

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ МОЛДОВА

Технический Университет Молдовы

Факультет Вычислительной Техники, Информатики и Микроэлектроники

Кафедра Программной Инженерии и Автоматики

Допущен к защите
Заведующий кафедрой:

Ион ФИОДОРОВ, доцент, кандидат наук

„___” _____ 2026

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ОПТИМИЗАЦИИ БАЗ ЗНАНИЙ ОРГАНИЗАЦИЙ

Магистерский проект

Студент: _____ **Дилан Юлиана, ТІА-241М**
Руководитель: _____ **Кунев Вячеслав, лект. унив., док.**
Консультант: _____ **Кожокару Светлана, ассист. унив.**

Chişinău, 2026

АННОТАЦИЯ

Данная работа по теме «ИССЛЕДОВАНИЕ И РАЗРАБОТКА СИСТЕМЫ ПО ОПТИМИЗАЦИИ БАЗ ЗНАНИЙ ОРГАНИЗАЦИЙ» посвящена проектированию и программной реализации интеллектуальной системы на базе локальной языковой модели Gemma 2, интегрированной с локальным хранилищем. В первой части работы проводится подробный анализ и определение предметной области, где обосновывается актуальность внедрения технологий генеративного искусственного интеллекта в рабочие процессы. Здесь рассматриваются схожие системы и существующие аналоги, на основе которых выявляются ключевые потребности в создании защищенного ИИ-ассистента, способного работать внутри конфиденциального контура организации.

Вторая глава посвящена глубокому моделированию и проектированию информационной системы. С помощью набора диаграмм, таких как варианты использования, диаграммы последовательности, классов и развертывания, а также функциональных моделей описывается архитектура будущего решения. Этот этап позволяет визуализировать логику взаимодействия компонентов, от момента отправки запроса пользователем до интеллектуальной обработки данных моделью Gemma 2.

В главе, описывающей реализацию системы, приводится детальное изложение использованных материалов и технологий, послуживших фундаментом для создания продукта. Здесь раскрывается процесс настройки программной среды, интеграция с локальным хранилищем. Описание сопровождается скриншотами функционирующей системы, которые наглядно демонстрируют интерфейс взаимодействия и корректность выполнения заложенных сценариев.

Заключительные разделы работы включают в себя детальный расчет затрат на разработку и внедрение проекта, где оценивается экономическая эффективность выбранных решений. В завершении формулируются выводы о проделанной работе, подтверждающие достижение поставленных целей, и приводится структурированный список источников, использованных в ходе исследования. Таким образом, работа представляет собой полное описание жизненного цикла разработки интеллектуального сервиса, от теоретического обоснования и проектирования до практической реализации и оценки стоимости управления проектом.

REZUMAT

Prezenta Această lucrare cu tema „CERCETAREA ȘI DEZVOLTAREA UNUI SISTEM PENTRU OPTIMIZAREA BAZELOR DE CUNOȘTINȚE DIN ORGANIZAȚII” este dedicată proiectării și implementării software a unui sistem inteligent bazat pe modelul lingvistic local Gemma 2, integrat cu un depozit local de date. În prima parte a lucrării se realizează o analiză detaliată și definirea domeniului de studiu, unde se argumentează relevanța implementării tehnologiilor de inteligență artificială generativă în procesele de lucru. Sunt examinate sisteme similare și analogii existente, pe baza cărora sunt identificate necesitățile-cheie pentru crearea unui asistent AI securizat, capabil să funcționeze în interiorul perimetrului confidențial al organizației.

Al doilea capitol este dedicat modelării aprofundate și proiectării sistemului informațional. Cu ajutorul unui set de diagrame, precum cazuri de utilizare, diagrame de secvență, de clase și de implementare (deployment), precum și modele funcționale, este descrisă arhitectura viitoarei soluții. Această etapă permite vizualizarea logicii de interacțiune a componentelor, de la momentul trimiterii unei solicitări de către utilizator până la procesarea inteligentă a datelor de către modelul Gemma 2.

Capitolul dedicat implementării sistemului oferă o prezentare detaliată a materialelor și tehnologiilor utilizate, care au constituit fundamentul pentru crearea produsului. Aici este descris procesul de configurare a mediului software și integrarea cu depozitul local. Descrierea este însoțită de capturi de ecran ale sistemului funcțional, care demonstrează vizual interfața de interacțiune și corectitudinea executării scenariilor implementate.

Secțiunile finale ale lucrării includ un calcul detaliat al costurilor de dezvoltare și implementare a proiectului, unde este evaluată eficiența economică a soluțiilor alese. În încheiere sunt formulate concluzii privind activitatea realizată, confirmând atingerea obiectivelor propuse, și este prezentată o listă structurată a surselor utilizate în cadrul cercetării. Astfel, lucrarea oferă o descriere completă a ciclului de viață al dezvoltării unui serviciu inteligent, de la fundamentarea teoretică și proiectare până la implementarea practică și evaluarea costurilor de gestionare a proiectului.

ABSTRACT

This work, titled “RESEARCH AND DEVELOPMENT OF A SYSTEM FOR THE OPTIMIZATION OF ORGANIZATIONAL KNOWLEDGE BASES,” is dedicated to the design and software implementation of an intelligent system based on the local language model Gemma 2, integrated with a local storage system. The first part of the paper provides a detailed analysis and definition of the subject domain, justifying the relevance of implementing generative artificial intelligence technologies in work processes. It examines similar systems and existing analogues, based on which key requirements are identified for creating a secure AI assistant capable of operating within an organization’s confidential environment.

The second chapter is devoted to in-depth modeling and design of the information system. Using a set of diagrams—such as use case diagrams, sequence diagrams, class diagrams, and deployment diagrams—as well as functional models, the architecture of the future solution is described. This stage allows for visualization of the interaction logic between components, from the moment a user submits a request to the intelligent processing of data by the Gemma 2 model.

The chapter describing system implementation provides a detailed overview of the materials and technologies used as the foundation for the product. It outlines the process of configuring the software environment and integrating with local storage. The description is accompanied by screenshots of the functioning system, which clearly demonstrate the interaction interface and the correct execution of the implemented scenarios.

The final sections of the paper include a detailed calculation of the costs of development and project implementation, evaluating the economic efficiency of the chosen solutions. In conclusion, the results of the work are summarized, confirming the achievement of the set objectives, and a structured list of sources used in the research is provided. Thus, the paper presents a comprehensive description of the lifecycle of developing an intelligent service, from theoretical justification and design to practical implementation and evaluation of project management costs.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	8
1 АНАЛИЗ И ОПРЕДЕЛЕНИЕ ОБЛАСТИ ИССЛЕДОВАНИЯ.....	9
1.1 Важность темы.....	9
1.2 Схожие системы.....	10
1.3 Цель, задачи и требования системы.....	13
2 МОДЕЛИРОВАНИЕ И ПРОЕКТИРОВАНИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ.....	16
2.1 Моделирование и поведенческое описание системы.....	16
2.1.1 Общее представление системы.....	17
2.1.2 Моделирование потоков внутри системы с помощью диаграмм деятельности.....	19
2.1.3 Описание сценариев использования приложения.....	21
2.2 Структурное описание системы.....	23
2.2.1 Отношения между компонентами.....	24
2.2.2 Описание процессов приложения.....	27
3 РЕАЛИЗАЦИЯ СИСТЕМЫ.....	31
3.1 Взаимодействие с системой.....	34
4 ОЦЕНКИ СТОИМОСТИ УПРАВЛЕНИЯ ПРОЕКТОМ ПРЕДЛАГАЕМОЙ К РЕАЛИЗАЦИИ СИСТЕМЫ.....	42
ВЫВОДЫ.....	45
БИБЛИОГРАФИЯ.....	46

ВВЕДЕНИЕ

Предметом данной научно-практической работы стала разработка и внедрение комплексного программного решения, которое нацелено на глубокую оптимизацию и структурирование корпоративных баз знаний с использованием передовых технологий искусственного интеллекта. В современных условиях избытка информации данная система существенно улучшает пользовательский опыт сотрудников, значительно упрощая и ускоряя процесс повседневного взаимодействия с накопленными интеллектуальными активами компании. Разработанная система эффективно объединяет в себе высокие вычислительные возможности локальной языковой модели семейства Gemma 2 и привычное удобство интерфейса мессенджера Telegram, создавая надежный и интуитивно понятный инструмент для оперативного анализа больших объемов неструктурированной корпоративной информации в режиме реального времени.

На начальном этапе проектирования будущей системы были четко определены, классифицированы и задокументированы функциональные и нефункциональные требования. Данные спецификации ориентированы на решение реальных бизнес-задач и устранение трудностей, которые наиболее часто встречаются в профессиональной деятельности сотрудникам крупных организаций. Важно заметить, что для качественной визуализации архитектурных решений и детальной проработки логики внутренних процессов был проведен комплексный этап системного моделирования, в ходе которого были составлены все необходимые графические диаграммы в соответствии со стандартами проектирования.

В данном отчете максимально подробно рассматривается пакет проектной документации, включающий диаграммы вариантов использования, деятельности, состояний, последовательности и классов. В своей совокупности эти модели дают наиболее полное и прозрачное представление о фундаментальных принципах работы системы и её внутренней иерархии. Проведенное моделирование позволило детально описать потоки данных и алгоритмы обработки запросов, выстроить четкое, отказоустойчивое взаимодействие между конечным пользователем, разработанным на языке Python программным скриптом и защищенным локальным хранилищем файлов, обеспечивая целостность всей инфраструктуры.

Основной и неизменной целью моей работы является создание максимально доступного, гибкого и при этом полностью защищенного инструмента автоматизации офисных процессов. Продукт позволяет в считанные секунды извлекать необходимую информацию из разноформатных документов (текстовых файлов, таблиц, отчетов), хранящихся исключительно во внутреннем контуре компании. Это полностью исключает риск утечки коммерческой тайны и передачу конфиденциального контента на сторонние облачные сервисы или публичные API сторонних разработчиков ИИ.

БИБЛИОГРАФИЯ

- [1] “Every SharePoint Site Now Has an Agent,” Lightning Tools Blog. [Онлайн]. Доступно: <https://lightningtools.com/blog/every-sharepoint-site-now-has-an-agent/>
- [2] “Glean Platform,” Glean. [Онлайн]. Доступно: <https://www.glean.com/>
- [3] M. Farcas, “LlamaIndex vs LangChain,” n8n Blog. [Онлайн]. Доступно: <http://blog.n8n.io/llamaindex-vs-langchain/>
- [4] А. У. Актаева, Управление проектами: информационные технологии. [Онлайн]. Доступно: <https://library.atu.edu.kz/files/41383.pdf>
- [5] И. Соммервиль, Инженерия программного обеспечения.
- [6] Р. Прессман, Инженерия программного обеспечения: практический подход.
- [7] Ф. Шолле, Глубокое обучение на Python.
- [8] М. Ожегов, Н. Эйсмонт, М. Соболева, “Кто такой бизнес-аналитик и чем он занимается,” Яндекс Практикум. [Онлайн]. Доступно: <https://practicum.yandex.ru/blog/professiya-biznes-analitika/>
- [9] C. Farabet и T. Warkentin, “Gemma 2 теперь доступна исследователям и разработчикам,” Google DeepMind Blog. [Онлайн]. Доступно: <https://blog.google/innovation-and-ai/technology/developers-tools/google-gemma-2/>
- [10] Т. Х. Кормен, Ч. Лейзерсон, Р. Ривест, К. Штайн, Алгоритмы: построение и анализ, 3-е изд. Cambridge, MA: MIT Press, 2009.
- [11] С. Рассел, П. Норвиг, Искусственный интеллект: современный подход, 4-е изд. Pearson, 2020.
- [12] И. Гудфеллоу, Й. Бенджио, А. Курвиль, Глубокое обучение. MIT Press, 2016.
- [13] М. Фаулер, Рефакторинг: улучшение существующего кода, 2-е изд. Addison-Wesley, 2018.
- [14] К. Швабер, Д. Сазерленд, “Руководство по Scrum,” 2020. [Онлайн]. Доступно: <https://scrumguides.org/>
- [15] “Документация LangChain,” LangChain. [Онлайн]. Доступно: <https://docs.langchain.com/>
- [16] “Документация LlamaIndex,” LlamaIndex. [Онлайн]. Доступно: <https://docs.llamaindex.ai/>
- [17] “Документация моделей GPT,” OpenAI, 2024. [Онлайн]. Доступно: <https://platform.openai.com/docs/>
- [18] “Документация Vertex AI,” Google Cloud, 2024. [Онлайн]. Доступно: <https://cloud.google.com/vertex-ai/docs>
- [19] “Документация Azure AI Services,” Microsoft, 2024. [Онлайн]. Доступно: <https://learn.microsoft.com/en-us/azure/ai-services/>
- [20] “Документация Transformers,” Hugging Face, 2024. [Онлайн]. Доступно: <https://huggingface.co/docs/transformers>