \{0,1\}$ – индикаторная переменная, равная единице, если класс c является истинной меткой j-го токена (one-hot encoding).

Задача обучения параметров модели формулируется как минимизация суммарной функции потерь на всем обучающем множестве:

$$\theta^* = \arg \underbrace{\min}_{\theta} \mathcal{L}(\theta).$$

Метрики экспериментальной оценки качества: Оценка эффективности формализованной интеллектуальной системы при извлечении структурированной информации выполняется на основе строгих критериев соответствия выделенных сущностей (Entity-level evaluation). Вычисление метрик точности (precision – $\mathcal{P}_r$), полноты (recall – $\mathcal{R}_c$) и $\mathcal{F}_1$-меры ($\mathcal{F}_1$-measure) осуществляется на основе подсчёта матрицы ошибок, соттветственно, по формулам [4-A]:

$$\mathcal{P}_r = \frac{\mathcal{T}_\mathcal{P}}{\mathcal{T}_\mathcal{P} + \mathcal{F}_\mathcal{P}};\ \mathcal{R}_c = \frac{\mathcal{T}_\mathcal{P}}{\mathcal{T}_\mathcal{P} + \mathcal{F}_\mathcal{N}};\ \mathcal{F}_1 = \frac{2\mathcal{P}_r\mathcal{R}_c}{\mathcal{P}_r + \mathcal{R}_c},$$

где $\mathcal{T}_\mathcal{P}$ (true positives) – количество истинно положительных исходов (границы и тип именованной сущности в таджикском тексте полностью совпали с эталонной разметкой); $\mathcal{F}_\mathcal{P}$ (false positives) – количество ложноположительных исходов (система ошибочно выделила несуществующую сущность или неверно определила её границы/тип); $\mathcal{F}_\mathcal{N}$ (false negatives) – количество ложноотрицательных исходов (система протустила эталонную именованную сущность).

Метрики оценки эффективности и мсаштабируемости распределённой системы: Для комплексной валидации разработанной интеллектуальной системы документооборота и оценки влияния синтетических данных на качество извлечения информации вводятся следующие аналитические показатели.

a) Относительный прирост качества извлечения. Изменение интегрального критерия эффективности ($\mathcal{F}_1$-меры) предложенного нейросетевого метода на базе модели mBERT по отношению к традиционному детерминированному подходу на основе регулярных выражений (Regex) рассчитывается по формуле:

$$\Delta\mathcal{F}_1 = \frac{\mathcal{F}_1^{BERT} - \mathcal{F}_1^{Regex}}{\mathcal{F}_1^{Regex}} \cdot 100\%.$$

b) Эффективность использования синтетических обучающих данных: Для оценки релевантности сгенерированного синтетического корпуса документов на таджикском языке вычисляется отношение качества модели, обученной на искусственных данных, к качеству модели, обученной на реальном размеченном подкорпусе:

$$\mathcal{Q}_{rel} = \frac{\mathcal{F}_1^{synthetic}}{\mathcal{F}_1^{real}}.$$

c) Временные и нагрузочные характеристики производительности системы: Среднее время обработки одного документа ($\mathcal{T}_{avg}$) и общая пропускная способность ИТ-инфраструктуры (Θ) при объёме выборки в $\mathcal{N}$ документов определяется как:

$$\mathcal{T}_{avg} = \frac{1}{\mathcal{N}}\sum_{j=1}^{\mathcal{N}} t_j\,,\ \Theta = \frac{\mathcal{N}}{\mathcal{T}_{total}},$$

где t_j – время обработки j-го документа индивидуальным микросервисом, а $\mathcal{T}_{total}$ – суммарное время выполнения пакета задач распределённой системой.

d) Показатели горизонтального масштабирования: При развёртывании k независимых экземпляров (подов/контейнеров) NLP-микросервиса в распределённой архитектуре теоретическая пропускная способность системы масштабируется линейно [4-A]:

$$\Theta(k) \approx k \cdot \Theta(1).$$

Реальный коэффициент масштабируемости системы $\mathcal{S}(k)$, отражающий эффективность параллелной обработки данных и накладные расходы на сетевой gRPC-коммуникатор, формализуется как:

$$\mathcal{S}(k) = \frac{\mathcal{T}(1)}{\mathcal{T}(k)},$$

где $\mathcal{T}(1)$ – время выполнения фиксированного объёма вычислений на одном вычислительном узле, а $\mathcal{T}(k)$ – время выполнения того же объёма вычислений на k параллелных узлах микросервисной архитектуры.

е) Гибридный механизм обеспечения отказоустойчивости (Fallback): Для минимизации рисков простоя системы при возникновении сетевых сбоев или перегрузки GPU-мощностей трансформерной модели, архитектура предусматривает автоматическое переключение на резервный алгоритмический слой. Итоговая функция извлечения структурированной информации $\mathcal{F}_{extract}(\mathcal{X})$ для входного документа $\mathcal{X}$ декомпозируется следующим образом [5-А]:

$$\mathcal{F}_{extract}(\mathcal{X}) = \begin{cases} \mathcal{F}_{BERT}(\mathcal{X}), \text{если микросервис } NLP \text{ доступен} \\ \mathcal{F}_{Regex}(\mathcal{X}), \text{если микросервис } NLP \text{ недоступен.} \end{cases}$$

Разработанная система обеспечивает формализованное представление процессов обработки документов в виде распределённой микросервисной архитектуры, описанной через ориентированный граф взаимодействующих сервисов. В рамках модели реализована вероятностная постановка задачи извлечения именованных сущностей, формализован механизм самовнимания трансформеров и предложены методы количественной оценки качества и производительности системы.

Интеграция многоязычных моделей BERT с механизмами генерации синтетических данных позволила достичь высокого уровня точности обработки документов (F1-score ≈ 0,93), что существенно превосходит традиционные rule-based подходы. При этом внедрение гибридного механизма обработки, сочетающего методы машинного обучения и правил-ориентированные алгоритмы, обеспечило дополнительную отказоустойчивость системы. Механизм адаптивного переключения (fallback) гарантирует непрерывность функционирования даже при временной недоступности нейросетевых сервисов, что критически важно для государственных информационных систем.

Модель адаптирована под структуру Центрального аппарата Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан (Комитет), включающего восемь функциональных подразделений (рисунок 4.1).

Практическая реализация микросервисной системы документооборота на базе моделей BERT, адаптированной под структуру Комитета, представляет собой комплексное решение, включающее (рисунок 4.2):

- *архитектурную модель* – описание функциональных и инфраструктурных компонентов;
- *математическую формализацию* – описание механизмов внимания, метрик эффективности и динамики информационных потоков;
- *логику внедрения* – интеграцию модулей в бизнес-процессы восьми подразделений с использованием гибридной модели отказоустойчивости.



Рисунок 4.1. Структура Центрального аппарата Комитета

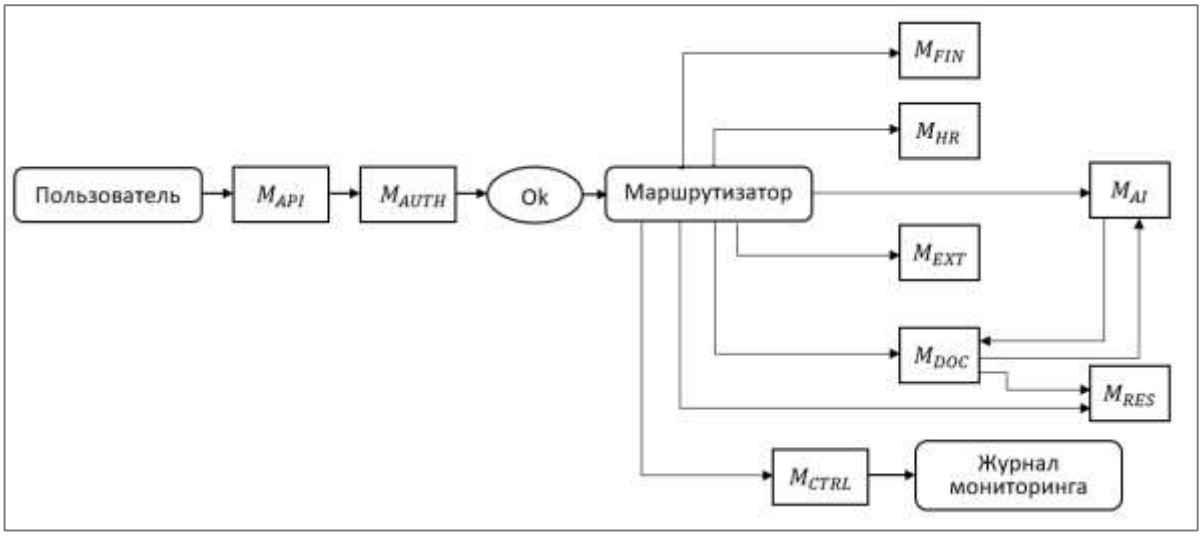


Рисунок 4.2. Алгоритм полного цикла реализации микросервисной модели интеллектуальной системы документооборота Комитета по профобразованию

Центральный аппарат Комитета рассматривается как распределённая система: $\mathcal{S} = \langle \mathcal{U}, \mathcal{D}, \mathcal{M}, \mathcal{A}, \mathcal{C} \rangle$, где $\mathcal{U}$ – множество подразделений Комитета, $\mathcal{D}$ – поток документов Комитета, $\mathcal{M}$ – множество микросервисов, $\mathcal{A}$ – интеллектуальный модуль (BERT), $\mathcal{C}$ – коммуникационная среда (gRPC/API).

Подразделения Комитета: $\mathcal{U} = \{U_1, U_2, U_3, U_4, U_5, U_6, U_7, U_8\}$, U_1 – «Начальное профессиональное образование», U_2 – «Среднее профессиональное образование», U_3 – «Финансы и бухгалтерия», U_4 – «Кадры и секретариат», U_5 – «Международные отношения», U_6 – «Право и контроль», U_7 – «IT и цифровизация», U_8 – «Хозяйственный отдел».

Поток документов (ключевая часть модели): Каждый документ Комитета представляется следующим образом: $d_i = (x_i, y_i, m_i, t_i)$, где, x_i – текст документа, y_i – извлечённые сущности, m_i – метаданные, t_i – время обработки.

Функция обработки документа: $y_i = f_{AI}(x_i)$, где f_{AI} – модель BERT: $\hat{y} = \arg\max_{y} \mathcal{P}(y \mid x;\ \theta)$, где θ – параметры модели BERT. Такое представление соответствует вероятностной модели последовательной разметки.

Архитектура: Система разбивается на множества микросервисов, соответствующих подразделениям Комитета (таблица 4.1): $\mathcal{M} = \{M_{DOC}, M_{FIN}, M_{HR}, M_{EXT}, M_{CTRL}, M_{AI}, M_{RES}\}$.

Помимо микросервисов, перечисленных в таблице 4.1, в математической модели учтены еще два инфраструктурных компонента (микросервиса) M_{API} (шлюз) и M_{AUTH} (безопасность), которые являются сквозными общесистемными функциями платформы. Они обеспечивают связность, взаимодействие и безопасность для всех функциональных микросервисов, но вынесены за рамки привязки к конкретным подразделениям. В этом случае математическая модель структуры системы представиться в виде кортежа:

$$\mathcal{M} = \langle M_{API}, M_{AUTH}, \{M_{DOC}, M_{FIN}, M_{HR}, M_{EXT}, M_{CTRL}, M_{AI}, M_{RES}\}\rangle.$$

Таблица 4.1. Соответствие микросервисов подразделениям Комитета

Подразделение Комитета	**Архитектурное название микросервиса**	**Функциональное название микросервиса**
Образование (U_1, U_2)	M_{DOC}	Сервис документов (Document Service)
Финансы и бухгалтерия (U_3)	M_{FIN}	Финансовый сервис (Financial Service)
Кадры и секретариат (U_4)	M_{HR}	Сервис кадров (HR Service)
Международные отношения (U_5)	M_{EXT}	Сервис внешних коммуникаций (External Communication Service)
Право и контроль (U_6)	M_{CTRL}	Сервис управления и контроля (Control Service)
IT и цифровизация (U_7) – информационные технологии	M_{AI}	Сервис ИИ и инфраструктуры (AI & Infra Service)
Хозяйство (U_8)	M_{RES}	Сервис ресурсов и архива (Resouce Service)

В таком виде модель обеспечивает модульность и масштабируемость системы, позволяя независимо развивать функциональность каждого подразделения Комитета без нарушения целостности общесистемных функций. Информационный обмен между компонентами модели строится на базе единой шины событий или REST-взаимодействия (интеллектуального AI-модула).

Для Комитета метрики эффективности системы (точность-$\mathcal{P}_r$, полнота-$\mathcal{R}_c, \mathcal{F}_1$-баллы) критически важны при обработке приказов (U_1, U_2) и финансовых документов (U_3), так как любая ошибка ($\mathcal{F}_{\mathcal{P}}$) или пропуск ($\mathcal{F}_{\mathcal{N}}$) в извлечении сумм или ФИО может привести к юридическим неточностям. Для обеспечения надежности документооборота в Комитете установлены следующие пороговые значения метрик: $\mathcal{P}_r \geq 0{,}92;\ \mathcal{R}_c \geq 0{,}90;\ \mathcal{F}_1 \geq 0{,}85$. Среднее время обработки документов $\mathcal{T}_{avg} = \frac{1}{\mathcal{N}}\sum_{j=1}^{\mathcal{N}} t_j$, где $\mathcal{N}$ – общее количество обработанных документов за расчетный период, t_j – время, затраченное на обработку конкретного j-го документа (от момента поступления в микросервис до выдачи результата). Пропускная способность $\Theta = \frac{\mathcal{N}}{\mathcal{T}_{total}}$, где Θ – количество документов, которые система способна обработать в единицу времени (например, *док/сек*), $\mathcal{N}$ – количество выполненных задач, $\mathcal{T}_{total}$ – общее время наблюдения или работы системы. Это главный показатель мощности (throughput). Если Θ меньше интенсивности входящего потока документов, в системе будет расти очередь.

Гибридная модель отказоустойчивости (Fallback-механизм):

$$f(x) = \begin{cases} f_{AI}(x), & \text{если AI доступен} \\ f_{rule}(x), & \text{если AI недоступен.} \end{cases}$$

При снижении уверенности f_{AI} ниже порога ρ<0,7 или технической недоступности GPU-мощностей, система переключается на предиктивный анализ по правилам f_{rule}.

Итоговая система (интегральная модель): Накопления очереди документов описывается дифференциальным уравнением:

$$\frac{dD}{dt} = \lambda_{in} - \lambda_{proc}(\mathcal{M}, \mathcal{A}),$$

где λ_{in} – интенсивность входящего потока, λ_{proc} – скорость обработки документов. Изменение объема необработанных документов dD/dt зависит от баланса входящего трафика λ_{in} и производительности микросервисной архитектуры λ_{proc}. Условием стабильности системы является $\lambda_{proc} > \lambda_{in}$, иначе очередь D будет расти бесконечно.

Значимым научным результатом выступает формализация и алгоритмическое описание сквозного конвейера обработки документации, базирующегося на событийно-управляемой (event-driven) архитектуре взаимодействия микрослужб, асинхронных механизмах обмена сообщениями и детерминированном контуре диспетчеризации процессов (рисунок 4.3).

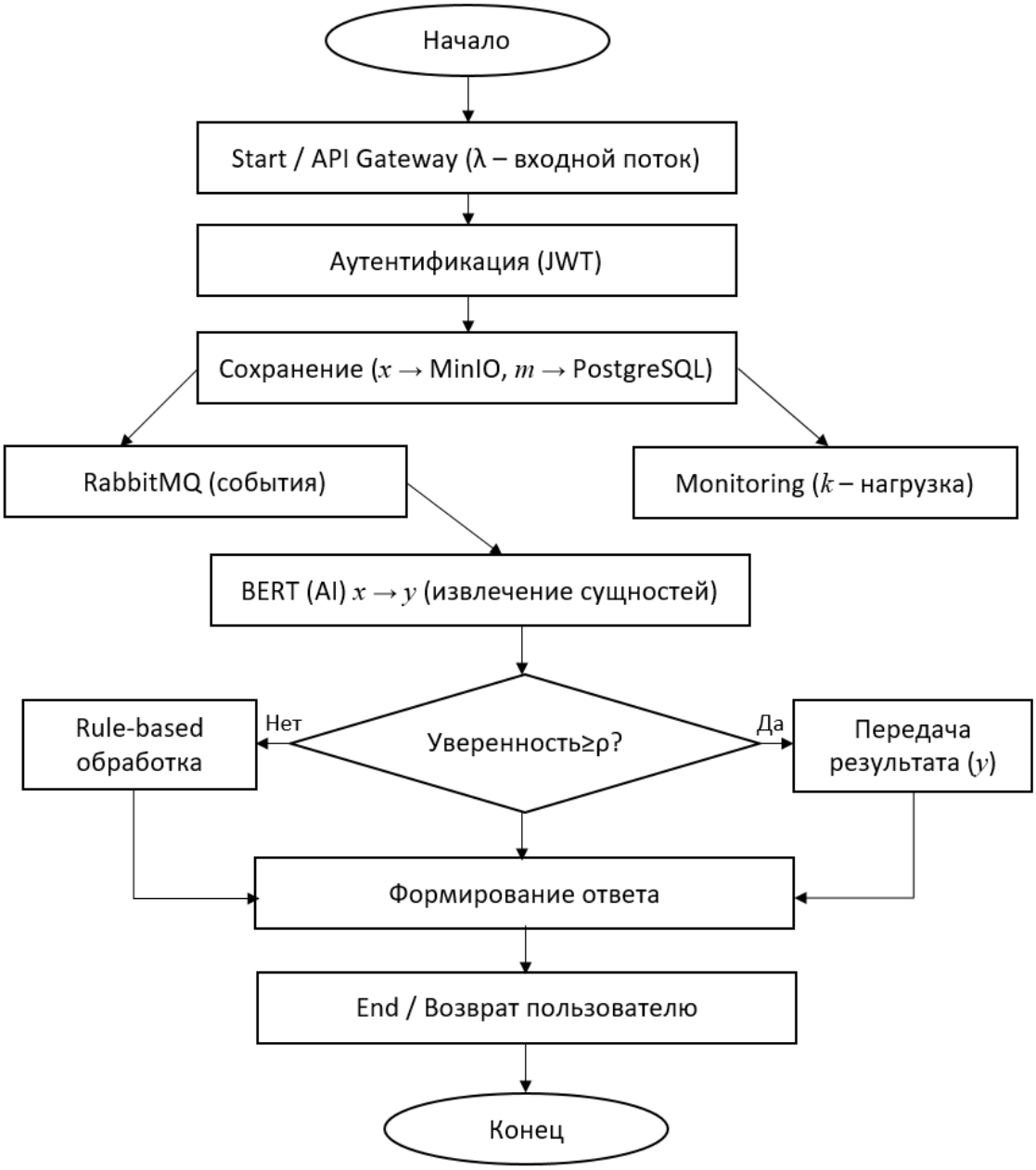


Рисунок 4.3. Блок-схема алгоритма функционирования микросервисного интеллектуального документооборота, формализующая процессы внутри системы

Реализация данного комплекса обеспечила линейное масштабирование платформы посредством горизонтального наращивания вычислительных узлов, минимизацию латентности при оперировании информационными потоками и интенсификацию общей пропускной способности. Опытно-промышленная эксплуатация подтвердила падение временных затрат на обработку документарных массивов на 60-70% при сохранении стабильно высоких показателей отказоустойчивости в условиях пиковых нагрузок.

Интеграция разработанной интеллектуальной микросервисной платформы в контур управления Комитета обеспечивает достижение фиксированных технико-экономических и эксплуатационных параметров:

1. *Интенсификация документооборота:* Ускорение процессов обработки входящей документации на 60-70% за счёт автоматизации процедур распознавания именованных сущностей (NER) и кластеризации, что позволяет радикально сократить долю рутинных ручных операций.

2. *Верификация и валидация данных:* Рост точности и релевантности извлекаемой информации до уровня 92-95% (в соответствии с контрольными критериями прецизионности $\mathcal{P}_r$ и интегральной метрики $\mathcal{F}_1$), что существенно нивелирует вероятность возникновения правовых и финансовых коллизий в деятельности структурных подразделений U_3 и U_4

3. Снижение операционных затрат на поддержку IT-инфраструктуры благодаря горизонтальному масштабированию и использованию свободных вычислительных мощностей для работы модели BERT.

4. Оптимизация контроля исполнения: сокращение цикла согласования документов между подразделениями ($U_i \to U_j$) на 40% за счет сквозного мониторинга в M_{CTRL}.

Предложенная система адаптируема к структуре государственных учреждений, учитывает особенности многоязычной среды (таджикский, русский, английский языки), а также ограничения вычислительной инфраструктуры. Возможность поэтапного внедрения, использование контейнеризации и оркестрации, а также оптимизация моделей (включая методы сжатия) делают решение применимым в условиях ограниченных ресурсов и развивающейся ИТ-инфраструктуры страны.

В то же время выявлены определённые ограничения: задержки нейросетевого инференса делают систему более эффективной для пакетной обработки документов, а обработка рукописных данных требует интеграции OCR-технологий. Анализ ошибок показал необходимость расширения обучающих выборок и внедрения доменно-специфичных правил для повышения качества извлечения сложных сущностей.

В целом предложенная модель представляет собой целостное архитектурно-алгоритмическое решение, обеспечивающее высокую точность, масштабируемость и надёжность. Она может служить основой для разработки интеллектуальных систем документооборота, способствуя развитию электронного управления и повышению эффективности цифровых процессов.

Перспективы дальнейших исследований связаны с оптимизацией трансформерных моделей, расширением мультиязычной поддержки, внедрением механизмов самообучения и совершенствованием распределения нагрузки в микросервисной среде.

ВЫВОДЫ

1. Основные научные результаты диссертации

В ходе диссертационного исследования получены следующие основные научные результаты:

1. Сформировано теоретическое и прикладное представление о развитии микросервисной архитектуры как основы построения масштабируемых интеллектуальных систем, позволяющей обеспечить декомпозицию функциональности, автономность компонентов и устойчивость к отказам [6-А; 7-А].

2. Разработана архитектурная модель интеллектуальной системы документооборота, основанная на принципах микросервисного подхода, включающая выделенный компонент интеллектуального анализа (AI Service), что обеспечивает независимую масштабируемость вычислительно сложных задач и гибкость развития системы [3-А].

3. Предложена математико-алгоритмическая модель обработки документов, основанная на применении трансформерной архитектуры BERT, обеспечивающая высокую точность извлечения именованных сущностей и структурирование текстовой информации [1-А; 4-А].

4. Разработан метод адаптации многоязычных моделей к условиям низкоресурсных языков (таджикского) на основе синтетической генерации обучающих данных и дообучения, позволяющий достигать качества, сопоставимого с высокоресурсными языками [5-А.

5. Предложен гибридный подход к обработке данных, объединяющий методы машинного обучения и правил, обеспечивающий одновременно высокую точность, устойчивость и интерпретируемость результатов [3-А].

6. Исследованы и систематизированы методы масштабирования микросервисных систем, включая горизонтальное и вертикальное масштабирование, а также обоснована эффективность их комбинированного применения в интеллектуальных системах [8-А; 10-А].

7. Обоснована роль контейнеризации и оркестрации как ключевых технологических механизмов управления жизненным циклом микросервисов, обеспечивающих автоматическое масштабирование, отказоустойчивость и оптимизацию использования ресурсов [3-А; 11-А].

8. Проведена экспериментальная валидация предложенной системы, подтвердившая её эффективность: достигнуто существенное повышение точности обработки документов и многократное сокращение времени их обработки [1-А].

9. Выявлены ограничения разработанного подхода, связанные с требованиями к вычислительным ресурсам, особенностями обработки неструктурированных и рукописных документов, а также необходимостью дополнительной адаптации моделей для новых предметных областей [2-А].

10. На основе разработанной математической модели интеллектуальной системы документооборота, базирующейся на микросервисной архитектуре и трансформерных моделях типа BERT, а также адаптированной под организационно-функциональную структуру Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан, разработан практический алгоритм реализации системы. Указанный алгоритм оформлен в инженерно-прикладном стиле и отражает последовательность этапов развертывания, интеграции и эксплуатации интеллектуальной системы в условиях реальной информационной инфраструктуры [1-А–5-А].

2. Рекомендации по практическому использованию результатов

Практическая ценность исследования существенно усиливается за счёт того, что предложенные решения не ограничиваются теоретическим уровнем, а доведены до стадии инженерной реализации (Приложение 1). Разработанный алгоритм внедрения, сформированный на основе математической модели системы и адаптированный к условиям функционирования Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан (Приложение 2), может быть использован в качестве типового подхода при цифровизации документооборота в государственных учреждениях.

Полученные в диссертационном исследовании результаты могут быть рекомендованы к практическому применению в следующих направлениях:

1. *Государственный сектор*: Рекомендуется внедрение разработанной системы в органах государственного управления, в том числе с использованием предложенного

инженерно-ориентированного алгоритма реализации, обеспечивающего поэтапную интеграцию интеллектуальной системы в существующую инфраструктуру документооборота.

2. *Коммерческие организации:* Целесообразно использование системы для автоматизации обработки договоров, счетов-фактур и других юридически значимых документов, что позволит повысить эффективность бизнес-процессов и снизить операционные издержки.

3. *Финансовый сектор:* Рекомендуется применение решений для реализации процедур KYC, обработки кредитных заявок и проведения комплаенс-проверок с использованием интеллектуального анализа документов.

4. *ИТ-инфраструктура организаций:* Рекомендуется использование микросервисной архитектуры, контейнеризации и оркестрации для построения масштабируемых и отказоустойчивых систем обработки данных.

5. *Развитие систем искусственного интеллекта:* Предложенные методы адаптации моделей BERT могут быть использованы при разработке NLP-систем для других низкоресурсных языков и специализированных предметных областей.

6. *Научные исследования и образование:* результаты могут применяться в исследованиях распределённых систем, машинного обучения и цифровой трансформации, а также в образовательных курсах по информационным технологиям.

7. *Дальнейшее развитие:* предполагается совершенствование предложенного подхода по нескольким направлениям – оптимизация моделей (дистилляция, квантование), внедрение мультимодальной обработки данных, повышение интерпретируемости моделей и интеграция с промышленными платформами управления жизненным циклом ML (Machine Learning – машинное обучение).

Разработанная в диссертации математическая модель, архитектурная реализация и прикладной алгоритм внедрения образуют целостную научно-прикладную основу для построения интеллектуальных систем документооборота нового поколения. Адаптация решения к структуре Комитета по начальному и среднему профессиональному образованию при Правительстве Республики Таджикистан подтверждает его практическую применимость и соответствие направлениям цифрового развития страны.

ЛИТЕРАТУРА

1. Список использованных источников

[1] Васвани, А. Внимание – это все, что вам нужно [Текст] / А. Васвани, Н. Шазир, Н. Пармар и др. // Достижения в системах обработки нейронной информации. – 2017. – Т. 30. – С. 5998-6008.

[2] Воинов, А. А. Имитационная модель Кайраккумского водохранилища [Текст] / А.А. Воинов, Ф.С. Комилов. – М.: ВЦ АН СССР, «Сообщения по прикладной математике», 1986. – 40 с.

[3] Гадоев, А. Ф. Разработка и реализация управленческих решений на предприятиях легкой промышленности Таджикистана в условиях ускоренной индустриализации. Монография [Текст] / А.Ф. Гадоев, А. Абдалимов, Ф.С. Комилиён. – Душанбе: Бахманруд, 2022. – 155 с.

[4] Девлин, Дж. BERT: предобучение глубоких двунаправленных трансформеров для понимания языка [Текст] / Дж. Девлин, М.В. Чанг, К. Ли, К. Тутанова // Труды конференции Североамериканского отделения Ассоциации компьютерной лингвистики 2019. – 2019. – С. 4171-4186.

[5] Ёров, М. Р. Татбиқи системаи хизматрасонии оммавӣ дар коркарди дархостҳои онлайнӣ [Матн] / М.Р. Ёров, Ф.С. Комилиён // Паёми Донишгоҳи миллии Тоҷикистон. Бахши илмҳои табиӣ. – 2023. – № 2. – С. 42-53. – EDN: QTPBDV.

[6] Комилиён, Ф. С. Амсилаи консептуалии низоми иттилоотии идоракунии сифати хизматрасониҳои таълимӣ дар муассисаи таҳсилоти олии касбӣ [Матн] / Ф.С. Комилиён, З.Ф. Раҳмонзода // Паёми Донишгоҳи миллии Тоҷикистон. Бахши илмҳои иҷтимоӣ-иқтисодӣ ва ҷамъиятӣ. – 2021. – № 3. – С. 127-135.

[7] Комилиён, Ф. С. Амсиласозии компютерии системаи хизматрасонии шабакавии дар вақти дискретӣ бо тартиби инверсионӣ ва афзалияти тасодуфӣ амалкунандаи ПД КОА [Матн] / Ф.С. Комилиён, М.Р. Ёров // Паёми Донишгоҳи миллии Тоҷикистон. Бахши илмҳои табиӣ. – 2020. – №2. – С. 68-79. – EDN: AJPQAI.

[8] Комилиён, Ф. С. Информатика. Китоби дарсӣ барои донишҷӯёни муассисаҳои таҳсилоти олии касбӣ. Ҷилди 2 [Матн] / Ф.С. Комилиён. – Душанбе: «Донишварон», 2019. – 464 с. – EDN: RNYCDB.

[9] Концепция формирования электронного правительства Республики Таджикистан. Утверждена постановлением Правительства Республики Таджикистан от 30 декабря 2011 года № 643.

[10] Национальная стратегия развития образования Республики Таджикистан на период до 2030 года. Утверждена постановлением Правительства Республики Таджикистан от 29 сентября 2020 года № 526. – 90 с.

[11] Программа реализации концепции государственной информационной политики Республики Таджикистан. Утверждена постановлением Правительства Республики Таджикистан от 28 мая 2009 года № 307.

[12] Саидзода, И. М. Барномасозии компютерии фаъолияти занбӯри асал вобаста ба таъсири бемориҳо ва зараррасонҳо [Матн] / И.М. Саидзода, Ф.С. Комилиён // Паёми донишгоҳи давлатии Хоруғ. – 2023. – № 1 (25). – С. 103-115.

[13] Burns, B. Kubernetes: Up and Running [Text] / B. Burns, J. Beda, K. Hightower. – Sebastopol: O'Reilly Media, 2019. – 330 p.

[14] Chiticariu, L. Rule-based information extraction [Text] / L. Chiticariu, Y. Li, F. Reiss // Proceedings of EMNLP 2013. – 2013. – P. 827-832.

[15] Devlin, J. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [Text] / J. Devlin, M. W. Chang, K. Lee, K. Toutanova // Proceedings of NAACL-HLT. – 2019. – P. 4171-4186.

[16] Earl, T. Service-Oriented Architecture: Concepts, Technologies, and Design [Text] / T. Earl. – Upper Saddle River: Prentice Hall, 2005. – 760 p.

[17] Fowler, M. Guide to Microservices [Electronic resource] / M. Fowler // URL: https://martinfowler.com/microservices/(Date of the application: 5.07.2022).

[18] Fowler, M. Microservices Architecture [Электронный ресурс] / M. Fowler // URL: https://martinfowler.com/microservices/ (Date of the application: 5.07.2022).

[19] Mazlami, G. Extraction of microservices from monolithic software architectures [Electronic resource] / G. Mazlami et al. // International Conference on Web Services (ICWS), 2017, IEEE. – P. 524-531.

[20] Newman, S. Building Microservices: Designing Fine-Grained Systems [Text] / S. Newman. – Sebastopol: O'Reilly Media, 2015. – 280 p.

[21] Ratner, A. Snorkel: Fast Training Data Generation [Text] / A. Ratner, S. Bach, H. Ehrenberg et al. // VLDB. – 2017. – Vol. 11. – No. 3. – P. 269-282.

[22] Richardson, K. Microservices: Development and Refactoring Patterns [Text] / K. Richardson. – Shelter Island: Manning Publications, 2018. – 520 p.

[23] Vaswani, A. Attention Is All You Need [Text] / A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin // Advances in Neural Information Processing Systems (NeurIPS). – 2017. – Vol. 30. – P. 5998-6008.

[24] Wolf, T. Transformers: State-of-the-Art Natural Language Processing [Text] / T. Wolf, L. Debut, V. Sanh et al. // Proceedings of EMNLP. – 2020. – P. 38-45.

2. Список научных публикаций соискателя ученой степени

а) Статьи, опубликованные в рецензируемых журналах Высшей аттестационной комиссии при Президенте Республики Таджикистан:

[1-А] Рахимов, М. Ф. Алгоритмизация микросервисной модели интеллектуального документооборота в профобразовании на базе BERT [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Вестник национального университета. Серия естественных наук. – 2026. – № 2. – С. 98-109. – EDN: OSUPXM.

[2-А] Рахимов, М. Ф. Внедрение микросервисной архитектуры для оптимизации распределения информационных ресурсов [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Наука и инновация. Серия геологических и технических наук. – 2024. – № 2. – С. 71-79. – EDN: UZJNVI.

[3-А] Рахимов, М. Ф. Интеллектуальная система документооборота на основе микросервисной архитектуры и моделей BERT [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Вестник национального университета. Серия естественных наук. – 2026. – № 2. – С. 76-85. – EDN: ROVNVV.

[4-А] Рахимов, М. Ф. Математическое моделирование интеллектуальной системы документооборота на основе микросервисной архитектуры и моделей BERT [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Вестник Технологического университета Таджикистана. – 2026. – № 3 (66). – С. 96-106. – EDN: RWEDOQ.

[5-А] Рахимов, М. Ф. Микросервисная архитектура и BERT-модели в интеллектуальном документообороте профобразования [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. – 2026. – № 2 (74). – С. 62-71. – EDN: VJVTAH.

[6-А] Рахимов, М. Ф. Микросервисная архитектура: от монолита к гибким распределенным системам [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Доклады Национальной академии наук Таджикистана. – 2023. – Т. 66. – № 11-12. – С. 659-667. – EDN: LSVHDL.

[7-А] Рахимов, М. Ф. Анализ характеристик монолитных и микросервисных архитектур [Текст] / М.Ф. Рахимов, Ф.С. Комилиён // Известия Национальной академии наук Таджикистана. Отделение физико-математических, химических, геологических и технических наук. – 2023. – № 4 (193). – С. 44-54. – EDN: GHUYHC.

б) Статьи, опубликованные в материалах конференций и в других журналах:

[8-А] Рахимов, М. Ф. Микросервисная оптимизация распределения информационных ресурсов с помощью четко определенного «API» [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Современные проблемы математического моделирования и её применения. Материалы XII международной научно-практической конференции, посвященной «2020-2040 годы – 20-летия изучения и развития естественных, точных и математических наук в области науки и образования» и «75-летию Таджикского национального университета». – Душанбе, 2024. – С. 28-32. – EDN: PZLKJR.

[9-А] Рахимов, М. Ф. Влияние микросервисных приложений на улучшение работы IT приложений, на примере использования Node.js [Текст] / М.Ф. Рахимов // Материалы республиканской научно-практической конференции «Актуальные проблемы и перспективы развития естественных и точных наук». – Душанбе, ФМГУ, 2023. – С. 12-15.

[10-А] Рахимов, М. Ф. Внедрение микросервисной архитектуры в монолитные системы [Текст] / М.Ф. Рахимов // Материалы международной научно-практической конференции «Наука и образование: тенденции развития в условиях

информационного общества», посвященной «75-ой годовщине со дня образования Таджикского национального университета». – Душанбе, ТНУ, 2023. – С. 8-11.

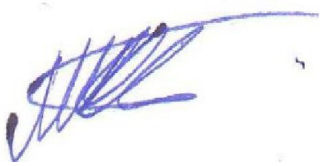
[11-А] Рахимов, М. Ф. Микросервисы в задачах оптимизации распределения информационных ресурсов организации [Текст] / М.Ф. Рахимов, Ф.С. Комилиён // Материалы республиканской научно-практической конференции: «Роль интернета в применении информационно-коммуникационных технологий и информационной безопасности в социальных сетях» – Душанбе, АГУ, 2024. – С. 16-24.

[12-А] Rahimov, M. F. Mathematical Modeling of an Intelligent Document Management System Based on Microservice Architecture and BERT Models [Electronic resource] / F. Rahimi, F.S. Komiliyon, M.F. Rahimov, M.R. Yorov // Applied and Computational Mathematics. – 2026. – Vol. 15. – No. 3. – P. 68-74. // https://doi.org/10.11648/j.acm. 20261503.11 (Date of the application: 14.05.2026).

АКАДЕМИЯИ МИЛЛИИ ИЛМҲОИ ТОҶИКИСТОН

Бо ҳуқуқи дастнавис

ТДУ 51:681.3+004.8
ТКБ 22.1+73
Р-27



Раҳимов Манучеҳр Фарҳодович

МОДЕЛСОЗИИ МАТЕМАТИКИИ СИСТЕМАИ ЗЕҲНИИ ГАРДИШИ САНАДҲО ДАР АСОСИ АРХИТЕКТУРАИ МИКРОСЕРВИСӢ ВА МЕХАНИЗМҲОИ НЕЙРОШАБАКАВИИ BERT

А В Т О Р Е Ф Е Р А Т И

диссертатсия барои дарёфти дараҷаи илмии доктори фалсафа (PhD)
аз рӯйи ихтисоси 6D060200 – Информатика (6D060204 – Асосҳои назариявии информатика)

Душанбе-2026

Диссертатсия дар шуъбаи моделсозии математикии Институти математикаи ба номи А. Ҷӯраеви Академияи миллии илмҳои Тоҷикистон иҷро шудааст

Роҳбари илмӣ: **Комилиён Файзалӣ Саъдулло** – доктори илмҳои физикаю математика, профессори кафедраи информатикаи Донишгоҳи миллии Тоҷикистон

Муқарризони расмӣ: **Худойбердиев Хуршед Атохонович –** доктори илмҳои техникӣ, дотсенти кафедраи барномарезӣ ва низомҳои иттилоотии Донишкадаи политехникии Донишгоҳи техникии Тоҷикистон ба номи академик М.С.Осимӣ дар шаҳри Хуҷанд,

Ғуломсафдаров Абдулназар Ғуломназарович – номзади илмҳои техникӣ, мудири кафедраи илмҳои компютерӣ ва технологияҳои рақамии Донишкадаи инноватсионии Душанбе

Муассисаи пешбар: Филиали Донишгоҳи давлатии Москва ба номи М.В. Ломоносов дар шаҳри Душанбе

Ҳимояи диссертатсия «17» декабри соли 2026, соати 14^{00} дар шурои диссертатсионии 6D.KOA-105 назди Донишгоҳи миллии Тоҷикистон, ки дар суроғаи: 734025, шаҳри Душанбе, кӯчаи Буни-Ҳисорак, бинои 17, аудиторияи 203 ҷойгир аст, баргузор мешавад.

E-mail: rohatullo1956@mail.ru; рақами телефони котиби илмӣ: +992981097033.

Бо диссертатсия ва автореферати он тавассути сомонаи http://www.tnu.tj ва дар китобхонаи илмии Донишгоҳи миллии Тоҷикистон шинос шудан мумкин аст.

Автореферат фиристода шудааст «___» «___» соли 2026

Котиби илмии шурои
диссертатсионии 6D.KOA-105, номзади
илмҳои физикаю математика, дотсент Садуллоев Р.И.

МУҚАДДИМА

Мубрамии мавзуи таҳқиқот. Афзоиши ҳаҷми маълумот, мураккаб гардидани коммуникатсияҳои дохилӣ ва талаботи доимӣ ба суръати қабули қарорҳои идоракунӣ боиси он гардидаанд, ки моделҳои маъмулии гардиши санадҳо беш аз пеш ба маҳдудиятҳои марбут ба маҳсулнокӣ ва устувории коркарди селҳои иттилоотӣ дучор мешаванд. Дар ташкилоту муассисаҳои дорои гардиши пуршиддати санадҳо ҳаҷми бузурги иттилооте, ки аз рӯйи сохтор ва таъйинот аз ҳам фарқ мекунанд, пайваста ҷамъ мешаванд. Мушкилоти зикршуда дар чунин шароит боз ҳам равшантар зоҳир мегарданд. Қисми зиёди расмиёт то ҳол ё ба таври дастӣ иҷро мегарданд, ё мунтазам бо дахолати омили инсонӣ сурат мегиранд. Аз ин сабаб, эҳтимолияти содир шудани хатогиҳо меафзояд, назорат душвор мегардад ва барои қабули қарорҳо вақти иловагӣ сарф карда мешавад.

Таҳти таъсири технологияҳои ҳисоббарории тақсимшуда талабот нисбат ба системаҳои муосир низ тағйир ёфтааст. Қиёспазирӣ ва устуворӣ ба нокомиҳо ҳамчун хусусиятҳои афзалиятнок арзёбӣ мегарданд. Ҳамзамон аз системаҳо қобилияти нигоҳ доштани маҳсулнокӣ дар шароити сарбории тағйирёбанда интизор меравад. Архитектураҳое, ки системаро ба сервисҳои мустақил ҷудо мекунанд, беш аз пеш интихоб мешаванд. Дар заминаи архитектураи микросервисӣ коркарди маълумот чандир гардида, рушди системаи барномавӣ аз ҳисоби декомпозитсия ба сервисҳои алоҳида содатар мешавад ва устуворияи фаъолият ҳатто дар шароити сарбориҳои назаррас таъмин мегардад.

Рақамикунонии идоракунии давлатӣ дар Ҷумҳурии Тоҷикистон ба самтҳои афзалиятноки рушд мансуб дониста мешавад. Санадҳои стратегӣ ва меъёрӣ ин афзалиятро муқаррар менамоянд, ки дар миёни онҳо Стратегияи миллии рушди Ҷумҳурии Тоҷикистон то соли 2030, Консепсияи ташаккули ҳукумати электронӣ дар Ҷумҳурии Тоҷикистон ва Барномаи татбиқи консепсияи сиёсати давлатии иттилоотии Ҷумҳурии Тоҷикистон қарор доранд. Дар санадҳои мазкур вазифаҳои рушди инфрасохтори рақамӣ ва истифодаи васеътари усулҳои зеҳнии коркарди иттилоот муайян гардидаанд.

Афзалиятҳои зикршуда талаботро ба системаҳои зеҳнии гардиши санадҳо меафзоянд. Чунин системаҳо бояд нигоҳдории санадҳо ва интиқоли онҳоро байни зерсохторҳо таъмин намоянд. Дар баробари ин, таҳлили автоматии муҳтаво бо истихроҷи маълумоти калидӣ, муайян намудани гурӯҳи санад ва дастгирии қабули қарорҳо талаб карда мешавад. Барои бунёди чунин системаҳо усулҳои моделсозии математикӣ бо технологияҳои ҳисоббарорӣ ва алгоритмҳои муосири коркарди матн ҳамгиро карда мешаванд.

Дар солҳои охир усулҳои коркарди забони табиӣ босуръат рушд ёфтанд. Моделҳои трансформерии навъи BERT (Bidirectional Encoder Representations from Transformers - Рамзгузор-намоишҳои дуҷониба аз трансформерҳо) абзорҳои таҳлили зеҳнии матнҳоро васеъ намуданд. Бо истифода аз онҳо масъалаҳои таснифоти семантикӣ, истихроҷи моҳиятҳои номдор, тафсири контекстии санадҳо ва таҳлили автоматикунонидашудаи муҳтаво нисбат ба равишҳои маъмулӣ дақиқтар ҳал карда мешаванд. Бо вуҷуди ин, ҷорӣ намудани моделҳои мазкур дар системаҳои гардиши санадҳо ба маҳдудиятҳои амалӣ рӯ ба рӯ мешавад. Ин маҳдудиятҳо хусусан барои шароити забонҳои камзахира, ки корпусҳои бузурги нишонагузоришуда надоранд ва аз норасоии моделҳои мутобиқгардонидашудаи забонӣ танқисӣ мекашанд, шадидтар зоҳир мегарданд.

Мушкилоти алоҳида аз интегратсияи алгоритмҳои зеҳнӣ ба муҳити ҳисоббарории тақсимшуда иборат аст. Дар аксари таҳқиқот тарҳрезии архитектурӣ ва таҳлили зеҳнии санадҳо ба таври ҷудогона омӯхта мешаванд. Бархе муаллифон диққати асосиро ба ташкили микросервисии системаҳои барномавӣ равона мекунанд, бархи дигар бошад - ба алгоритмҳои алоҳидаи NLP ё ба моделҳои нейрошабакавии коркарди матн. Дар ин замина, масъалаҳои интегратсияи комплексии архитектураи микросервисӣ ва моделҳои математикии таҳлили зеҳнии санадҳо ҳанӯз ба таври кофӣ таҳқиқ нашудаанд.

Аҳаммияти мавзуъ инчунин ба мутобиқсозии системаҳои зеҳнӣ ба шароити воқеии идоракунии давлатӣ алоқаманд мебошад. Дар доираи таҳқиқоти мазкур ҳамчун муҳити

татбиқ Кумита оид ба таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон баррасӣ мегардад. Чунин интихоб имкон медиҳад, ки қоидаҳои амалкунандаи коркарди санадҳо, сохтори ташкилии муассиса ва селҳои иттилоотие, ки дар раванди фаъолияти он ташаккул меёбанд, ба инобат гирифта шаванд. Шароити воқеӣ роҳҳои ҳалли таҳияшударо барои ҷорисозии минбаъда боз ҳам мувофиқтар мегардонад.

Мубрамии диссертатсия ба омилҳои зерин вобаста аст: баланд бардоштани самаранокии гардиши санадҳо зарур мебошад; ҳамзамон талабот ба архитектураҳои қиёспазир ва ба нокомиҳо устуворе, ки барои коркарди пуршиддати маълумот пешбинӣ шудаанд, меафзояд; абзорҳои таҳлили зеҳнии иттилооти матнӣ низ васеъ мегарданд; роҳҳои ҳалли мавҷудаи NLP ба муҳитҳои забонҳои камзахира заиф мутобиқ гардидаанд ва ин ҳолат татбиқи амалии онҳоро маҳдуд месозад; вазифаи бунёди моделҳои аз ҷиҳати математикӣ асоснокшудаи коркарди зеҳнии санадҳо бошад, ҳамчунон мубрам боқӣ мемонад..

.

Дараҷаи таҳқиқи мавзуи илмӣ. Мавзуи илмии «Моделсозии математикии системаи зеҳнии гардиши санадҳо дар асоси архитектураи микросервисӣ ва механизмҳои нейрошабакавии BERT» дар сатҳи ҷанбаҳои бунёдӣ ва амалӣ таҳқиқ гардида, таҳияи моделҳои математикӣ ва архитектурии система, усулҳои мутобиқсозии моделҳои трансформериро барои забонҳои камзахира, инчунин татбиқи амалӣ ва тасвиби онҳоро дар шароити инфрасохтори воқеии иттилоотӣ дар бар мегирад.

Зимни таҳияи маҷмуи моделҳои компютерӣ аз натиҷаҳои таҳқиқотии муҳаққиқони соҳаи коркарди забонҳои табиӣ васеъ истифода шудааст. Саҳми назаррасро дар ин самт Ashish Vaswani [23], Jacob Devlin [15] ва Thomas Wolf [24] гузоштаанд, ки таҳқиқотҳои онҳо асосҳои архитектурии трансформер ва модели BERT-ро ташкил дода, дақиқии баланди таҳлили маълумоти матниро таъмин менамоянд.

Ҷанбаҳои архитектурӣ ба корҳои илмии Sam Newman [19], Martin Fowler [17] ва Thomas Erl [16] такя мекунанд, ки дар онҳо архитектураи микросервисӣ ҳамчун усули самараноки сохтани системаҳои қиёспазир ва ба нокомиҳо устувор баррасӣ шудааст. Рушди ин ғояҳо дар таҳқиқоти Chris Richardson [22] идома ёфта, намунаҳои асосии барномаҳои тақсимшуда формализатсия шудаанд.

Дар заминаи системаҳои зеҳнӣ технологияҳои контейнеризатсия ва оркестратсия аҳаммияти махсус пайдо кардаанд. Истифодаи платформаҳои Docker ва Kubernetes, ки дар таҳқиқотҳои Chris Richardson [22], Brendan Burns ва Joe Beda [13] рушд ёфтаанд, изолятсияи сервисҳо, автоматизатсияи ҷойгиркунӣ ва идоракунии самараноки қиёспазириро таъмин менамоянд. Онҳо имкон медиҳанд, ки компонентҳои серхароҷоти омӯзиши мошинӣ ба инфрасохтори тақсимшуда ворид карда шаванд ва компонентҳои ҳисоббарориҳои мураккаб бо имконияти азнавтақсимшавии динамикии захираҳо ба ҳисоб гирифта шаванд.

Таҳлили гузаронидашуда нишон дод, ки интегратсияи архитектураи микросервисӣ бо моделҳои BERT заминаи системаҳои зеҳнии гардиши санадҳоро ташкил медиҳад, ки истихроҷ ва сохторбандии автоматии маълумотро таъмин менамоянд.

Ҳамзамон, бояд қайд кард, ки масъалаҳои мутобиқсозии моделҳо ба забонҳои камзахира ва интегратсияи самараноки онҳо ба муҳитҳои тақсимшуда ба таври кофӣ таҳқиқ нашудаанд.

Дар таҳқиқоти муҳаққиқони ватанӣ танҳо масъалаҳои тарҳрезии системаҳои тақсимшуда ва оптимизатсияи коркарди ҳаҷмҳои калони маълумоти матнӣ дар шароити маҳдудияти захираҳои ҳисоббарорӣ баррасӣ шудаанд [2-3; 5-6; 12; 1-А–12-А].

Асоси методологии таҳқиқоти мазкурро корҳои илмии олимони ватанӣ ва хориҷӣ дар соҳаи моделсозии математикӣ, системаҳои тақсимшуда ва омӯзиши мошинӣ ташкил додаанд: Ф.С. Комилиён, М.К. Юнуси, З.Ҷ. Усмонов, М.Р. Ёров, В.В. Меншуткин, M. Fowler, S. Newman, L. Bass, I. Weber, L. Zhu, B. Burns, J. Beda, K. Hightower, N. Dragoni, C. Pahl, Q. Zhang, M. Chen, J. Devlin, M.W. Chang, K. Lee, K. Toutanova, A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, A. Rogers, T. Wolf, T.

Brown, P. Jamshidi, D. Taibi, J. Thönes, N. Alshuqayran, H. Zhang, S. Giallorenzo, A.L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, L. Safina, J. Lewis ва дигарон.

Дар доираи таҳқиқоти диссертатсионӣ самтҳои зерин рушд дода шуданд: формализатсияи раванди коркарди санадҳо, таҳияи архитектураи микросервисӣ бо компоненти ҷудошудаи зеҳнӣ, мутобиқсозии моделҳои BERT барои забонҳои камзахира ва таҳияи алгоритми амалии татбиқи система, ки ба шароити фаъолияти сохторҳои давлатӣ мутобиқ карда шудааст. Дар таҳқиқот диққати махсус ба масъалаҳои қиёспазирӣ, устуворӣ ба нокомиҳо ва истифодаи самараноки захираҳои ҳисоббарорӣ дода шудааст.

Бо вуҷуди натиҷаҳои бадастомада, масъалаҳои баланд бардоштани самаранокӣ ҳангоми коркарди ҳаҷмҳои калони маълумот, тавсеаи дастгирии бисёрзабонӣ ва интегратсия бо системаҳои давлатии гардиши санадҳои электронӣ ҳанӯз ҳам кушода боқӣ мемонанд.

Ҳамин тариқ, таҳқиқоти гузаронидашуда мавқеи миёнаро байни таҳияҳои бунёдии моделҳои трансформерӣ ва татбиқи амалии онҳо дар системаҳои зеҳнии гардиши санадҳо ишғол намуда, як қатор ҷанбаҳои қаблан нокифоя таҳқиқшударо фаро мегирад, ки ба мутобиқсозии моделҳои BERT ба забонҳои камзахира ва қиёспазирии комбинатсионии архитектуриҳои микросервисӣ алоқаманд мебошанд.

Робитаи таҳқиқот бо барномаҳо ва мавзуъҳои илмӣ. Таҳқиқоти диссертатсионӣ мутобиқи самтҳои афзалиятноки рушди илмӣ-техникӣ ва рақамии Ҷумҳурии Тоҷикистон иҷро гардида, ба рушди усулҳои моделсозии математикӣ, системаҳои иттилоотии зеҳнӣ ва технологияҳои коркарди маълумот равона шудааст.

Мавзуи диссертатсионӣ бевосита бо татбиқи стратегияҳо ва барномаҳои давлатии соҳаи рақамисозӣ алоқаманд мебошад, аз ҷумла Стратегияи миллии рушди Ҷумҳурии Тоҷикистон то соли 2030 [10], Консепсияи ташаккули ҳукумати электронӣ дар Ҷумҳурии Тоҷикистон [9] ва Барномаи иқтисоди рақамии Ҷумҳурии Тоҷикистон [11], ки дар доираи онҳо ҷорӣ намудани технологияҳои муосири иттилоотӣ, системаҳои гардиши электронии санадҳо ва усулҳои зеҳни сунъӣ дар равандҳои идоракунии давлатӣ пешбинӣ шудааст.

Диссертатсия дар доираи самти илмии кафедраи информатикаи Донишгоҳи миллии Тоҷикистон «Таҳия ва таҳқиқи моделҳои математикӣ-компютерии экосистемаҳои обанборҳо, мамнуъгоҳҳо ва низоми маориф», ки дар Маркази миллии иттилоотии патентии Ҷумҳурии Тоҷикистон (№ 0116TJ00657) ба қайд гирифта шудааст, иҷро гардида, онро тавассути истифодаи усулҳои моделсозии математикӣ ва зеҳни сунъӣ дар ҳалли масъалаҳои автоматизатсияи гардиши санадҳо густариш медиҳад.

Натиҷаҳои таҳқиқот ба таҳия ва ҷорӣ намудани системаҳои зеҳнии гардиши санадҳо, ки бар асоси архитектураи микросервисӣ ва моделҳои трансформерии навъи BERT сохта шудаанд, равона гардида, ба тамоюлҳои муосири трансформатсияи рақамӣ ва баланд бардоштани самаранокии идоракунии захираҳои иттилоотӣ мутобиқ мебошанд.

ТАВСИФИ УМУМИИ ТАҲҚИҚОТ

Таҳқиқоти диссертатсионӣ ба таҳия ва асосноккунии математикии системаи зеҳнии гардиши санадҳо бахшида шуда, дар асоси архитектураи микросервисӣ ва моделҳои трансформерии навъи BERT амалӣ гардидааст ва ба баланд бардоштани самаранокии коркарди санадҳо, хусусиятҳои қиёспазирӣ ва мутобиқшавӣ ба шароити тағйирёбанда равона гардидааст.

Мақсади таҳқиқот аз таҳия ва асосноккунии назариявию амалии модели математикии системаи зеҳнии гардиши санадҳо иборат аст, ки дар асоси архитектураи микросервисӣ ва моделҳои BERT кор мекунад ва ба баланд бардоштани дақиқии коркарди санадҳо, қиёспазирӣ ва устувории фаъолияти онҳо бахшида шудааст.

Вазифаҳои таҳқиқот дар асоси ҳадафи таҳқиқот муайян гардида, ҳалли масъалаҳои зеринро дар бар мегиранд:

1. Гузаронидани таҳлили равишҳои муосири сохтани системаҳои иттилоотии тақсимшуда, аз ҷумла таҳқиқи хусусиятҳои архитектураи микросервисӣ, бартариҳо, маҳдудиятҳо ва самтҳои татбиқи амалии он.
2. Таҳқиқи усулҳои мавҷудаи қиёспазирии барномаҳои микросервисӣ ва механизмҳои таъмини устуворӣ ба нокомиҳо дар муҳитҳои барномавии тақсимшуда бо шиддати баланди коркарди дархостҳо.
3. Таҳияи модели архитектурии системаи зеҳнии гардиши санадҳо бо дарназардошти талабот нисбат ба модулнокӣ, васеъшавандагӣ, устуворӣ ва ҳамкории мутақобилаи сервисҳои алоҳида.
4. Амалӣ намудани мутобиқсозии моделҳои трансформерии навъи BERT барои ҳалли масъалаҳои истихроҷи моҳиятҳо, таҳлили семантикӣ ва таснифоти санадҳо дар шароити маҳдудияти захираҳои забонӣ.
5. Таҳияи усулҳои ташаккул ва тавлиди маълумоти омӯзишӣ, ки ба баланд бардоштани самаранокии омӯзиши моделҳои нейрошабакавии коркарди иттилооти матнӣ нигаронида шудаанд.
6. Амалисозии прототипи барномавии системаи зеҳнии гардиши санадҳо бо арзёбии минбаъдаи таҷрибавии самаранокии моделҳо, алгоритмҳо ва роҳҳои ҳалли архитектурии пешниҳодшуда дар асоси маҷмуи нишондиҳандаҳои миқдорӣ

Объекти таҳқиқот системаҳои зеҳнии гардиши санадҳо мебошанд, ки дар шароити трансформатсияи рақамӣ фаъолият намуда, дар асоси архитектуриҳои тақсимшудаи микросервисӣ сохта шудаанд.

Мавзуи таҳқиқот усулҳо ва моделҳои таъминоти математикӣ, алгоритмҳо ва роҳҳои ҳалли архитектурии сохт ва қиёспазирии системаи зеҳнии гардиши санадҳоро бо истифода аз моделҳои трансформерии навъи BERT дар муҳити микросервисӣ дар бар мегирад.

Асосҳои назариявии таҳқиқотро консепсияҳои амсиласозии математикию компютерии системаҳои иттилоотӣ, архитектураи микросерсисӣ ва диди зеҳнӣ ба равандҳои автоматикунонии гардиши санадҳо ташкил медиҳанд. Робитаи мавзуъ бо назарияҳои мавҷуда дар таҳлили интиқодии моделҳои трансформерии навъи BERT ва ҳалли мушкилоти мутобиқсозии моделҳои бисёрзабона дар шароити маҳдудияти захиравии забонҳои камманбаъ зоҳир мегардад. Таҳияи комплекси амсилаҳои математикӣ, усулҳои гибридии коркарди матн ва роҳҳои ҳалли архитектурӣ заминаи назариявиро барои гузаронидани таҷрибаҳои компютерӣ ва идоракунии илмии гардиши санадҳо фароҳам овардааст.

Навгонии илмӣ дар таҳияи асосҳои назариявӣ ва амалии сохтани системаҳои зеҳнии гардиши санадҳо дар заминаи архитектураи микросервисӣ ва моделҳои трансформерии коркарди забони табиӣ ифода ёфта, бо дастёбӣ ба натиҷаҳои зерин зоҳир гардидааст:

1. Модели математикии системаи зеҳнии гардиши санадҳо таҳия шудааст, ки принсипҳои архитектураи микросервисӣ ва усулҳои трансформерии коркарди матниро ҳамгиро намуда, формализатсияи равандҳои истихроҷ, сохторбандӣ ва тафсири маълумотро дар муҳити ҳисоббарории тақсимшуда таъмин менамояд [6-А; 7-А].

2. Роҳҳои ҳалли архитектурии системаи микросервисӣ бо ҷудо намудани компоненти махсуси таҳлили зеҳнӣ (AI Service) пешниҳод гардидааст, ки бо имконияти қиёспазирии мустақили модулҳои ҳисоббарории мураккаб фарқ намуда, самаранокии коркарди маълумот ва чандирии рушди системаро таъмин менамояд [3-А].

3. Усули мутобиқсозии моделҳои бисёрзабонаи BERT ба шароити забонҳои камманбаъ таҳия шудааст, ки бар асоси тавлиди синтетикии намунаҳои омӯзишӣ ва бозомӯзӣ асос ёфта, имкон медиҳад, ки дақиқии баланди истихроҷи моҳиятҳои номдор дар шароити ҳаҷми маҳдуди маълумоти нишонагузоришуда ба даст оварда шавад [5-А].

4. Усули гибридии коркарди санадҳо пешниҳод гардидааст, ки моделҳои трансформерии омӯзиши мошинӣ ва алгоритмҳои қоидаасосро муттаҳид сохта, эътимоднокӣ, тафсирпазирӣ ва устувории фаъолияти системаро дар шароити нокифоягӣ ё дастнорасии маълумот баланд мебардорад [3-А].

5. Усулҳои қиёспазирии системаҳои микросервисӣ, аз ҷумла қиёспазириҳои уфуқӣ ва амудӣ ба низом оварда шуда, аз ҷиҳати назариявӣ асоснок карда шудаанд ва самаранокии истифодаи якҷояи онҳо барои таъмини маҳсулнокии баланд ва устуворӣ ба нокомиҳо дар замимаҳои зеҳнӣ исбот гардидааст [8-А; 10-А].

6. Нақши контейнеризатсия ва оркестратсия ҳамчун асоси технологии идоракунии системаҳои тақсимшуда асоснок карда шудааст, ки қиёспазирии автоматӣ, мувозинати сарборӣ, устуворӣ ба нокомӣ ва оптимизатсияи истифодаи захираҳои ҳисоббарориро таъмин менамояд [11-А].

7. Дар асоси модели математикии таҳияшудаи системаи зеҳнии гардиши санадҳо, ки бар архитектураи микросервисӣ бо истифодаи моделҳои BERT такя менамояд ва ба сохтори Кумитаи таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон мутобиқ гардонида шудааст, алгоритми амалии татбиқ таҳия гардидааст, ки бо услуби муҳандисӣ-амалӣ пешниҳод шуда, ҷорӣ намудани системаро дар шароити инфрасохтори воқеии иттилоотӣ таъмин менамояд [1-А–5-А].

8. Тавсиби таҷрибавии роҳи ҳалҳои пешниҳодшуда гузаронида шудааст. Дар натиҷа, баланд гаштани дақиқии коркарди санадҳо ва коҳишёбии назарраси вақти коркарди онҳо нисбат ба равишҳои маъмул тасдиқ гардидааст, ки аҳаммияти амалӣ ва самаранокии усулу моделҳои таҳияшударо нишон медиҳад [1-А].

Натиҷаҳои пешниҳодшуда, дар маҷмуъ, асоси аз ҷиҳати илмӣ асоснокшударо барои тарҳрезӣ ва ҷорӣ намудани системаҳои зеҳнии нави гардиши санадҳо дар шароити иқтисоди рақамӣ ташкил медиҳанд.

Нуктаҳои ба ҳимоя пешниҳодшаванда:

1. Модели математикӣ ва роҳи ҳалли архитектурии системаи зеҳнии гардиши санадҳо таҳия ва асоснок карда шудааст, ки дар асоси архитектураи микросервисӣ бо истифодаи моделҳои трансформерии навъи BERT бунёд ёфта, формализатсияи равандҳои истихроҷ, сохторбандӣ ва коркарди маълумотро дар шароити инфрасохтори воқеии иттилоотӣ таъмин менамояд. Ҳамчунин бо ҷудо намудани компоненти махсуси таҳлили зеҳнӣ ва истифодаи қиёспазирии уфуқию амудии микросервисҳо он маҳсулнокии баланд ва устувории системаро ба сарборӣ тағйирёбанда кафолат медиҳад.

2. Собит карда шудааст, ки усули мутобиқсозии моделҳои бисёрзабонаи трансформерии BERT дар асоси тавлиди синтетикии маълумоти омӯзишӣ ва бозомӯзӣ барои шароити забонҳои камманбаъ самаранок мебошад. Дар ин замина усули гибридии коркарди санадҳо, ки алгоритмҳои омӯзиши мошинӣ ва қоидаасосро муттаҳид месозад, имкон медиҳад, ки сифат ва дақиқии коркарди матнҳо дар шароити маҳдудияти захиравии забонӣ ба таври назаррас баланд бардошта шавад.

3. Муқаррар карда шудааст, ки татбиқи алгоритми амалии системаи зеҳнии гардиши санадҳо дар сохтори Кумитаи таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон коҳиши вақти коркарди санадҳо ва боло рафтани дақиқии истихроҷи иттилоотро таъмин менамояд, ки ин аҳаммияти баланди амалии натиҷаҳои таҳқиқот ва устувории таҷрибавии онҳоро тасдиқ мекунад.

Аҳамияти назариявӣ ва амалии таҳқиқот. Аҳаммияти назариявии таҳқиқот дар рушди минбаъдаи асосҳои илмии моделсозии математикии системаҳои зеҳнии гардиши санадҳо барои муҳитҳои ҳисоббарории тақсимшуда ифода меёбад. Натиҷаҳои бадастомада тасаввуроти мавҷударо оид ба принсипҳои сохтани архитектураҳои микросервисӣ тавассути интегратсияи онҳо бо усулҳои муосири коркарди забони табиӣ дар асоси моделҳои трансформерӣ боз ҳам васеътар менамоянд.

Модели математикии таҳияшудаи системаи зеҳнии гардиши санадҳо равандҳои истихроҷ ва сохторбандии маълумотро формалӣ намуда, дар рушди усулҳои таҳлили

иттилооти матнӣ ва назарияи системаҳои иттилоотии тақсимшуда нақши муассир мебозад. Асосноккунии назариявии равиши гибридӣ, ки усулҳои омӯзиши мошинӣ ва алгоритмҳои қоидаасосро муттаҳид месозад, устуворӣ ва тафсирпазирии системаҳои зеҳниро болотар мебарад.

Натиҷаи муҳими назариявӣ муназзамӣ ва асоснокии усулҳои қиёспазирии уфуқӣ ва амудии системаҳои микросервисӣ ва инчунин исботи самаранокии истифодаи якҷояи онҳо дар масъалаҳои коркарди маълумот мебошад. Илова бар ин, усули мутобиқсозии моделҳои трансформерӣ ба шароити забонҳои камманбаъ ва камзахира таҳия гардидааст, ки имконоти истифодаи технологияҳои муосири NLP-ро дар муҳитҳои бисёрзабона васеъ мегардонад.

Дар маҷмуъ, натиҷаҳои бадастомада заминаи назариявиро барои таҳқиқоти минбаъда дар соҳаи системаҳои зеҳнии тақсимшуда, гардиши рақамии санадҳо ва истифодаи амалии усулҳои зеҳни сунъӣ ташаккул медиҳанд.

Аҳамияти амалии таҳқиқот бо имконияти истифодаи мустақими моделҳо, усулҳо ва роҳҳои ҳалли архитектурии таҳияшуда ҳангоми эҷод ва ҷорӣ намудани системаҳои зеҳнии гардиши санадҳо дар соҳаҳои гуногун муайян карда мешавад.

Дар асоси модели математикӣ ва равиши архитектурии пешниҳодшуда системаи зеҳнӣ таҳия гардидааст, ки автоматизатсияи равандҳои коркарди санадҳо, аз ҷумла истихроҷи маълумот, таснифот ва ташаккули иттилооти сохторбандишударо таъмин месозад. Ҷорӣ намудани система имкон медиҳад, ки вақти коркарди санадҳо ба таври назаррас коҳиш дода шавад, шумораи хатогиҳои воридкуниҳои дастӣ кам карда шавад ва самаранокии равандҳои идоракунӣ баланд бардошта шавад.

Арзиши махсуси амалӣ ба алгоритми муҳандисӣ-амалии татбиқи система тааллуқ дорад, ки ба сохтори Кумитаи таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон мутобиқ гардонида шуда, татбиқпазирии натиҷаҳои таҳқиқотро дар шароити воқеии инфрасохтори давлатии иттилоотӣ тасдиқ менамояд.

Самтҳои истифодаи роҳи ҳалҳои пешниҳодшуда:

– дар бахши давлатӣ, барои автоматизатсияи гардиши санадҳо ва пешниҳоди хизматрасониҳои электронӣ;

– дар ташкилотҳои тиҷоратӣ, барои коркарди санадҳои шартномавӣ ва молиявӣ;

– дар муассисаҳои молиявӣ, ҳангоми татбиқи расмиёти назорат ва таҳлили маълумот.

Илова бар ин, аҳамияти амалии таҳқиқот дар имконияти қиёспазирии система, интегратсияи он бо платформаҳои мавҷудаи иттилоотӣ ва мутобиқсозии он ба соҳаҳои гуногуни мавзуӣ зоҳир мегардад.

Дараҷаи эътимоднокии натиҷаҳои диссертатсия бо амалӣ гаштани таҳқиқи комплексии масъалаҳои илман асосноки банақшагирифташуда, истифодаи технологияҳои муосири амсиласозии математикию компютерӣ ва моделҳои трансформерии BERT асоснок карда мешавад. Эътимоднокии натиҷаҳо инчунин бо роҳандозии озмоишҳои компютерӣ дар асоси маълумоти воқеии инфрасохтори иттилоотӣ, тавлиди синтетикии маълумоти омӯзишӣ ва татбиқи таҷрибавии алгоритмҳои таҳияшуда дар сохтори Кумитаи таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон тасдиқ мегардад.

Мутобиқати диссертатсия бо шиносномаи ихтисоси илмӣ. Мазмун ва муҳтавои диссертатсия самтҳои таҳқиқоти илмии ихтисоси 6D060204 – Асосҳои назариявии информатикаро фаро гирифта, ба бандҳои зерини шиноснома мувофиқат менамояд:

Банди 4. Таҳқиқоти комплексии масъалаҳои илмӣ-техникӣ ва бунёдӣ-амалӣ бо истифода аз технологияҳои муосири моделсозии математикӣ ва таҷрибаи ҳисоббарорӣ.

Банди 11. Омӯзиши мошинӣ ва зеҳни сунъӣ: асосҳои назариявии системаҳои омӯзанда, методҳои ошкорсозии донишҳои нав (Data Mining), хулосабарории мантиқӣ.

Банди 15. Моделҳои равандҳо ва сохторҳои иттилоотӣ.

Банди 18. Воситаҳои нави бадастории дониш ва ташкили онтологияҳо, воситаҳои зеҳнигардонии равандҳои бизнес.

Банди 25. Системаҳои тақсимшудаи бисёркорбарӣ.

Банди 26. Амалисозии методҳо ва алгоритмҳои самараноки ададӣ дар шакли комплексҳои барномаҳои масъалавӣ барои гузаронидани таҷрибаи ҳисоббарорӣ.

Саҳми шахсии довталаби дараҷаи илмӣ. Довталаб дар таҳияи мавзуи диссертатсионӣ иштирок намуда, таҳлили адабиёти соҳавиро анҷом додааст. Мустақилона модели математикии системаи зеҳнии гардиши санадҳоро дар асоси архитектураи микросервисӣ ва BERT таҳия кардааст, алгоритми амалиро татбиқ намудааст, таҷрибаҳои компютерӣ гузаронидааст, натиҷаҳоро коркард ва тафсир кардааст, инчунин тавсияҳои методӣ ва матни диссертатсияро омода намудааст.

Тасвиб ва амалисозӣ. Натиҷаҳои таҳқиқоти диссертатсионӣ дар семинарҳои муштараки кафедраҳои «Информатика», «Математикаи ҳисоббарорӣ ва механика», «Моделсозии математикӣ ва компютерӣ» ва «Технологияҳои иттилоотӣ ва коммуникатсионӣ»-и Донишгоҳи миллии Тоҷикистон (ДМТ) баррасӣ ва муҳокима гардида, дар конфронсҳои ҷумҳуриявӣ ва байналмилалии солҳои 2022-2026 дар ДМТ, Академияи идоракунии давлатии назди Президенти Ҷумҳурии Тоҷикистон, Филиали Донишгоҳи давлатии Москва ба номи М.В. Ломоносов дар ш. Душанбе, Донишгоҳи давлатии Бохтар ба номи Носири Хусрав, Донишгоҳи байналмилалии сайёҳӣ ва соҳибкории Тоҷикистон, Донишгоҳи технологии Тоҷикистон ва дигар муассисаҳои соҳавӣ ба тасвиб расидаанд.

Интишорот аз рӯйи мавзуи диссертатсия. Натиҷаҳои асосии таҳқиқоти диссертатсионӣ дар 12 кори илмӣ [1-А–12-А] нашр шудаанд, аз ҷумла 7 мақола дар маҷаллаҳои тақризшавандаи Ҷумҳурии Тоҷикистон, 4 тезиси маърӯза дар маводи конфронсҳои байналмилалӣ ва ҷумҳуриявӣ ва як мақола дар маҷаллаи хориҷӣ.

Сохтор ва ҳаҷми диссертатсия. Диссертатсия аз бахшҳои «Муқаддима», «Тавсифи умумии таҳқиқот», чор боб, бахши «Хулосаҳо» бо зербахшҳои «Натиҷаҳои асосии илмии диссертатсия» ва «Тавсияҳо оид ба истифодаи амалии натиҷаҳо», бахши «Адабиёт» бо зербахшҳои «Феҳристи сарчашмаҳои истифодашуда» ва «Феҳристи интишороти илмии довталаби дарёфти дараҷаи илмӣ», инчунин бахши «Замимаҳо» иборат мебошад. Таҳқиқот дар ҳаҷми 197 саҳифаи матни компютерӣ дар виростори матнии MS Word иҷро шуда, дорои 9 ҷадвал ва 15 расм мебошад, ки рақамгузории онҳо вобаста ба бобҳои диссертатсия сурат гирифтааст. Рӯйхати адабиёт бо шумули интишороти довталаб 182 номгӯйро дар бар мегирад.

МУҲТАВОИ АСОСИИ ДИССЕРТАТСИЯ

Дар муқаддима мубрамияти мавзуъ асоснок гардида, объект ва мавзуи таҳқиқот муайян карда шудаанд. Ҳадаф ва вазифаҳо, асосҳои методологӣ ва дараҷаи таҳқиқи илмии масъала ошкор гардидаанд. Робитаи таҳқиқот бо барномаҳои давлатӣ нишон дода шудааст. Навоварии илмӣ, муқаррароте, ки ба ҳимоя пешниҳод мешаванд, аҳамияти назариявӣ ва амалии таҳқиқот муайян карда шудаанд. Маълумот оид ба мутобиқати таҳқиқот ба шиносномаи ихтисос, тасвиби натиҷаҳо дар нашрияҳо оварда шуда, инчунин сохтор ва ҳаҷми диссертатсия тавсиф гардидаанд.

Дар боби якум таҳлили адабиёти илмӣ ва равишҳои муосири технологӣ гузаронида шудааст. Таҳлил имкон дод, ки оид ба рушди архитектураи микросервисӣ ва системаҳои зеҳнии гардиши санадҳо дар шароити трансформасияи рақамӣ тасаввуроти системавӣ ташаккул дода шавад [16; 18; 20; 22]. Муайян карда шуд, ки архитектураи микросервисӣ марҳилаи қонунии эволютсияи системаҳои барномавӣ буда, гузаришро ба роҳи ҳалҳои тақсимшуда ва мутобиқшаванда таъмин менамояд (расми 1.1).

Исбот карда шудааст, ки бартариҳои архитектураи микросервисӣ дар модулнокӣ, худмухторӣ ва мустақилияти ҷойгиркунии сервисҳо зоҳир мегарданд. Ҳамзамон дар кори он маҳдудиятҳое муайян карда шудаанд, ки бо мураккабшавии идоракунӣ, зарурати ҷорӣ намудани амалияҳои DevOps ва таъмини мувофиқати маълумот вобаста мебошанд.

Асоснок карда шудааст, ки фаъолияти самараноки системаҳои микросервисӣ ҳангоми риояи принсипҳои алоқамандии суст, изолятсияи маълумот ва автоматизатсияи давраи ҳаётӣ ба даст меояд. Нишон дода шудааст, ки истифодаи контейнеризатсия, оркестратсия,

механизмҳои асинхронии ҳамкорӣ ва технологияҳои абрӣ сатҳи зарурии қиёспазирӣ ва устуворӣ ба нокомиҳоро таъмин менамоянд.

Муайян карда шудааст, ки баланд бардоштани самаранокии архитектураи мазкур зимни истифодаи комплексии қиёскуниҳои уфуқӣ, архитектураи рӯйдодгаро, кэширонӣ ва мувозинатсозии сарборӣ ба даст меояд. Барои ҷузъҳои ҳисоббарориҳои мураккаб, ба монанди батчиронҳо ва суръатбахшандаҳои сахтафзорӣ, зарурати истифодаи усулҳои иловагии оптимизатсионӣ пешниҳод шудааст.



Расми 1.1. Эволютсияи архитектурҳои системаҳои барномавӣ

Дар заминаи архитектураи микросервисӣ бо компоненти таҳлили зеҳнӣ системаи зеҳнии гардиши санадҳо таҳия гардидааст, ки он қиёспазирии мустақил ва чандирияти рушдро таъмин сохта метавонад.

Исбот карда шудааст, ки истифодаи модели BERT ба даст овардани дақиқии баланди истихроҷи моҳиятҳо (то 93% аз рӯи метрикаи F1)-ро имкон медиҳад. Дар асоси синтези маълумот ва омӯзиши иловагӣ имконияти мутобиқсозии моделҳои бисёрзабона ба шароити забони камзахираи тоҷикӣ асоснок карда шудааст.

Навоварии илмии боби мазкур дар интегратсияи архитектураи микросервисӣ ва усулҳои омӯзиши мошинӣ, таҳияи равиши гибридии коркарди маълумот ва мутобиқсозии моделҳои трансформерӣ ба забонҳои камзахира ифода меёбад. Аҳамияти амалии он бо коҳишёбии вақти коркарди санадҳо ва паст шудани хатогиҳо тасдиқ гардида, самаранокии истифодаи онро дар соҳаҳои гуногуни татбиқӣ таъмин месозад.

Маҳдудиятҳое, ки ба ҳаҷми захираҳои моделҳо ва коркарди маълумоти ғайри сохторӣ вобастаанд, муайян карда шудаанд. Самтҳои рушди минбаъдаи таҳқиқот, аз ҷумла оптимизатсияи моделҳо, рушди равишҳои мултимодалӣ ва такмили идоракунии давраи ҳаётии компонентҳои ML, муайян гардидаанд.

Ҳамин тариқ, исбот карда шудааст, ки интегратсияи архитектураи микросервисӣ ва усулҳои зеҳни сунъӣ самти асоснок ва ояндадори рушди системаҳои гардиши санадҳо буда, ба талаботи муосири иқтисоди рақамӣ мутобиқ мебошад.

Боби дуюми диссертатсия ба таҳқиқи роҳҳои ҳалли архитектурӣ, ки ба таъмини қиёсшавандагии системаҳои микросервисӣ равона шудаанд, бахшида шудааст. Таҳлили баргузоршуда имкон дод, ки тасаввуроти системавӣ оид ба равишҳои муосири баланд бардоштани маҳсулнокӣ, устуворӣ ба нокомиҳо ва мутобиқшавандагии системаҳои иттилоотии тақсимшуда, ки ба коркарди ҳаҷми зиёди маълумот ва истифодаи усулҳои зеҳни сунъӣ нигаронида шудаанд, ташаккул дода шавад [5; 7; 8; 11-А].

Муайян карда шудааст, ки таъмини қиёспазирӣ дар архитектураи микросервисӣ ба истифодаи комплексии қиёскуниҳои уфуқӣ ва амудӣ ҳамчун стратегияҳои якдигарро пуркунанда асос меёбад. Исбот карда шудааст, ки миқёсгирии уфуқӣ чандирият ва устуворӣ

ба нокомиҳоро тавассути такрорсозии сервисҳо ва тақсимоти сарборӣ таъмин менамояд, дар ҳоле ки қиёскунии амудӣ зимни ҳаллу фасли масъалаҳои ҳисоббарории захираталаб, ки хосси компонентҳои омӯзиши мошинӣ мебошанд, самаранок аст. Нишон дода шудааст, ки самаранокии баландтарин зимни истифодаи якҷояи онҳо бо назардошти ғайри якҷинсагии сарборӣ ва хусусиятҳои функсионалии сервисҳо ба даст меояд (ҷадвали 2.1).

Асоснок карда шудааст, ки асоси технологии системаҳои микросервисии қиёсшаванда контейнеризатсия мебошад, ки изолятсияи компонентҳо, бартарафсозии ихтилофи вобастагиҳо ва интиқолпазирии барномаҳоро таъмин менамояд. Муайян карда шудааст, ки идоракунии инфрасохтори контейнерӣ тавассути воситаҳои оркестратсия амалӣ мегардад, ки автоматизатсияи ҷойгиркунӣ, мувозинатсозии сарборӣ ва назорати ҳолати сервисҳоро таъмин мекунанд.

Ҷадвали 2.1. Таҳлили муқоисавии равишҳо

Меъёр	**Қиёскунии уфуқӣ**	**Қиёскунии амудӣ**
Принсип	Афзоиши шумораи гиреҳҳо	Афзоиши захираҳои гиреҳ
Устуворӣ ба нокомиҳо	Баланд	Паст
Қиёспазирӣ	Амалан номаҳдуд	Маҳдуд
Дараҷаи мураккабии татбиқ	Баланд	Паст
Арзиш	Оптимизатсияшаванда	Зимни афзоиш баланд

Исбот шудааст, ки татбиқи механизмҳои қиёскунии автоматӣ (HPA, VPA, Cluster Autoscaling) имкон медиҳанд, ки система ба таври динамикӣ ба тағйирёбии сарборӣ мутобиқ шуда, истифодаи захираҳоро оптимизатсия намояд. Нишон дода шудааст, ки самаранокии ин механизмҳо аз танзими дурусти ченакҳо ва тарзҳои идоракунӣ вобаста аст.

Муайян карда шудааст, ки дар системаҳои зеҳнии микросервисӣ масъалаҳои қиёскунӣ хусусияти махсус доранд: сервисҳои коркарди дархостҳо асосан ба таври уфуқӣ қиёс мешаванд, дар ҳоле ки ҷузъҳои ҳисоббарории мураккаб ба қиёскунии амудӣ бо истифодаи захираҳои махсус ниёз доранд. Мақсаднок будани истифодаи роҳи ҳалли гибридии архитектурӣ, ки ҳамкориҳои асинхронӣ, кэширонӣ ва коркарди пакетии маълумотро дар бар мегиранд, асоснок карда шудааст.

Нақши калидии оркестратсия ҳамчун унсури идоракунии системаҳои тақсимшуда муайян карда шудааст, ки худбарқароршавӣ, навсозии бефосила ва татбиқи равиши декларативиро дар идоракунии инфрасохтор таъмин намуда, пешгӯишавандагӣ ва идорашавандагии системаро баланд мебардорад.

Дар баробари бартариҳо, маҳдудиятҳое низ ошкор карда шуданд, ки ба мураккабии идоракунии инфрасохтор, зарурати ҷорисозии амалияҳои DevOps, талаботи баланд зоҳир намудан ба мониторинг ва амният, инчунин ба душворқиёспазирии компонентҳои ҳолатгаро (stateful) вобаста мебошанд.

Аҳамияти амалии натиҷаҳо дар имконияти истифодаи онҳо ҳангоми лоиҳакашии системаҳои зеҳнии қиёсшаванда, аз ҷумла системаҳои гардиши санадҳо, бо таъмини маҳсулнокии баланд, устуворӣ ба нокомиҳо ва истифодаи самараноки захираҳо ифода меёбад.

Ҳамин тариқ, исбот карда шудааст, ки интегратсияи усулҳои қиёскуниҳои уфуқӣ ва амудӣ, контейнеризатсия ва оркестратсияи автоматӣ асоси технологии сохтани системаҳои муосири микросервисиро ташкил медиҳад. Истифодаи онҳо шарти зарурии таҳияи роҳи ҳалли мутобиқшаванда, боэътимод ва қиёсшаванда буда, ба талаботи иқтисоди рақамӣ ҷавобгӯ мебошад [2-А; 9-А].

Боби сеюми диссертатсия ба таҳлили муқоисавии архитектуриҳои монолитӣ ва микросервисӣ дар заминаи трансформасияи системаҳои иттилоотӣ ва оптимизатсияи захираҳои ТИ бахшида шудааст. Таҳқиқот имкон дод хулосаҳои умумӣ оид ба бартариҳо ва маҳдудиятҳои равишҳои баррасишуда шакл гиранд.

Дар шароити рушди босуръати технологияҳои иттилоотӣ талабот ба чандирӣ, қиёспазирӣ, устуворӣ ба нокомиҳо ва суръати навсозии таъминоти барномавӣ ба таври назаррас меафзояд. Системаҳои маъмули монолитӣ, ки тамоми функсионалро дар як барнома муттаҳид мекунанд, дар дурнамо маҳдудиятҳои худро ошкор месозанд, ки бо мушкилоти миқёсгирӣ, воридсозии тағйирот ва интегратсияи технологияҳои нав вобаста мебошанд [1; 4; 5].

Архитектураи микросервисӣ ҳамчун алтернативаи самаранок ба равиши монолитӣ рӯйи кор омада, декомпозитсияи системаро ба сервисҳои мустақил таъмин менамояд, ки ҳар кадоми онҳо як вазифаи алоҳидаи тиҷориро иҷро мекунад. Ин ба баланд гаштани чандирии таҳаввулотҳо, содагардонии навсозиҳо, беҳтар шудани устуворӣ ба нокомиҳо ва имконияти истифодаи технологияҳои гуногун дар доираи як система мусоидат мекунад.

Пояи асосии технологии гузариш ба микросервисҳо воситаҳои муосири контейнеризатсия ва оркестратсия (масалан, Docker ва Kubernetes) мебошанд, ки стандартсозии ҷойгиркунӣ, миқёсгирӣ ва идоракунии сервисҳои тақсимшударо таъмин менамоянд. Ҳамзамон нақши амалияҳои DevOps, аз ҷумла Continuous Integration ва Continuous Deployment, меафзояд. Онҳо имкон медиҳанд, ки равандҳои таҳия автоматӣ гардонида шуда, вақти ба бозор баровардани маҳсулот ба таври назаррас коҳиш дода шавад.

Бо вуҷуди ин, татбиқи муваффақонаи архитектураи микросервисӣ на танҳо бо омилҳои технологӣ, балки бо зарурати тағйироти амиқи ташкилӣ ва фарҳангӣ муайян мегардад. Гузариш ба моделҳои чандиртари идоракунии таҳия, тақвияти ҳамкории байни гурӯҳҳои таҳия ва истифода, инчунин ташаккули муҳити ҳамкорӣ ва мубодилаи донишҳо талаб карда мешавад.

Омезиши комплексии тағйироти техникӣ, ташкилӣ ва стратегӣ ба ташкилоту муассисаҳо имкон медиҳад, ки самаранокии истифодаи захираҳои ТИ баланд бардошта шавад, фаъолияти онҳо ба талаботи босуръат тағйирёбандаи бозор мутобиқ сохта шавад, равандҳои инноватсионӣ тезонида шаванд, сифати хизматрасониҳо беҳтар карда шаванд.

Ҳамин тариқ, гузариш аз архитектурии монолитӣ ба микросервисӣ омили муҳими баланд бардоштани рақобатпазирии системаҳои иттилоотии муосир ва мутобиқати онҳо ба талаботи иқтисоди рақамӣ ба ҳисоб меравад [6-А; 8-А].

Боби чоруми диссертатсия ба моделсозии математикӣ ва архитектурии системаи зеҳнии микросервисии гардиши санадҳо бо истифода аз моделҳои трансформерии BERT бахшида шудааст. Дар заминаи трансформатсияи рақамии сохторҳои давлатӣ ва корпоративии Ҷумҳурии Тоҷикистон натиҷаҳои бадастомада ҳам аҳамияти назариявӣ ва ҳам арзиши амалӣ пайдо мекунанд.

Дар таҳқиқот синтези тарроҳии архитектурии системаҳои тақсимшуда ва технологияҳои NLP (Natural Language Processing – коркард ва таҳлили забони табиӣ) истифода шудааст. Абзори мазкур барои таъмини самаранокии баландтари система интихоб шудааст: қобилияти тобоварӣ ба сарбориҳо, минимизатсияи нокомиҳо ва расидан ба эътимоднокии максималӣ зимни таҳлили матнҳо. Технологияи NLP барои коркарди автоматии матнҳо усулҳои лингвистӣ ва омӯзиши мошиниро муттаҳид месозад: тарҷума, таҳлили маъно ва лаҳҷа, сохтани чат-ботҳо ва ёрдамчиёни зеҳнӣ [14; 21].

Формализатсияи масъалаи шинохти моҳиятҳо (NER). Мавриди баррасӣ пайдарпайии токенҳо $\mathcal{X} = (x_1, x_2, \cdots, x_n)$ қарор дода мешавад. Талаб карда мешавад, ки пайдарпайии нишонаҳо $\mathcal{Y} = (y_1, y_2, \cdots, y_n), y_j \in \mathcal{C}$ муайян карда шавад: $\mathcal{C}$ – маҷмуи синфҳои нишонагузории навъи BIO (B – оғоз, I – дохилшавӣ ба объект, O – беруноӣ аз объект).

Масъала ҳамчун масъалаи максимизатсияи шартӣ шакл дода мешавад: $\hat{\mathcal{Y}} = \arg \underbrace{\max}_{\mathcal{Y}} \mathcal{P}(\mathcal{Y} \mid \mathcal{X};\ \theta)$: θ – параметрҳои модели BERT. Ин гуна тасвир бо модели эҳтимолии нишонгузории пайдарпай мувофиқ аст.

Механизми «худэътибор» (self-attention) дар трансформер: Барои матритсаи эмбеддингҳо $\mathcal{X} \in \mathbb{R}^{n \times d}$ матритсаҳои $\mathcal{Q} = \mathcal{X}\mathcal{W}^{\mathcal{Q}}, \mathcal{K} = \mathcal{X}\mathcal{W}^{\mathcal{K}}, \mathcal{V} = \mathcal{X}\mathcal{W}^{\mathcal{V}}$ ҳисоб карда мешаванд: $\mathcal{Q}$ – матрисаи суроғаҳо, $\mathcal{K}$ – матрисаи калидҳо, $\mathcal{V}$ – матрисаи қиматҳо.

Механизми таваҷҷуҳ (attention) ба таври зерин ҳисоб карда мешавад: $\mathcal{A}(Q,\mathcal{K},\mathcal{V}) = \text{softmax}\left(\frac{Q\mathcal{K}^T}{\sqrt{d_k}}\right)\mathcal{V}$, ки дар ин ҷо d_k – андозаи фазои калидҳо мебошад [5-А].

Таваҷҷуҳи бисёрҷониба (multi-head attention): $\mathcal{M}(Q,\mathcal{K},\mathcal{V}) = \text{Concat}(\mathcal{H}_1,\mathcal{H}_2,\cdots,\mathcal{H}_h)\mathcal{W}^o$, ки $\mathcal{H}_i = \mathcal{A}\left(Q\mathcal{W}_i^Q, \mathcal{K}\mathcal{W}_i^\mathcal{K}, \mathcal{V}\mathcal{W}_i^\mathcal{V}\right)$ аст. Ин механизм имкон медиҳад, ки вобастагиҳои глобалӣ дар матн моделсозӣ шаванд [5-А].

Баромадгоҳи таснифотии токенҳо. Иҷроишҳои ниҳонии модели BERT: $\mathcal{H} = (h_1, h_2, \cdots, h_n)$. Логитҳои таснифӣ: $z_j = \mathcal{W}h_j + b$. Модели эҳтимолии таснифот: $\mathcal{P}\left(y_j = c \mid h_j\right) = \frac{exp(z_{j,c})}{\sum_{c'} exp\left(z_{j,c'}\right)}$.

Функсияи талафот. Бо назардошти хориҷ шудани зертокенҳо аз ҳисобкунӣ, кросс-энтропияи $\mathcal{L} = -\sum_{j\in\mathcal{J}}\sum_{c\in\mathcal{C}} y_{j,c}\log\mathcal{P}(y_j, c)$ истифода мешавад, ки $\mathcal{J}$ маҷмуи шохисҳои зертокенҳои аввалин, $y_{j,c}$ нишони ҳақиқӣ (one-hot) мебошанд. Оптимизатсия:

$$\theta^* = \arg\underbrace{\min}_{\theta}\mathcal{L}.$$

Ченакҳои арзёбии сифат: дақиқӣ (precision): $\mathcal{P}_r = \frac{\mathcal{T}_\mathcal{P}}{\mathcal{T}_\mathcal{P}+\mathcal{F}_\mathcal{P}}$; пуррагии система (recall): $\mathcal{R}_c = \frac{\mathcal{T}_\mathcal{P}}{\mathcal{T}_\mathcal{P}+\mathcal{F}_\mathcal{N}}$; $\mathcal{F}_1$-баллҳо (F-measure): $\mathcal{F}_1 = \frac{2\mathcal{P}_r\mathcal{R}_c}{\mathcal{P}_r+\mathcal{R}_c}$, $\mathcal{T}_\mathcal{P}$ – натиҷаҳои мусбати воқеӣ (true positives), $\mathcal{F}_\mathcal{P}$ – натиҷаҳои мусбати бардурӯғ (false positives), $\mathcal{F}_\mathcal{N}$ – натиҷаҳои манфии бардурӯғ (false negatives) [4-А].

Арзёбии афзоиши сифат. Афзоиш $\mathcal{F}_1$ нисбат ба методи пойгоҳӣ:

$$\Delta\mathcal{F}_1 = \frac{\mathcal{F}_1^{BERT} - \mathcal{F}_1^{Regex}}{\mathcal{F}_1^{Regex}}\cdot 100\%.$$

Самаранокии система. Вақти миёнаи коркарди санад: $\mathcal{T}_{avg} = \frac{1}{\mathcal{N}}\sum_{j=1}^{\mathcal{N}} t_j$.

Иқтидори гузаронандагии система: $\Theta = \frac{\mathcal{N}}{\mathcal{T}_{total}}$.

Миқёсгирии уфуқӣ. Зимни миқдори нусхаҳои сервис ба k баробар будан $\Theta(k) \approx k\cdot\Theta(1)$. Коэффитсиенти қиёс: $\mathcal{S}(k) = \frac{\mathcal{T}(1)}{\mathcal{T}(k)}$.

Механизми гибридии формализатсия. Функсияи ниҳоии истихроҷ:

$$\mathcal{F}_{extract}(\mathcal{X}) = \begin{cases}\mathcal{F}_{BERT}(\mathcal{X}), \text{агар сервис дастрас бошад}\\ \mathcal{F}_{Regex}(\mathcal{X}), \text{агар сервис дастрас набошад.}\end{cases}$$

Самаранокии маълумоти синтезӣ. Сифати нисбӣ:

$$\mathcal{Q}_{rel} = \frac{\mathcal{F}_1^{synthetic}}{\mathcal{F}_1^{real}}.$$

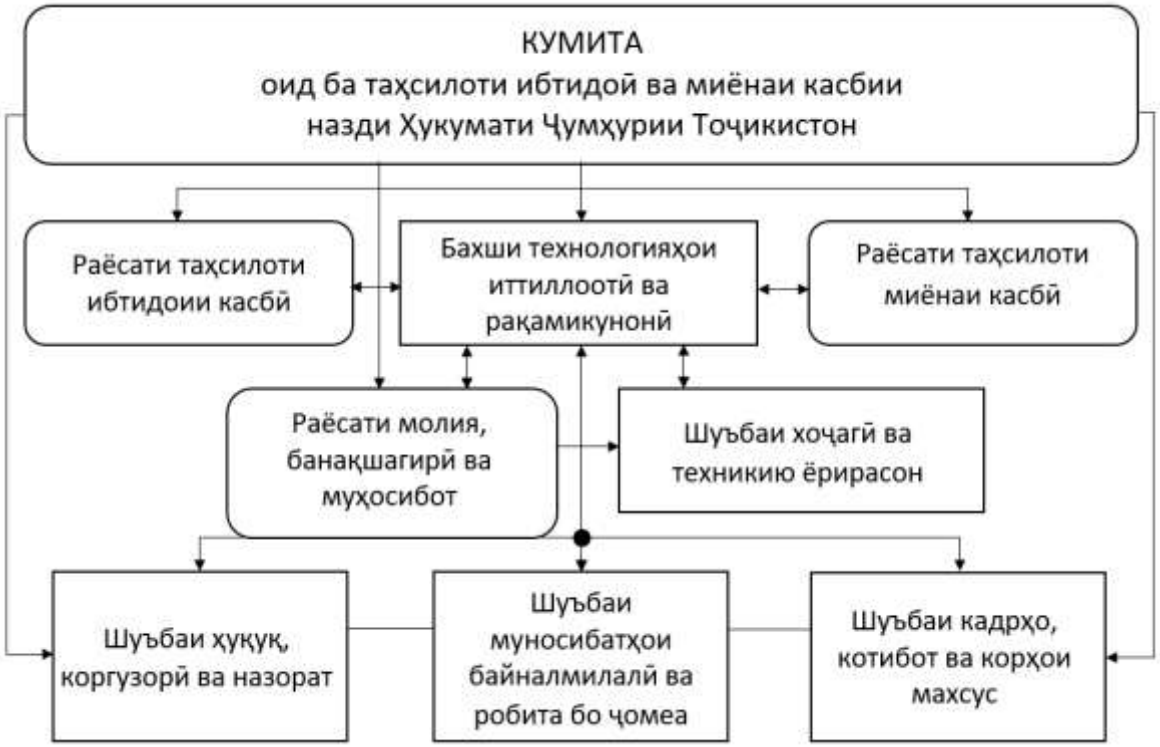
Системаи таҳияшуда равандҳои коркарди санадҳоро, ки дар шакли формалӣ пешниҳод шуда, тавассути графи ба хизматрасониҳои мутақобила равонашуда тавсиф мешаванд, дар намуди архитектураи микросервисии тақсимшуда таъмин менамояд. Дар доираи модел гузориши эҳтимолии масъалаи истихроҷи моҳиятҳои номдор амалӣ гардида, механизми худэътибори трансформерҳо формализатсия шудааст ва усулҳои арзёбии миқдории сифат ва самаранокии система кор карда баромада шудаанд.

Интегратсияи моделҳои бисёрзабонаи BERT бо механизмҳои тавлиди маълумоти синтезӣ имкон дод, ки сатҳи баланди дақиқии коркарди санадҳо (F1-score ≈ 0,93) ба даст оварда шавад. Ҳамгироии мазкур аз усулҳои маъмули қоидаасос (rule-based) ба таври назаррас бартарӣ дорад. Ҳамзамон, ҷорӣ намудани механизми гибридии коркард, ки усулҳои омӯзиши мошинӣ ва алгоритмҳои ба қоидаҳо асосёфтаро муттаҳид месозад, устувории иловагии системаро ба нокомиҳо таъмин менамояд. Механизми гузариши мутобиқшаванда (fallback) пайвастагии фаъолияти системаро ҳатто дар ҳолати муваққатан дастнорас будани хизматрасониҳои нейронӣ кафолат медиҳад, ки барои системаҳои иттилоотии давлатӣ аҳамияти калидӣ дорад.

Модел ба сохтори Дастгоҳи марказии Кумитаи таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон (Кумита) мутобиқ гардонида шудааст, ки он ҳашт воҳиди функционалиро дар бар мегирад (расми 4.1).

Татбиқи амалии системаи микросервисии гардиши санадҳои ба моделҳои BERT асосноккардашуда, ки ба сохтори Кумита мутобиқ гардонида шудааст, ҳалли комплексии масъалаҳои зеринро ташкил медиҳад (расми 4.2):

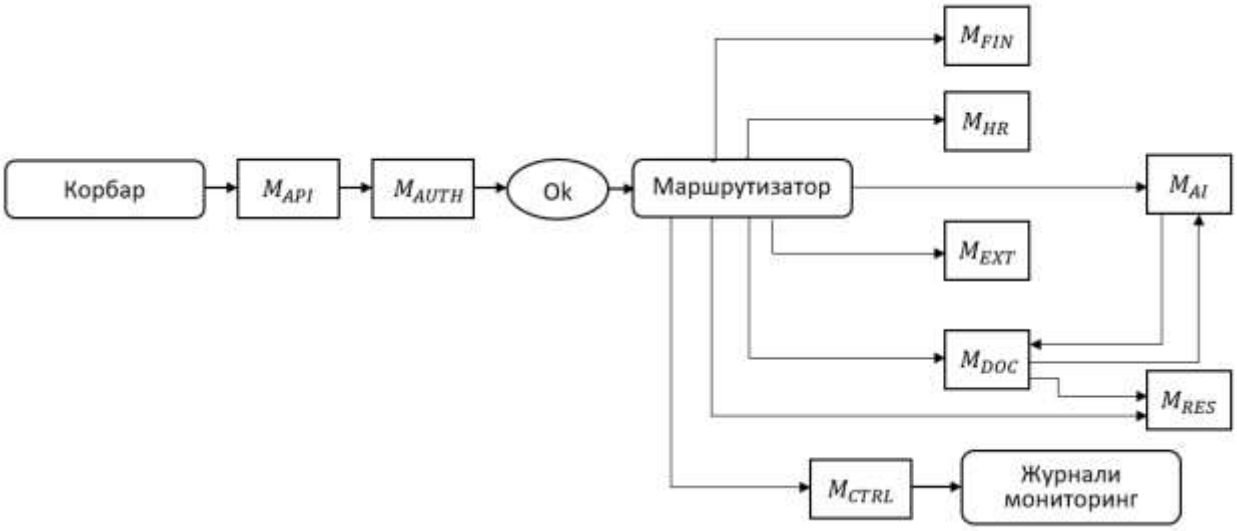
- *модели архитектурӣ* – тавсифи компонентҳои функционалӣ ва инфрасохторӣ;
- *формализатсияи математикӣ* – тавсифи механизмҳои таваҷҷуҳ (attention), ченакҳои самаранокӣ ва динамикаи селҳои иттилоотӣ;
- *мантиқи татбиқ* – интегратсияи модулҳо дар равандҳои бизнесии ҳашт зерсохтор бо истифодаи модели гибридии устуворӣ ба нокомиҳои система.



Расми 4.1. Сохтори Дастгоҳи марказии Кумита

Дар ин маврид Дастгоҳи марказии Кумита ҳамчун системаи тақсимшуда тасвир карда мешавад: $\mathcal{S} = \langle \mathcal{U}, \mathcal{D}, \mathcal{M}, \mathcal{A}, \mathcal{C} \rangle$: $\mathcal{U}$ – маҷмуи зерсохторҳо, $\mathcal{D}$ – маҷмуи санадҳо, $\mathcal{M}$ – маҷмуи микросервисҳо, $\mathcal{A}$ – модул зеҳнӣ (BERT), $\mathcal{C}$ – муҳити коммуникатсионӣ (gRPC/API). Зерсохторҳо: $\mathcal{U} = \{U_1, U_2, U_3, U_4, U_5, U_6, U_7, U_8\}$, U_1 – «Таҳсилоти ибтидоии касбӣ», U_2 – «Таҳсилоти миёнаи касбӣ», U_3 – «Молия ва муҳосибот», U_4 – «Кадрҳо ва котибот», U_5 – «Муносибатҳои байналмилалӣ», U_6 – «Ҳуқуқ ва назорат», U_7 – «ТИ ва рақамикунонӣ», U_8 – «Шуъбаи хоҷагидорӣ».

Интиқоли санадҳо (қисми асосии модел): Ҳар як санади Кумита чунин интиқол дода мешавад: $d_i = (x_i, y_i, m_i, t_i)$, ки x_i – сандаи матнӣ, y_i – моҳияти дастрасшуда, m_i – метамаълумот, t_i – вақти коркард мебошанд.



Расми 4.2. Алгоритми сикли пурраи татбиқи модели микросервисии системаи зеҳнии гардиши санадҳои Кумита оид ба таҳсилоти касбӣ

Функсияи коркарди санад: $y_i = f_{AI}(x_i)$: f_{AI} – модели BERT, $\hat{\mathcal{Y}} = \arg\underbrace{\max}_{\mathcal{Y}} \mathcal{P}(\mathcal{Y} \mid \mathcal{X};\, \theta)$, θ – параметрҳои модели BERT. Ин гуна тасвир ба модели эҳтимолии нишонагузории пайдарпай мувофиқат мекунад.

Архитектура: Система мутобиқан ба зерсохторҳои Кумита ба маҷмуи микросервисҳо тақсим карда мешавад (ҷадвали 4.1): $\mathcal{M} = \{M_{DOC}, M_{FIN}, M_{HR}, M_{EXT}, M_{CTRL}, M_{AI}, M_{RES}\}$.

Ҷадвали 4.1. Мувофиқати микросервисҳо бо зерсохторҳои Кумита

Зерсохторҳои Кумита	**Номи архитектурии микросервис**	**Номи функсионалии микросервис**
Таҳсилот (U_1, U_2)	M_{DOC}	Сервиси санадҳо (Document Service)
Молия ва муҳосибот (U_3)	M_{FIN}	Сервиси молиявӣ (Financial Service)
Кадрҳо ва котибот (U_4)	M_{HR}	Сервиси кадрӣ (HR Service)
Муносибатҳои байналмилалӣ (U_5)	M_{EXT}	Сервиси коммуникатсияи беруна (External Communication Service)
Ҳуқуқ ва назорат (U_6)	M_{CTRL}	Сервиси идоракунӣ ва назорат (Control Service)
ТИ ва рақамикунонӣ (U_7) – технологияҳои иттилоотӣ	M_{AI}	Сервиси зеҳни сунъӣ ва инфрасохторӣ (AI & Infra Service)
Хоҷагидорӣ (U_8)	M_{RES}	Сервиси захираҳо ва бойгонӣ (Resouce Service)

Ба ҷуз микросервисҳое, ки дар ҷадвали 4.1 зикр шудаанд, дар модели математикӣ ду компоненти инфрасохтории (микросервиси) дигар низ ба ҳисоб гирифта шудаанд: M_{API} (шлюз) ва M_{AUTH} (амният), ки функсияҳои умумисистемавии платформаро иҷро мекунанд. Онҳо ҳампайвандӣ, ҳамкорӣ ва амниятро барои ҳамаи микросервисҳои функсионалӣ таъмин менамоянд, вале ба воҳидҳои мушаххаси Кумита бастагӣ надоранд. Дар ин ҳолат модели математикӣ барои сохтори система ба шакли кортежи зерин пешниҳод карда мешавад: $\mathcal{M} = \langle M_{API}, M_{AUTH}, \{M_{DOC}, M_{FIN}, M_{HR}, M_{EXT}, M_{CTRL}, M_{AI}, M_{RES}\}\rangle$.

Дар чунин шакл модел модулнокӣ ва қиёспазирии системаро таъмин сохта, бе халалдор гардонидани якпорчагии функсияҳои умумисистемавӣ имкон медиҳад, ки иқтидори функсионалии ҳар як воҳиди Кумита мустақилона рушд дода шавад. Табодули иттилоот байни компонентҳои модел дар асоси шинаи ягонаи рӯйдодҳо ё ҳамкории REST (модули зеҳнии AI) ташкил карда мешавад.

Барои Кумита нишондиҳандаҳои самаранокии система (дақиқӣ – $\mathcal{P}_r$, пуррагӣ – $\mathcal{R}_c$, $\mathcal{F}_1$-баллҳо) ҳангоми коркарди фармонҳо (U_1, U_2) ва санадҳои молиявӣ (U_3) аҳамияти ҳалкунанда доранд, зеро ҳар гуна хатогӣ ($\mathcal{F}_\mathcal{P}$) ё нокомӣ ($\mathcal{F}_\mathcal{N}$) дар раванди истихроҷи маблағҳо ё ному насаб метавонад ба номуайяниҳои ҳуқуқӣ оварда расонад. Барои таъмини эътимоднокии гардиши санадҳо дар Кумита, қиматҳои ҳудудии ченкуниҳои зерин муқаррар карда шудаанд: $\mathcal{P}_r \geq 0{,}92$; $\mathcal{R}_c \geq 0{,}90$; $\mathcal{F}_1 \geq 0{,}85$.

Вақти миёнаи коркарди санадҳо бо формулаи $\mathcal{T}_{avg} = \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} t_j$ ҳисоб карда мешавад, ки дар он $\mathcal{N}$ – шумораи умумии санадҳои дар давраи ҳисобкунӣ коркардшуда ва t_j – вақти барои коркарди санади мушаххаси j-юм (аз лаҳзаи ворид шудан ба микросервис то гирифтани натиҷа) сарфшударо ифода менамоянд.

Иқтидори гузарониши санадҳо бо формулаи $\Theta = \frac{\mathcal{N}}{\mathcal{T}_{total}}$ ҳисоб карда мешавад: Θ – шумораи санадҳои аз тарафи система дар як воҳиди вақт (масалан, *санад/сония*) коркардшаванда, $\mathcal{N}$ – шумораи вазифаҳои иҷрошуда, $\mathcal{T}_{total}$ – вақти умумии мушоҳида ё кори система мебошад. Ин нишондиҳандаи асосии иқтидори (throughput) система ба ҳисоб меравад. Агар қимати Θ аз шиддати сели воридшавандаи санадҳо хурд бошад, дар система навбат афзоиш меёбад.

Модели гибридии ба нокомиҳо устувор (механизми эҳтиётӣ – Fallback):

$$f(x) = \begin{cases} f_{AI}(x), & \text{агар AI дастрас бошад} \\ f_{rule}(x), & \text{агар AI дастрас набошад.} \end{cases}$$

Ҳангоми паст шудани дараҷаи эътимоднокии f_{AI} аз ҳадди ρ<0,7 ё дар ҳолати дастнорас будани иқтидорҳои GPU, система ба таҳлили пешгӯишавандаи қоидаасос f_{rule} мегузарад.

Системаи ниҳоӣ (модели интегралӣ): Ҷамъшавии навбати санадҳо бо муодилаи дифференсиалӣ тавсиф карда мешавад:

$$\frac{dD}{dt} = \lambda_{in} - \lambda_{proc}(\mathcal{M}, \mathcal{A}),$$

ки λ_{in} – шиддати (сели) санадҳои воридшаванда, λ_{proc} – суръати коркарди санадҳо мебошанд. Тағйирёбии ҳаҷми санадҳои коркарднашуда dD/dt аз тавозуни иттилооти воридшаванда λ_{in} ва маҳсулнокии архитектураи микросервисӣ λ_{proc} вобаста аст. Шарти устувории система $\lambda_{proc} > \lambda_{in}$ мебошад, дар акси ҳол навбат D беохир афзоиш меёбад.

Натиҷаи муҳим алгоритмиронии сикли пурраи коркарди санадҳо мебошад, ки ҳамбастагии ба рӯйдодҳо асосёфтаи микросервисҳо, коркарди асинхронии маълумот ва контури формализатсияшудаи идоракунии равандҳоро дар бар мегирад (расми 4.3).

Ин имкон дод, ки қиёспазирии система аз ҳисоби васеъшавии уфуқӣ таъмин гардад, таъхирҳои коркард коҳиш ёбанд ва самаранокии умумӣ баланд бардошта шавад. Тасвиби амалӣ коҳиши вақти коркарди санадҳоро ба андозаи 60-70% ва фаъолияти устувори системаро ҳангоми афзоиши сарборӣ муаррифӣ кард.

Татбиқи системаи зеҳнии микросервисӣ дар фаъолияти Кумита имкон медиҳад, ки нишондиҳандаҳои зерин ба даст оварда шаванд:

1. Коҳиш ёфтани вақти коркарди санадҳо ба андозаи 60-70% аз ҳисоби истихроҷи автоматии моҳиятҳо (NER) ва таснифбандӣ, ки воридсозии дастии маълумотро ба ҳадди ақал мерасонад.

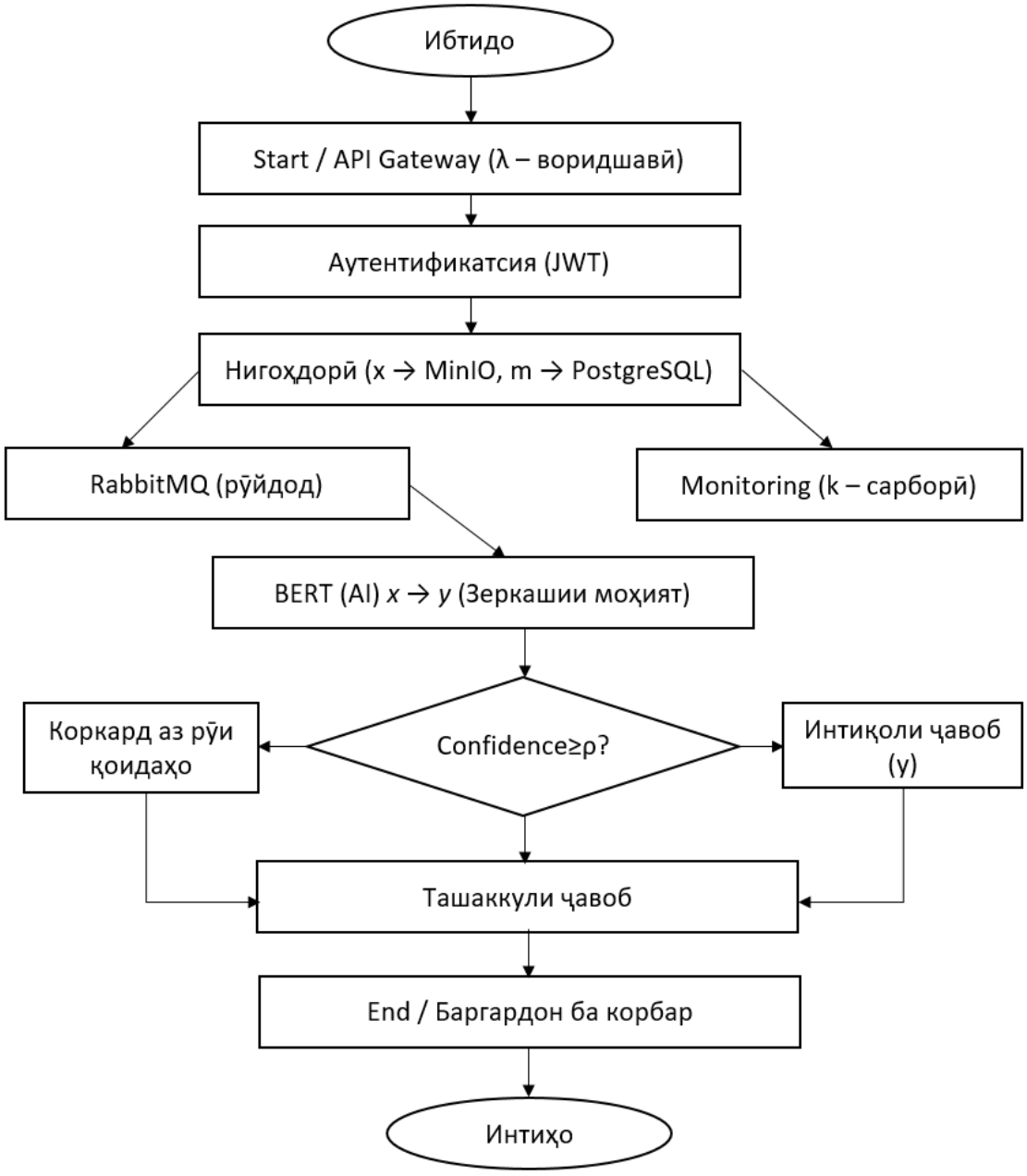
2. Боло рафтани дақиқии маълумот то 92-95% (тибқи ченакҳои ҳадафии $\mathcal{P}_r$ ва $\mathcal{F}_1$), ки хавфи хатогиҳои ҳуқуқӣ ва молиявиро дар зерсохторҳои U_3 ва U_4 коҳиш медиҳад.

3. Паст шудани хароҷоти амалиётӣ барои дастгирии инфрасохтори ТИ аз ҳисоби миқёсгирии уфуқӣ ва истифодаи иқтидорҳои озоди ҳисоббарорӣ барои кори модели BERT.

4. Оптимизатсияи назорати иҷро: кӯтоҳ шудани сикли мувофиқаи санадҳо байни зерсохторҳо ($U_i \to U_j$) то 40% аз ҳисоби мониторинги фарогир дар M_{CTRL}.

Системаи пешниҳодшуда ба сохтори кулли муассисаҳои давлатӣ мутобиқшаванда буда, хусусиятҳои муҳити бисёрзабона (тоҷикӣ, русӣ, англисӣ) ва маҳдудиятҳои инфрасохтори ҳисоббарориро ба назар мегирад. Имконияти ҷорисозии марҳила ба марҳила, истифодаи

контейнеризатсия ва оркестратсия, инчунин оптимизатсияи моделҳо (аз ҷумла усулҳои фишурдасозӣ) роҳи ҳалли пешниҳодшударо барои шароити захираҳои маҳдуд ва инфрасохтори рушдёбандаи технологияҳои иттилоотии кишвар қобили татбиқ мегардонад.



Расми 4.3. Блок-нақшаи алгоритми идоракунии гардиши санадҳои зеҳнии микросервисӣ, ки равандҳои дохилисистемавиро ба расмият меорад

Ҳамзамон, як қатор маҳдудиятҳо муайян карда шудаанд: таъхири хулосаҳои шабакаи нейронӣ системаро барои коркарди баставии санадҳо самараноктар мегардонад, дар ҳоле ки коркарди маълумоти дастнавис интегратсияи технологияҳои OCR-ро талаб мекунад. Таҳлили хатогиҳо зарурати васеъ намудани интихобҳои омӯзишӣ ва ҷорӣ сохтани қоидаҳои махсуси доменнӣ-соҳавиро барои баланд бардоштани сифати истихроҷи моҳиятҳои мураккаб пеш оварад.

Дар маҷмуъ, модели пешниҳодшуда як роҳи ҳалли мукаммали архитектурӣ-алгоритмӣ мебошад, ки дақиқии баланд, қиёспазирӣ ва эътимоднокиро таъмин месозад. Он метавонад ҳамчун асос барои таҳияи системаҳои зеҳнии гардиши санадҳо хизмат намуда, ба рушди

идоракунии электронӣ ва баланд бардоштани самаранокии равандҳои рақамӣ мусоидат намояд.

Дурнамои таҳқиқоти минбаъда ба оптимизатсияи моделҳои трансформерӣ, васеъ намудани дастгирии бисёрзабона, ҷорӣ намудани механизмҳои худомӯзӣ ва такмили тақсимоти сарборӣ дар муҳити микросервисӣ вобаста мебошад.

ХУЛОСАҲО

1. Натиҷаҳои асосии илмии диссертатсия

Асоси таҳқиқоти анҷомдодашударо натиҷаҳои илмии зерин ташкил додаанд:

1. Тасаввуроти назариявӣ ва амалӣ оид ба рушди архитектураи микросервисӣ ҳамчун ба асоси бунёди системаҳои зеҳнии қиёсшаванда ташаккул ёфтааст, ки декомпозитсияи функсионалӣ, мустақилияти компонентҳо ва устуворӣ ба нокомиҳоро таъмин менамояд [6-А; 7-А].

2. Модели архитектурии системаи зеҳнии гардиши санадҳо таҳия гардидааст, ки ба принсипҳои равиши микросервисӣ асос ёфта, дорои компоненти ҷудошудаи таҳлили зеҳнӣ (AI Service) мебошад, ки он мустақилона миқёсгирии масъалаҳои ҳисоббарории мураккаб ва чандирии рушди системаро ташкил менамояд [3-А].

3. Модели математикӣ-алгоритмии коркарди санадҳо пешниҳод гардидааст, ки ба татбиқи архитектураи трансформерии BERT асос ёфта, дақиқии баланди истихроҷи моҳиятҳои номдор ва сохторбандии иттилооти матниро таъмин менамояд [1-А; 4-А].

4. Усули мутобиқсозии моделҳои бисёрзабона барои шароити забонҳои камзахира (масалан, тоҷикӣ) таҳия гардидааст, ки дар асоси генератсияи синтетикии маълумоти омӯзишӣ ва омӯзиши иловагӣ имконияти ба даст овардани сифатеро таъмин менамояд, ки он бо сифати забонҳои дорои захираҳои баланд муқоисашаванда мебошад [5-А].

5. Равиши гибридии коркарди маълумот пешниҳод гардидааст, ки усулҳои омӯзиши мошинӣ ва қоидаасосро бо ҳам муттаҳид месозад ва ҳамзамон дақиқии баланд, устуворӣ ва тафсиршавандагии натиҷаҳоро таъмин мекунад [3-А].

6. Усулҳои қиёскунии системаҳои микросервисӣ, аз ҷумла қиёскуниҳои уфуқӣ ва амудӣ таҳқиқ ва танзим карда шуда, самаранокии истифодаи якҷояи онҳо дар системаҳои зеҳнӣ асоснок гардидааст [8-А; 10-А].

7. Нақши контейнеризатсия ва оркестратсия ҳамчун механизмҳои калидии технологӣ барои идоракунии сикли ҳаётии микросервисҳо асоснок карда шудааст, ки қиёскунии автоматӣ, устуворӣ ба нокомиҳо ва оптимизатсияи истифодаи захираҳоро таъмин менамоянд [3-А; 11-А].

8. Санҷиши таҷрибавии системаи пешниҳодшуда гузаронида шудааст, ки самаранокии кори онро тасдиқ менамояд: дақиқии коркарди санадҳо ба таври назаррас афзоиш ёфта, вақти коркарди онҳо якчанд маротиба коҳиш ёфтааст [1-А].

9. Маҳдудиятҳои равиши таҳияшуда муайян карда шудаанд, ки онҳо ба талаботи баланд нисбат ба захираҳои ҳисоббарорӣ, хусусиятҳои коркарди санадҳои ғайри сохторӣ ва дастнависҳо ва инчунин ба зарурати мутобиқсозии иловагии моделҳо барои соҳаҳои нави мавзуӣ алоқаманд мебошанд [2-А].

10. Дар асоси модели математикии системаи зеҳнии гардиши санадҳо, ки ба архитектураи микросервисӣ ва моделҳои трансформерии навъи BERT такя карда, ба сохтори ташкилӣ-функсионалии Кумита оид ба таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон мутобиқ гардонида шудааст, алгоритми татбиқии система таҳия гардидааст. Алгоритми мазкур бо услуби муҳандисӣ-амалӣ пешниҳод шуда, пайдарпайии марҳилаҳои ҷойгиркунӣ, интегратсия ва истифодаи системаи зеҳниро дар шароити инфрасохтори воқеии иттилоотӣ инъикос менамояд [1-А–5-А].

2. Тавсияҳо оид ба истифодаи амалии натиҷаҳо

Арзиши амалии таҳқиқот бо он таҳким меёбад, ки роҳи ҳалли пешниҳодшуда дар сатҳи назариявӣ маҳдуд набуда, то марҳилаи татбиқи муҳандисӣ расонида шудааст (Замимаи 1).

Алгоритми татбиқии таҳияшуда, ки бар асоси модели математикии система ва мутобиқ ба шароити фаъолияти Кумита оид ба таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон сохта шудааст (Замимаи 2), метавонад ҳамчун равиши намунавӣ барои рақамигардонии гардиши санадҳо дар муассисаҳои давлатӣ истифода шавад.

Натиҷаҳои диссертатсионӣ метавонанд барои истифодаи амалӣ дар самтҳои зерин тавсия шаванд:

1. *Сектори давлатӣ:* Пешниҳод карда мешавад, ки система дар мақомотҳои идоракунии давлатӣ татбиқ карда шавад, аз ҷумла бо истифода аз алгоритми муҳандисӣ-амалӣ, ки интегратсияи пайдарпайи системаи зеҳниро дар инфрасохтори мавҷудаи гардиши санадҳо таъмин менамояд.

2. *Ташкилотҳои тиҷоратӣ:* Истифодаи системаи мазкур барои автоматизатсияи коркарди шартномаҳо, ҳисобфактураҳо ва дигар санадҳои ҳуқуқӣ муфид мебошад, зеро он болоравии самаранокии равандҳои тиҷоратӣ ва коҳишёбии хароҷоти амалиётиро таъмин сохта метавонад.

3. *Сектори молиявӣ:* Роҳи ҳалли пешниҳодшударо барои иҷрои равандҳои KYC, коркарди дархостҳои қарзӣ ва гузаронидани санҷишҳои мувофиқатӣ бо истифода аз таҳлили зеҳнии санадҳо истифода бурдан мумкин аст.

4. *Инфрасохтори ТИ дар созмонҳо:* Барои сохтани системаҳои қиёсшаванда ва устувор ба нокомиҳо пешниҳод карда мешавад, ки усули архитектураи микросервисӣ, контейнеризатсия ва оркестратсия истифода бурда шавад.

5. *Рушди системаҳои зеҳни сунъӣ:* Усулҳои мутобиқсозии моделҳои BERT метавонанд зимни таҳияи системаҳои NLP барои дигар забонҳои камзахира ва соҳаҳои махсуси мавзуӣ истифода шаванд.

6. *Тадқиқоти илмӣ ва таҳсилот:* Натиҷаҳои диссертатсиониро дар тадқиқоти илмии соҳаи системаҳои тақсимшуда, омӯзиши мошинӣ ва трансформатсияи рақамӣ ва инчунин дар барномаҳои таълимии фанни технологияи иттилоотӣ истифода бурдан ба мақсад мувофиқ аст.

7. *Рушди минбаъда:* Пешниҳод карда мешавад, ки равиши таҳияшуда минбаъд дар самтҳои зерин инкишоф дода шавад:

- оптимизатсияи моделҳо (дистиллятсия, квантикунонӣ);
- татбиқи усулҳои мултимодалӣ барои коркарди маълумот;
- баланд бардоштани тафсиршавандагии моделҳо;
- интегратсия бо платформаҳои саноатӣ барои идоракунии сикли ҳаёти ML.

Ҳамин тариқ, модели математикии таҳияшуда, татбиқи архитектурии он ва алгоритми амалӣ-муҳандисии ҷорисозандаи система асоси илмию амалӣ ва мукаммали сохтани системаҳои зеҳнии гардиши санадҳои насли навро ташаккул медиҳанд. Мутобиқсозии роҳи ҳалли пешниҳодшуда ба сохтори Кумита оид ба таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон самаранокии онро дар шароити воқеӣ ва мувофиқати онро бо самтҳои стратегии рушди рақамии Ҷумҳурии Тоҷикистон тасдиқ мекунад.

АДАБИЁТ

1. Феҳристи сарчашмаҳои истифодашуда

[1] Васвани, А. Внимание – это все, что вам нужно [Текст] / А. Васвани, Н. Шазир, Н. Пармар и др. // Достижения в системах обработки нейронной информации. – 2017. – Т. 30. – С. 5998-6008.

[2] Воинов, А. А. Имитационная модель Кайраккумского водохранилища [Текст] / А.А. Воинов, Ф.С. Комилов. – М.: ВЦ АН СССР, «Сообщения по прикладной математике», 1986. – 40 с.

[3] Гадоев, А. Ф. Разработка и реализация управленческих решений на предприятиях легкой промышленности Таджикистана в условиях ускоренной индустриализации.

Монография [Текст] / А.Ф. Гадоев, А. Абдалимов, Ф.С. Комилиён. – Душанбе: Бахманруд, 2022. – 155 с.

[4] Девлин, Дж. BERT: предобучение глубоких двунаправленных трансформеров для понимания языка [Текст] / Дж. Девлин, М.В. Чанг, К. Ли, К. Тутанова // Труды конференции Североамериканского отделения Ассоциации компьютерной лингвистики 2019. – 2019. – С. 4171-4186.

[5] Ёров, М. Р. Татбиқи системаи хизматрасонии оммавӣ дар коркарди дархостҳои онлайнӣ [Матн] / М.Р. Ёров, Ф.С. Комилиён // Паёми Донишгоҳи миллии Тоҷикистон. Бахши илмҳои табиӣ. – 2023. – № 2. – С. 42-53. – EDN: QTPBDV.

[6] Комилиён, Ф. С. Амсилаи консептуалии низоми иттилоотии идоракунии сифати хизматрасониҳои таълимӣ дар муассисаи таҳсилоти олии касбӣ [Матн] / Ф.С. Комилиён, З.Ф. Раҳмонзода // Паёми Донишгоҳи миллии Тоҷикистон. Бахши илмҳои иҷтимоӣ-иқтисодӣ ва ҷамъиятӣ. – 2021. – № 3. – С. 127-135.

[7] Комилиён, Ф. С. Амсиласозии компютерии системаи хизматрасонии шабакавии дар вақти дискретӣ бо тартиби инверсионӣ ва афзалияти тасодуфӣ амалкунандаи ПД КОА [Матн] / Ф.С. Комилиён, М.Р. Ёров // Паёми Донишгоҳи миллии Тоҷикистон. Бахши илмҳои табиӣ. – 2020. – №2. – С. 68-79. – EDN: AJPQAI.

[8] Комилиён, Ф. С. Информатика. Китоби дарсӣ барои донишҷӯёни муассисаҳои таҳсилоти олии касбӣ. Ҷилди 2 [Матн] / Ф.С. Комилиён. – Душанбе: «Донишварон», 2019. – 464 с. – EDN: RNYCDB.

[9] Концепция формирования электронного правительства Республики Таджикистан. Утверждена постановлением Правительства Республики Таджикистан от 30 декабря 2011 года № 643.

[10] Национальная стратегия развития образования Республики Таджикистан на период до 2030 года. Утверждена постановлением Правительства Республики Таджикистан от 29 сентября 2020 года № 526. – 90 с.

[11] Программа реализации концепции государственной информационной политики Республики Таджикистан. Утверждена постановлением Правительства Республики Таджикистан от 28 мая 2009 года № 307.

[12] Саидзода, И. М. Барномасозии компютерии фаъолияти занбӯри асал вобаста ба таъсири бемориҳо ва зараррасонҳо [Матн] / И.М. Саидзода, Ф.С. Комилиён // Паёми донишгоҳи давлатии Хоруғ. – 2023. – № 1 (25). – С. 103-115.

[13] Burns, B. Kubernetes: Up and Running [Text] / B. Burns, J. Beda, K. Hightower. – Sebastopol: O'Reilly Media, 2019. – 330 p.

[14] Chiticariu, L. Rule-based information extraction [Text] / L. Chiticariu, Y. Li, F. Reiss // Proceedings of EMNLP 2013. – 2013. – P. 827-832.

[15] Devlin, J. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [Text] / J. Devlin, M. W. Chang, K. Lee, K. Toutanova // Proceedings of NAACL-HLT. – 2019. – P. 4171-4186.

[16] Earl, T. Service-Oriented Architecture: Concepts, Technologies, and Design [Text] / T. Earl. – Upper Saddle River: Prentice Hall, 2005. – 760 p.

[17] Fowler, M. Guide to Microservices [Electronic resource] / M. Fowler // URL: https://martinfowler.com/microservices/(Date of the application: 5.07.2022).

[18] Fowler, M. Microservices Architecture [Электронный ресурс] / M. Fowler // URL: https://martinfowler.com/microservices/ (Date of the application: 5.07.2022).

[19] Mazlami, G. Extraction of microservices from monolithic software architectures [Electronic resource] / G. Mazlami et al. // International Conference on Web Services (ICWS), 2017, IEEE. – P. 524-531.

[20] Newman, S. Building Microservices: Designing Fine-Grained Systems [Text] / S. Newman. – Sebastopol: O'Reilly Media, 2015. – 280 p.

[21] Ratner, A. Snorkel: Fast Training Data Generation [Text] / A. Ratner, S. Bach, H. Ehrenberg et al. // VLDB. – 2017. – Vol. 11. – No. 3. – P. 269-282.

[22] Richardson, K. Microservices: Development and Refactoring Patterns [Text] / K. Richardson. – Shelter Island: Manning Publications, 2018. – 520 p.

[23] Vaswani, A. Attention Is All You Need [Text] / A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin // Advances in Neural Information Processing Systems (NeurIPS). – 2017. – Vol. 30. – P. 5998-6008.

[24] Wolf, T. Transformers: State-of-the-Art Natural Language Processing [Text] / T. Wolf, L. Debut, V. Sanh et al. // Proceedings of EMNLP. – 2020. – P. 38-45.

2. Феҳристи интишороти илмии довталаби дарёфти дараҷаи илмӣ

а) Мақолаҳое, ки дар маҷаллаҳои тақризшавандаи Комиссияи олии аттестатсионии назди Президенти Ҷумҳурии Тоҷикистон чоп шудаанд:

[1-А] Рахимов, М. Ф. Алгоритмизация микросервисной модели интеллектуального документооборота в профобразовании на базе BERT [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Вестник национального университета. Серия естественных наук. – 2026. – № 2. – С. 98-109. – EDN: OSUPXM.

[2-А] Рахимов, М. Ф. Внедрение микросервисной архитектуры для оптимизации распределения информационных ресурсов [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Наука и инновация. Серия геологических и технических наук. – 2024. – № 2. – С. 71-79. – EDN: UZJNVI.

[3-А] Рахимов, М. Ф. Интеллектуальная система документооборота на основе микросервисной архитектуры и моделей BERT [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Вестник национального университета. Серия естественных наук. – 2026. – № 2. – С. 76-85. – EDN: ROVNVV.

[4-А] Рахимов, М. Ф. Математическое моделирование интеллектуальной системы документооборота на основе микросервисной архитектуры и моделей BERT [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Вестник Технологического университета Таджикистана. – 2026. – № 3 (66). – С. 96-106. – EDN: RWEDOQ.

[5-А] Рахимов, М. Ф. Микросервисная архитектура и BERT-модели в интеллектуальном документообороте профобразования [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. – 2026. – № 2 (74). – С. 62-71. – EDN: VJVTAH.

[6-А] Рахимов, М. Ф. Микросервисная архитектура: от монолита к гибким распределенным системам [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Доклады Национальной академии наук Таджикистана. – 2023. – Т. 66. – № 11-12. – С. 659-667. – EDN: LSVHDL.

[7-А] Рахимов, М. Ф. Анализ характеристик монолитных и микросервисных архитектур [Текст] / М.Ф. Рахимов, Ф.С. Комилиён // Известия Национальной академии наук Таджикистана. Отделение физико-математических, химических, геологических и технических наук. – 2023. – № 4 (193). – С. 44-54. – EDN: GHUYHC.

б) Мақолаҳое, ки дар маводи конференсияҳо ба табъ расидаанд:

[8-А] Рахимов, М. Ф. Микросервисная оптимизация распределения информационных ресурсов с помощью четко определенного «API» [Текст] / Ф.С. Комилиён, М.Ф. Рахимов // Современные проблемы математического моделирования и её применения. Материалы XII международной научно-практической конференции,

посвященной «2020-2040 годы – 20-летия изучения и развития естественных, точных и математических наук в области науки и образования» и «75-летию Таджикского национального университета». – Душанбе, 2024. – С. 28-32. – EDN: PZLKJR.

[9-А] Рахимов, М. Ф. Влияние микросервисных приложений на улучшение работы IT приложений, на примере использования Node.js [Текст] / М.Ф. Рахимов // Материалы республиканской научно-практической конференции «Актуальные проблемы и перспективы развития естественных и точных наук». – Душанбе, ФМГУ, 2023. – С. 12-15.

[10-А] Рахимов, М. Ф. Внедрение микросервисной архитектуры в монолитные системы [Текст] / М.Ф. Рахимов // Материалы международной научно-практической конференции «Наука и образование: тенденции развития в условиях информационного общества», посвященной «75-ой годовщине со дня образования Таджикского национального университета». – Душанбе, ТНУ, 2023. – С. 8-11.

[11-А] Рахимов, М. Ф. Микросервисы в задачах оптимизации распределения информационных ресурсов организации [Текст] / М.Ф. Рахимов, Ф.С. Комилиён // Материалы республиканской научно-практической конференции по теме «Роль интернета в применении информационно-коммуникационных технологий и информационной безопасности в социальных сетях» – Душанбе, АГУ, 2024. – С. 16-24.

[12-А] Rahimov, M. F. Mathematical Modeling of an Intelligent Document Management System Based on Microservice Architecture and BERT Models [Electronic resource] / F. Rahimi, F.S. Komiliyon, M.F. Rahimov, M.R. Yorov // Applied and Computational Mathematics. – 2026. – Vol. 15. – No. 3. – P. 68-74. // https://doi.org/10.11648/j.acm. 20261503.11 (Date of the application: 14.05.2026).

АННОТАЦИЯ

диссертации Рахимова Манучехра Фарходовича на тему «Математическое моделирование интеллектуальной системы документооборота на основе микросервисной архитектуры и нейросетевых механизмов BERT» на соискание учёной степени доктора философии (PhD) по специальности 6D060200 – Информатика (6D060204 – Теоретические основы информатики)

Ключевые слова: микросервисная архитектура, монолитная архитектура, гибкость, надежность, отказоустойчивость, управляемость, масштабируемость, контейнеризация, оркестрация, интеграция, инфраструктура, BERT, NER, DevOps, CI/CD, документооборот.

Актуальность темы исследования: Диссетация посвящена разработке интеллектуальной микросервисной системы управления документооборотом с интеграцией машинного обучения на основе BERT для автоматического извлечения структурированных данных из документов. Предложенная архитектура сочетает преимущества микросервисного подхода с возможностями трансформерных нейронных сетей, что обеспечивает масштабируемость и высокую точность обработки документов.

Целью исследования является разработка и теоретико-прикладное обоснование математической модели интеллектуальной системы документооборота, реализованной на основе микросервисной архитектуры и моделей BERT, обеспечивающей повышение точности обработки документов, масштабируемость и устойчивость функционирования.

Объект исследования – интеллектуальные системы документооборота, функционирующие в условиях цифровой трансформации и построенные на основе распределённых микросервисных архитектур.

Предмет исследования – методы и модели математического обеспечения, алгоритмы и архитектурные решения построения и масштабирования интеллектуальной системы документооборота с использованием трансформерных моделей типа BERT в микросервисной среде.

Методы исследования: моделирование, теоретический, эмпирический, научные выводы, экспертные оценки, сравнительно-сопоставительный анализ, компьютерный эксперимент, количественный и качественный анализ результатов исследований, обобщение результатов.

Аппаратура: персональные компьютеры типа IBM PC, различные сервера.

Научная новизна диссертационной работы заключается в разработке теоретических и прикладных основ построения интеллектуальных систем документооборота на базе микросервисной архитектуры и трансформерных моделей обработки естественного языка, а также в получении следующих оригинальных результатов:

Разработана математическая модель интеллектуальной системы документооборота, интегрирующая принципы микросервисной архитектуры и трансформерные методы обработки текстовой информации, обеспечивающая формализацию процессов извлечения, структурирования и интерпретации данных в распределённой вычислительной среде.

Предложено архитектурное решение микросервисной системы с выделением специализированного компонента интеллектуального анализа (AI Service), отличающееся возможностью независимой масштабируемости вычислительно сложных модулей и обеспечивающее повышение эффективности обработки данных и гибкость развития системы.

Степень применения: Разработанный алгоритм внедрения, адаптированный к условиям функционирования Комитета по профобразованию РТ, может быть использован в качестве типового подхода при цифровизации документооборота в государственных учреждениях.

Область применения: моделирование, принятие управленческих решений, научные исследования, государственные, коммерческие, финансовые, экономические, ИТ-инфраструктура, ИИ, расчетно-диагностические и образовательные задачи.

АННОТАТСИЯИ

диссертатсияи Раҳимов Манучеҳр Фарҳодович дар мавзуи «Моделсозии математикии системаи зеҳнии гардиши санадҳо дар асоси архитектураи микросервисӣ ва механизмҳои нейрошабакавии BERT» барои дарёфти дараҷаи илмии доктори фалсафа (PhD) аз рӯйи ихтисоси 6D060200 – Информатика (6D060204 – Асосҳои назариявии информатика)

Калидвожаҳо: архитектураи микросервисӣ, архитектураи монолитӣ, чандирӣ, эътимоднокӣ, устувории система, идорашавӣ, васеъшавӣ, контейнерсозӣ, оркестратсия, интегратсия, инфрасохтор, BERT, NER, DevOps, CI/CD, гардиши санадщо.

Мубрамии мавзуи диссертатсия: Диссертатсия ба таҳияи системаи зеҳнии идоракунии гардиши санадҳо бо архитектураи микросервисӣ ва интегратсияи омӯзиши мошинӣ бо истифодаи BERT барои истихроҷи автоматии маълумоти сохторёфта аз санадҳо бахшида шудааст. Архитектураи пешниҳодшуда комбинатсияи афзалиятҳои равиши микросервисиро бо имкониятҳои шабакаҳои нейронӣ-трансформерӣ омезиш медиҳад, ки ба васеъшавӣ ва дақиқии баланди коркарди санадҳо мусоидат мекунад.

Мақсади таҳқиқот аз таҳия ва асосноккунии назариявӣ ва амалии модели математикӣ барои системаи зеҳнии гардиши санадҳо иборат аст, ки он бо истифодаи архитектураи микросервисӣ ва моделҳои BERT амалӣ мешавад ва дақиқии коркарди санадҳо, васеъшавӣ ва устувории фаъолияти системаро таъмин менамояд.

Объекти таҳқиқот системаҳои зеҳнии гардиши санадҳо мебошад, ки дар доираи трансформатсияи рақамӣ амал мекунанд ва дар асоси архитектураи микросервисии тақсимшуда сохта шудаанд.

Мавзуи таҳқиқот ба усулҳо ва моделҳои таъминоти математикӣ, алгоритмҳо ва роҳи ҳалли архитектурӣ барои сохтан ва васеъ намудани системаҳои зеҳнии гардиши санадҳо бахшида шудааст, ки бо истифодаи моделҳои трансформерии BERT дар муҳити микросервисӣ амалӣ мегарданд.

Усулҳои таҳқиқот: моделсозӣ, таҳлили назариявӣ ва эмпирикӣ, хулосабарориҳои илмӣ, арзёбиҳои экспертӣ, таҳлили нисбӣ-муқоисавӣ, озмоишҳои компютерӣ, таҳлили миқдорӣ ва сифатии натиҷаҳои таҳқиқотӣ, умумигардонии натиҷаҳо.

Сахтафзор: компютери фардии навъи IBM PC, серверҳои гуногун.

Навгониҳои илмии таҳқиқот: Навовариҳои таҳқиқоти диссертатсионӣ дар таҳияи асосҳои назариявӣ ва амалии коркарди системаҳои зеҳнии гардиши санадҳо бо истифодаи архитектураи микросервисӣ ва моделҳои трансформерӣ барои коркарди забони табиӣ, инчунин ба даст овардани натиҷаҳои аслӣ зоҳир мешавад:

Модели математикии системаи зеҳнии гардиши санадҳо таҳия шудааст, ки принсипҳои архитектураи микросервисӣ ва усулҳои коркарди матн бо трансформерҳоро муттаҳид сохта, равандҳои истихроҷ, сохторсозӣ ва маънидодкунии маълумотро дар муҳити ҳисоббарории тақсимшуда ба расмият медарорад.

Роҳи ҳалли архитектурии системаи микросервисӣ пешниҳод шудааст, ки дорои компоненти махсуси таҳлили зеҳнӣ (AI Service) буда, имконияти васеъшавии мустақили модулҳои ҳисоббарории мураккабро фароҳам меорад ва самаранокии коркарди маълумот ва чандирии рушди системаро таъмин месозад.

Дараҷаи татбиқ: Алгоритми таҳияи системаи зеҳнии гардиши санадҳо, ки ба шароити фаъолияти Кумитаи таҳсилоти ибтидоӣ ва миёнаи касбии назди Ҳукумати Ҷумҳурии Тоҷикистон мутобиқ карда шудааст, метавонад ҳамчун равиши стандартӣ барои рақамигардонии ҳуҷҷатгузорӣ дар муассисаҳои давлатӣ истифода шавад.

Соҳаи истифода: моделсозӣ, қабули қарорҳои идоракунӣ, таҳқиқоти илмӣ, соҳаҳои давлатӣ, тиҷоратӣ, молиявӣ ва иқтисодӣ, инфрасохтори ТИ, зеҳни сунъӣ, вазифаҳои ҳисоббарорӣ ва таълимӣ.

ANNOTATION

of the dissertation by Manuchekhr Rahimov Farhodovich on the topic «Mathematical modeling of an intelligent document management system based on microservice architecture and BERT neural network mechanisms» for the degree of Doctor of Philosophy (PhD) in the specialty 6D060200 – Informatics (6D060204 – Theoretical foundations of Informatics)

Keywords: microservice architecture, monolithic architecture, flexibility, reliability, fault tolerance, manageability, scalability, containerization, orchestration, integration, infrastructure, BERT, NER, DevOps, CI/CD, document management.

The relevance of the research topic: The dissertation is devoted to the development of an intelligent microservice-based document management system integrated with machine learning using BERT for automatic extraction of structured data from documents. The proposed architecture combines the advantages of the microservice approach with the capabilities of transformer neural networks, ensuring system scalability and high accuracy in document processing.

Objective of the Research: The aim of the study is to develop and provide a theoretical and applied justification for a mathematical model of an intelligent document management system implemented on a microservice architecture and BERT models, which ensures increased accuracy in document processing, scalability, and robustness of operation.

Research Object: Intelligent document management systems operating in the context of digital transformation and built on distributed microservice architectures.

Research Subject: Methods and models of mathematical support, algorithms, and architectural solutions for building and scaling an intelligent document management system using transformer models of the BERT type within a microservice environment.

Research Methods: Modeling, theoretical and empirical analysis, scientific reasoning, expert assessments, comparative analysis, computer experiments, quantitative and qualitative analysis of research results, and synthesis of findings.

Equipment: IBM PC-type personal computers and various servers.

Scientific Novelty of the Dissertation: The novelty lies in the development of theoretical and applied foundations for constructing intelligent document management systems based on microservice architecture and transformer models for natural language processing, as well as in obtaining the following original results:

A mathematical model of an intelligent document management system has been developed, integrating the principles of microservice architecture and transformer-based text processing methods, formalizing processes for data extraction, structuring, and interpretation in a distributed computing environment.

An architectural solution for the microservice system has been proposed, featuring a dedicated intelligent analysis component (AI Service), which enables independent scalability of computationally intensive modules and improves data processing efficiency and system flexibility.

Degree of Application: The developed implementation algorithm, adapted to the operational conditions of the Committee on Professional Education of the Republic of Tajikistan, can be used as a standard approach for the digitalization of document management in governmental institutions.

Field of Application: Modeling, managerial decision-making, scientific research, governmental, commercial, financial, and economic sectors, IT infrastructure, artificial intelligence, computational diagnostics, and educational tasks.