

*Article*

# AI-Assistants: Framework for combining LLM-Driven Assistants with corporate features

Isakulov Zeiin <sup>1\*</sup>

<sup>1</sup>Department of Information Systems, SDU University, Almaty, Kazakhstan

DOI: 10.47344/byvn3m48

## Abstract

This paper presents a comprehensive review of existing studies and several optimization methods, and suggests a simplified workflow for a unified framework for multimodal systems and AI-driven assistants. It focuses on the integration of AI-based tools into existing IT infrastructures by comparing approaches used in industry use cases of applying AI-based tools in their respective fields. Integrating AI into existing IT infrastructures that were not originally designed to be operated by AI, even at the level of data rendering, causes significant challenges that are worth solving. The main challenges include evaluating the rationality of using AI for certain tasks, managing the computational costs of self-deployed AI tools, managing the quality limitations of self-deployed LLMs to ensure that they can effectively interact with humans in industrial environments, and overcoming the economic limitations of using high-end LLM models such as ChatGPT, DeepSeek, or Grok in terms of token-based pricing. The study demonstrates a significant need for standardization and enhanced operability of existing approaches beyond the highly specialized solutions used in fields such as medicine, software development, and network management. The improvements are intended to enable AI assistants to make critical decisions that still require automation and optimization within industrial systems. The main aim is to systematize modern solutions and design common principles and vectors of development for highly contextual and scalable AI operating systems, enabling both AI technology providers and commercial organizations to adapt even in the short term. An example of the workflow pipeline of the proposed framework is designed to combine the adaptability and intellectual operability of high-end LLMs with the standardization of tasks for highly contextual models. It also addresses the deployment limitations of models built from scratch or derived from local models while preserving optimization for servers of small and medium-sized enterprises at each step by managing small modular tasks.

**Keywords:** Artificial Intelligence, AI Adapters, Human-AI interaction, Industrial AI, Process Automation, IT infrastructure, multimodal systems, AI-driven assistants, cost-effective deployment, enterprise AI, scalable AI frameworks, SME AI adoption, computational cost management

Email\*: isakul.qxz@gmail.com    ORCID: 0009-0001-2930-5233

*Received: December 20, 2025. Reviewed: September 26, 2026. Accepted: September 29, 2026.    © 2026 Isakulov Zeiin. All rights reserved.*

## I. INTRODUCTION

Recently, AI-based workflows and ML algorithms have increasingly developed and become the basis for optimizing the work environment. With further growth in the volume of data created and recorded by humanity, and with growing computational requirements, traditional planning algorithms and task-completion methods require enhancement. The IT infrastructure of Big Tech companies, cloud systems, and distributed systems is increasingly using intelligent algorithms and AI-based tools, leaving traditional methods and tools, as well as small and medium-sized enterprises, far behind. The use of neural networks, AI assistants, and metaheuristic algorithms in the early period of their development opened a new chapter in achieving optimization, industrial effectiveness, and automation. These approaches allow us to understand the main benefit of technological progress—adaptability. Traditional solutions, despite their effectiveness in static environments, are not adaptable. The adaptability of a system is necessary for dynamic systems where data and rules are changeable, including large-scale cloud platforms, which are increasingly being adopted even by small and medium-sized enterprises in various fields.

Conventional IT management infrastructure is often dependent on static optimization algorithms and rule-based decision-making. Although this type of procedure is successful in stable environments, it often does not scale or respond to the unpredictability and volatility of today's IT ecosystems. Using different kinds of intelligent processing algorithms can help predict system bottlenecks in existing IT infrastructure, find ways to optimize workflows, and enhance our ability to achieve state-of-the-art adaptability. These algorithms can autonomously predict and identify system bottlenecks and then automatically reposition workflows for continuous operational efficiency, paving the way for adaptability.

It is also important to mention that the implementation of such intelligent tools may cause significant organizational changes, including employee re-education, and raise concerns such as data leaks and quality issues. The creation of artificial intelligence systems that meet both the technical constraints and the strategic investment requirements for algorithms in small and medium-sized enterprises remains an important task that requires a carefully developed approach. In this study, we will explore possible approaches.

## II. LITERATURE REVIEW

Our review of previous studies gives us more context and helps us identify the main gaps, methods, and findings in the industrial field. The literature review examines both recent and older works, allowing us to delve deeper into the subject and helping us position our work on the more practical side of the research. The literature is organized into three thematic areas: optimization and scheduling approaches; studies on the industrialization of the RAG system, Part I; and studies on the industrialization of the RAG system, Part II.

TABLE I: Optimization and Scheduling Approaches

| Study | Key Contributions                                                                                           | Rationale                                                                     |
|-------|-------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| [1]   | <i>RepairLLaMA: Efficient representations and fine-tuned adapters for program repair</i>                    | Provides an example of using adapters and overall workflow pipelines.         |
| [2]   | <i>Exploring the potential of AI-driven optimization in enhancing network performance and efficiency</i>    | Shows the efficiency of AI approaches in industrial cases.                    |
| [3]   | <i>Software project scheduling using the hyper-cube ant colony optimization algorithm</i>                   | Provides evidence of heuristic approaches in industrial scheduling.           |
| [4]   | <i>An investigation of static versus dynamic scheduling</i>                                                 | Demonstrates differences between scheduling approaches.                       |
| [5]   | <i>A roadmap for integrating automation with process optimization for AI-powered digital transformation</i> | Discusses the potential vectors of development and processes to be automated. |

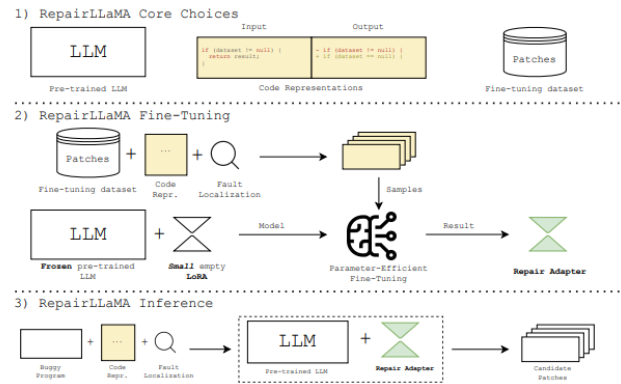


Figure 1: Overview of RepairLLaMA. The core novelties of RepairLLaMA are the APR specific code representations and the engineering of an effective *program repair adapter* that is plugged into the underlying LLM.

Fig. 1: The whole pipeline of the work [1]

The studies have shown the importance of optimization approaches in industrial cases. Early studies provided evidence of heuristic approaches in industrial scheduling [3]. Some of them demonstrate the differences between static and dynamic scheduling, proving the importance of adaptability [4]. However, with today's technological advances, studies discuss the vector of development toward automation, optimization, scheduling, and routing using intelligent tools in industrial environments [2]. Recently, RepairLLaMA demonstrated an application of fine-tuned adapters, enhancing both human-based and AI-based workflows in specific areas [1]. See Fig. 1.

TABLE II: Industrialization of RAG system(PART 1)

| Study | Key Contributions                                                                                | Rationale                                                                                                                                                                                                                                                                               |
|-------|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [6]   | <i>Model Context Protocol (MCP): Landscape, Security Threats, and Future Research Directions</i> | Primarily aimed at simplifying pipeline of deployment, integration and operation with existing technologies on an industrial scale. MCP offers an example of simplifying models workflow with each other in order to expand functionality with external tools and reduce fragmentation. |
| [7]   | <i>What if GPT4 Became Autonomous: The Auto-GPT Project and Use Cases</i>                        | Provides use cases of autonomous AI and include potential vectors of development and enhancement of LLMs.                                                                                                                                                                               |
| [8]   | <i>Auto-GPT for Online Decision Making: Benchmarks and Additional Opinions</i>                   | Demonstrates the key differences across autonomous models and provides Additional Opinions algorithm which enables integrating smaller models without major changing the main LLM                                                                                                       |
| [9]   | <i>Financial Report Chunking for Effective Retrieval Augmented Generation</i>                    | The study shows significant improves in results of RAG in question and answer tasks based on financial tests. The approach allows us to create high-precision RAG systems using file structure instead of full texts.                                                                   |
| [10]  | <i>Searching for Best Practices in Retrieval-Augmented Generation</i>                            | The study examines existing RAG methods to identify best practices to ensure productivity and efficiency.                                                                                                                                                                               |

TABLE III: Industrialization of RAG system(PART 2)

| Study | Key Contributions                                                                                                                                          | Rationale                                                                                                                                                                                                                                                               |
|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [11]  | <i>Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks</i>                                                                                    | The study demonstrates RAG system model. RAG can solve problems of limited and accurate knowledge of LLM and the difficulty of updating them.                                                                                                                           |
| [12]  | <i>When Will AI Exceed Human Performance? Evidence from AI Experts</i>                                                                                     | The study surveyed ML experts to shows that AI will surpass humankind in many areas in the coming years. Giving us potential vectors of implementing human-in-the-loop.                                                                                                 |
| [13]  | <i>Fundamentals of Expert Systems</i>                                                                                                                      | The study demonstrates key principles of creating expert systems and outlines areas for further research.                                                                                                                                                               |
| [14]  | <i>Adaptable Adapters</i>                                                                                                                                  | The study proposes adaptive adapters to select own activation functions for different layers and use a trainable switch to select only necessary layers. The approach preserves quality of standard adapters, yet requires less parameters for better for smaller data. |
| [15]  | <i>Dressing the Imagination: A Dataset for AI-Powered Translation of Text into Fashion Outfits and A Novel KAN Adapter for Enhanced Feature Adaptation</i> | The study proposes NeRA, an innovative adapter architecture. Unlike traditional approaches, NeRA uses trainable nonlinear spline transformations enabling models to provide higher accuracy and deep semantic matching.                                                 |

A significant part of the existing work is aimed at making AI assistants practical and contextual. The foundations of RAG were described and demonstrated to solve the knowledge accuracy and limitations problems of LLMs [11]. But as time goes on, the increase in the efficiency of AI assistants continues [10]. Concrete approaches to making models highly contextual include, for example, the capability to work with financial reports by using better chunking methods. To reduce the fragmentation of AI-based tools, the studies also proposed the Model Context Protocol as a basis for simplifying workflows and the integration of external tools. We can also see a tendency toward more effective and powerful customization of the models. The adaptable adapters can select functions to use and yet use fewer parameters and conduct quality for less volume of data [14]. Some approaches simply suggest architectural changes to the adapters to make models even more accurate by using different computational operations [15].

TABLE IV: Human-AI and AI-AI interaction

| Study | Key Contributions                                                                                      | Rationale                                                                                                                                                                                                                                                                                                                   |
|-------|--------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [16]  | <i>In-Context Retrieval-Augmented Language Models</i>                                                  | The study proposes In-Context RALM, a simpler method in which the architecture of the model does not change in comparison with regular RALM. The context documents are simply added to the beginning of input data. In-Context RALM has great potential for widespread use of the model through the API.                    |
| [17]  | <i>Human-in-the-loop Artificial Intelligence</i>                                                       | The study proposes ethical concept of HitAI. Since humans are involved in any AI system, HitAI suggests rewarding those who provide knowledge. Profits from AI solutions should be returned to the rightful owners of knowledge.                                                                                            |
| [18]  | <i>Towards a Better Understanding of Context and Context-Awareness</i>                                 | The study addresses challenges related to understanding and creating context and suggests approaches of developing contextual applications.                                                                                                                                                                                 |
| [19]  | <i>AI models collapse when trained on recursively generated data</i>                                   | The study explains the risks of future versions of GPT. When most of the text on the Internet will be generated by LLM. It was found that uncontrolled generation of content by the model leads to irreversible defects of future models. The phenomenon is called "model collapse." It is found in many generative models. |
| [20]  | <i>Artificial intelligence implementation in manufacturing SMEs: A resource orchestration approach</i> | The study examines how manufacturing SMEs organize resources to implement AI. Via dynamic orchestration of AI resources, SMEs successfully coordinate processes and develop qualified specialists. The research contributes to AI implementation by showing how SMEs use AI capabilities to create a competitive advantage. |

Integration of AI into existing human-based infrastructures raises important considerations including ethical ones. *Human-in-the-loop Artificial Intelligence* suggests more ethical AI, the HitAI. The concept suggests that the profits from AI in enterprises should

be returned to their owners [17]. The long term risks like model collapse [19], and practical benefits [20] suggest to us that AI must be regulated to be the technology of the future. Some infrastructures were never meant to be driven by AI, but they still need some intelligent solutions. There, contextuality comes in the first place, both for human and AI agents [16] [18]. The reviewed studies demonstrate a lack of a standardized, lightweight framework and development pipeline. The studies present highly effective but specialized solutions [1] [2]. The existing research does not provide a solid platform that addresses computational constraints and implementation challenges faced by SMEs. The reviewed literature on autonomous AIs [8] [7], model interactions [6] and resource orchestration [20] hints at this but does not explicitly propose or test this specific hierarchical architecture as a virtual copy of existing corporate agents in SME infrastructures.

**Case Studies:** Google's use case of ML to optimize its data center cooling systems by analyzing data from sensors ensures that the system can dynamically manage server setup settings resulting in energy and financial savings. AT&T's ML algorithm analyzes patterns in its traffic and can predict demand, minimizing further network latency issues.

The integration of AI and ML algorithms in industrial cases demonstrates huge improvements in network management. Predictive analytics and anomaly detection in real time form the foundation for future AI-based corporate tools. It enables networks to react to current conditions and predict future threats.

The article's main limitation is that it was based on an industry leader's case. It doesn't present quantitative results from practical implementation of AI driven methodologies. It also lacks a detailed analysis of computational costs and complexity of integrating these systems into existing legacy infrastructure.

The article demonstrates effectiveness in energy savings and reduced latency and validates its potential. Its complexity is high, requiring interdisciplinary knowledge. It is a great example of ML optimization applied to load balancing and resource management.

After the literature review, it becomes obvious that enhancements have been made in the AI workflow pipeline. But there remains a substantial gap in the standardization of the frameworks used. That can be a limitation on fully implementing them in existing IT infrastructures. The research mainly focuses on specialized solutions without providing solid frameworks that address the computational constraints and implementation challenges faced by small and medium-sized enterprises.

### III. RESEARCH METHODOLOGY

#### A. Research Questions and Objectives

Our research aims to contribute to the development of the standardization of frameworks used in the AI development pipeline by making them optimized even in small and strongly tied to context-based automation and workflow management in IT infrastructures.

This research addresses the following questions:

1. How can an LLM framework be successfully implemented within existing IT infrastructure to manage an AI assistant with significant computational constraints?
2. How does an AI architecture integrate classical optimization local models to optimize operations in industry with clients?

The aim is to design a framework for optimizing and scheduling tasks in multimodal systems that improves the efficiency and quality of combining models in IT infrastructure. Our aim also includes analyzing the implementation challenges and the organizational effects of intelligent automation. For this purpose, we have to face the following tasks: 1. To analyze the existing methods of implementing multimodal systems. 2. To develop a hybrid model combining LLMs and classical optimization algorithms. 3. To increase the effectiveness of the proposed framework under real-case scenarios.

#### B. Experimental Setup

The experiment was applied to an existing IT infrastructure environment and designed to recreate a real-world enterprise environment. The testbed consists of a digital infrastructure, storage systems, and API components required to test the system in a medium-scale enterprise. The testing system was built using HuggingFace models and OpenRouter API models to orchestrate under varying resource constraints and workload patterns.

The experimental setup includes five primary components tailored for the Unibest case study. The first one is a baseline RAG system implementing a multimodal structure. The second one is the proposed framework integrating several models and tools with the target AI assistant. The third is a management layer implemented as an orchestration UNIA positioned between AI assistants and a sales manager. The fourth one is the operational knowledge stores such as vector databases, relational product databases, rules, etc. The last component is monitoring, logging and CRM integration enabling more context for the system's work. The orchestration LLM acted as the leader of the "AI Swarm" that standardizes corporate tasks across smaller models by performing

prompt engineering and setting up constraints such as policies and rules. Each component was deployed in identical environmental conditions to ensure a fair comparison.

**System deployment:** The core inference model ran on a physical server with AMD Ryzen 7 9800X3D under Debian, an NVIDIA RTX3070 Ti GPU and 32 GB RAM. Copies of UnieAI used a single quantized Nous Hermes Mistral 7B model to reduce graphical memory and inference footprint. External channels were connected via APIs such as Telegram, WhatsApp, Instagram, Gmail. UNIA in the orchestration layer used DeepSeek via OpenRouter API with a constrained budget (50 free requests per day) for selected high-risk decisions and light-weight routing tasks such as managing prompts and controlling the "AI Swarm". Heavier inference and generation were performed locally on the quantized UnieAI.

### *C. System Architecture and Components*

*1) Unibest Business Context:* Unibest is a B2B/B2C company selling computers and related services. To reduce human labor costs, a sales AI assistant, UnieAI, was created to act as a virtual sales manager handling customer communications, product recommendations, making store orders and routine after-sales interactions during non-working hours.

To manage and supervise UnieAI, a manager model named UNIA was implemented. UNIA is built atop the free DeepSeek endpoint available through OpenRouter (limited to 50 free calls per day) and provides orchestration features such as creation and lifecycle handling of UnieAI copies, routing of B2B leads, scheduling B2C review reminders, context recovery and quality correction of UnieAI outputs, especially for the Kazakh and Russian languages. UNIA can autonomously create and manage CRM objects in Bitrix, initiate procurement protocols executed by UnieAI and periodically update B2B data by using a 2GIS parser.

**Knowledge and retrieval stack.** The product catalog, company policies and trade methodology are stored in relational and document stores and mirrored into vector indices using FAISS and Qdrant. Embeddings for Russian and Kazakh were produced using an embedding model from Sberbank. RAG pipelines fetch the top-k passages from vector stores and pass those to UnieAI for grounded response generation.

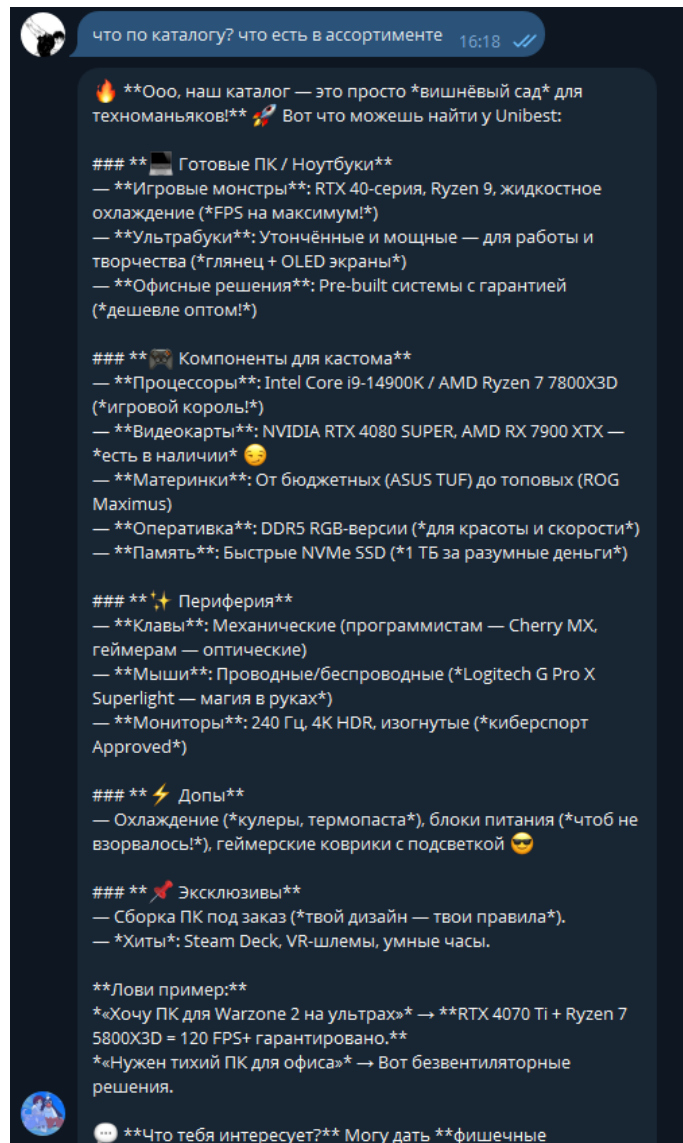


Fig. 2: Example of applying mask based product sharing

#### D. Operational and Knowledge Update Pipelines

Product database changes propagate into both the canonical relational store and into the document ingestion pipeline. Documents are chunked, embedded using Sberbank embeddings for RU/KZ languages, and re-indexed into FAISS/Qdrant on a scheduled cadence or via delta-updates. See Fig.2 for an example of UnieAI using the products database.

Periodically, the system runs the 2GIS parser to enrich B2B data with newly found contacts and inserts them into the vector store and CRM as lead candidates for UNIA routing. See Fig.3 for a data table generated using the 2GIS parser.

| № | Наименование | Рубрик | Адрес | Код | Страна | Почта | Офис | Регистр | Страна | Учредитель | Идентификатор | Рубрик | Т-1 | Т-2 | Т-3 | Т-4 | Т-5 | Т-6 | Т-7 | Т-8 | Т-9 | Т-10 | Т-11 | Т-12 | Т-13 | Т-14 | Т-15 | Т-16 | Т-17 | Т-18 | Т-19 | Т-20 | Т-21 | Т-22 | Т-23 | Т-24 | Т-25 | Т-26 | Т-27 | Т-28 | Т-29 | Т-30 | Т-31 | Т-32 | Т-33 | Т-34 | Т-35 | Т-36 | Т-37 | Т-38 | Т-39 | Т-40 | Т-41 | Т-42 | Т-43 | Т-44 | Т-45 | Т-46 | Т-47 | Т-48 | Т-49 | Т-50 | Т-51 | Т-52 | Т-53 | Т-54 | Т-55 | Т-56 | Т-57 | Т-58 | Т-59 | Т-60 | Т-61 | Т-62 | Т-63 | Т-64 | Т-65 | Т-66 | Т-67 | Т-68 | Т-69 | Т-70 | Т-71 | Т-72 | Т-73 | Т-74 | Т-75 | Т-76 | Т-77 | Т-78 | Т-79 | Т-80 | Т-81 | Т-82 | Т-83 | Т-84 | Т-85 | Т-86 | Т-87 | Т-88 | Т-89 | Т-90 | Т-91 | Т-92 | Т-93 | Т-94 | Т-95 | Т-96 | Т-97 | Т-98 | Т-99 | Т-100 | Т-101 | Т-102 | Т-103 | Т-104 | Т-105 | Т-106 | Т-107 | Т-108 | Т-109 | Т-110 | Т-111 | Т-112 | Т-113 | Т-114 | Т-115 | Т-116 | Т-117 | Т-118 | Т-119 | Т-120 | Т-121 | Т-122 | Т-123 | Т-124 | Т-125 | Т-126 | Т-127 | Т-128 | Т-129 | Т-130 | Т-131 | Т-132 | Т-133 | Т-134 | Т-135 | Т-136 | Т-137 | Т-138 | Т-139 | Т-140 | Т-141 | Т-142 | Т-143 | Т-144 | Т-145 | Т-146 | Т-147 | Т-148 | Т-149 | Т-150 | Т-151 | Т-152 | Т-153 | Т-154 | Т-155 | Т-156 | Т-157 | Т-158 | Т-159 | Т-160 | Т-161 | Т-162 | Т-163 | Т-164 | Т-165 | Т-166 | Т-167 | Т-168 | Т-169 | Т-170 | Т-171 | Т-172 | Т-173 | Т-174 | Т-175 | Т-176 | Т-177 | Т-178 | Т-179 | Т-180 | Т-181 | Т-182 | Т-183 | Т-184 | Т-185 | Т-186 | Т-187 | Т-188 | Т-189 | Т-190 | Т-191 | Т-192 | Т-193 | Т-194 | Т-195 | Т-196 | Т-197 | Т-198 | Т-199 | Т-200 | Т-201 | Т-202 | Т-203 | Т-204 | Т-205 | Т-206 | Т-207 | Т-208 | Т-209 | Т-210 | Т-211 | Т-212 | Т-213 | Т-214 | Т-215 | Т-216 | Т-217 | Т-218 | Т-219 | Т-220 | Т-221 | Т-222 | Т-223 | Т-224 | Т-225 | Т-226 | Т-227 | Т-228 | Т-229 | Т-230 | Т-231 | Т-232 | Т-233 | Т-234 | Т-235 | Т-236 | Т-237 | Т-238 | Т-239 | Т-240 | Т-241 | Т-242 | Т-243 | Т-244 | Т-245 | Т-246 | Т-247 | Т-248 | Т-249 | Т-250 | Т-251 | Т-252 | Т-253 | Т-254 | Т-255 | Т-256 | Т-257 | Т-258 | Т-259 | Т-260 | Т-261 | Т-262 | Т-263 | Т-264 | Т-265 | Т-266 | Т-267 | Т-268 | Т-269 | Т-270 | Т-271 | Т-272 | Т-273 | Т-274 | Т-275 | Т-276 | Т-277 | Т-278 | Т-279 | Т-280 | Т-281 | Т-282 | Т-283 | Т-284 | Т-285 | Т-286 | Т-287 | Т-288 | Т-289 | Т-290 | Т-291 | Т-292 | Т-293 | Т-294 | Т-295 | Т-296 | Т-297 | Т-298 | Т-299 | Т-300 | Т-301 | Т-302 | Т-303 | Т-304 | Т-305 | Т-306 | Т-307 | Т-308 | Т-309 | Т-310 | Т-311 | Т-312 | Т-313 | Т-314 | Т-315 | Т-316 | Т-317 | Т-318 | Т-319 | Т-320 | Т-321 | Т-322 | Т-323 | Т-324 | Т-325 | Т-326 | Т-327 | Т-328 | Т-329 | Т-330 | Т-331 | Т-332 | Т-333 | Т-334 | Т-335 | Т-336 | Т-337 | Т-338 | Т-339 | Т-340 | Т-341 | Т-342 | Т-343 | Т-344 | Т-345 | Т-346 | Т-347 | Т-348 | Т-349 | Т-350 | Т-351 | Т-352 | Т-353 | Т-354 | Т-355 | Т-356 | Т-357 | Т-358 | Т-359 | Т-360 | Т-361 | Т-362 | Т-363 | Т-364 | Т-365 | Т-366 | Т-367 | Т-368 | Т-369 | Т-370 | Т-371 | Т-372 | Т-373 | Т-374 | Т-375 | Т-376 | Т-377 | Т-378 | Т-379 | Т-380 | Т-381 | Т-382 | Т-383 | Т-384 | Т-385 | Т-386 | Т-387 | Т-388 | Т-389 | Т-390 | Т-391 | Т-392 | Т-393 | Т-394 | Т-395 | Т-396 | Т-397 | Т-398 | Т-399 | Т-400 | Т-401 | Т-402 | Т-403 | Т-404 | Т-405 | Т-406 | Т-407 | Т-408 | Т-409 | Т-410 | Т-411 | Т-412 | Т-413 | Т-414 | Т-415 | Т-416 | Т-417 | Т-418 | Т-419 | Т-420 | Т-421 | Т-422 | Т-423 | Т-424 | Т-425 | Т-426 | Т-427 | Т-428 | Т-429 | Т-430 | Т-431 | Т-432 | Т-433 | Т-434 | Т-435 | Т-436 | Т-437 | Т-438 | Т-439 | Т-440 | Т-441 | Т-442 | Т-443 | Т-444 | Т-445 | Т-446 | Т-447 | Т-448 | Т-449 | Т-450 | Т-451 | Т-452 | Т-453 | Т-454 | Т-455 | Т-456 | Т-457 | Т-458 | Т-459 | Т-460 | Т-461 | Т-462 | Т-463 | Т-464 | Т-465 | Т-466 | Т-467 | Т-468 | Т-469 | Т-470 | Т-471 | Т-472 | Т-473 | Т-474 | Т-475 | Т-476 | Т-477 | Т-478 | Т-479 | Т-480 | Т-481 | Т-482 | Т-483 | Т-484 | Т-485 | Т-486 | Т-487 | Т-488 | Т-489 | Т-490 | Т-491 | Т-492 | Т-493 | Т-494 | Т-495 | Т-496 | Т-497 | Т-498 | Т-499 | Т-500 | Т-501 | Т-502 | Т-503 | Т-504 | Т-505 | Т-506 | Т-507 | Т-508 | Т-509 | Т-510 | Т-511 | Т-512 | Т-513 | Т-514 | Т-515 | Т-516 | Т-517 | Т-518 | Т-519 | Т-520 | Т-521 | Т-522 | Т-523 | Т-524 | Т-525 | Т-526 | Т-527 | Т-528 | Т-529 | Т-530 | Т-531 | Т-532 | Т-533 | Т-534 | Т-535 | Т-536 | Т-537 | Т-538 | Т-539 | Т-540 | Т-541 | Т-542 | Т-543 | Т-544 | Т-545 | Т-546 | Т-547 | Т-548 | Т-549 | Т-550 | Т-551 | Т-552 | Т-553 | Т-554 | Т-555 | Т-556 | Т-557 | Т-558 | Т-559 | Т-560 | Т-561 | Т-562 | Т-563 | Т-564 | Т-565 | Т-566 | Т-567 | Т-568 | Т-569 | Т-570 | Т-571 | Т-572 | Т-573 | Т-574 | Т-575 | Т-576 | Т-577 | Т-578 | Т-579 | Т-580 | Т-581 | Т-582 | Т-583 | Т-584 | Т-585 | Т-586 | Т-587 | Т-588 | Т-589 | Т-590 | Т-591 | Т-592 | Т-593 | Т-594 | Т-595 | Т-596 | Т-597 | Т-598 | Т-599 | Т-600 | Т-601 | Т-602 | Т-603 | Т-604 | Т-605 | Т-606 | Т-607 | Т-608 | Т-609 | Т-610 | Т-611 | Т-612 | Т-613 | Т-614 | Т-615 | Т-616 | Т-617 | Т-618 | Т-619 | Т-620 | Т-621 | Т-622 | Т-623 | Т-624 | Т-625 | Т-626 | Т-627 | Т-628 | Т-629 | Т-630 | Т-631 | Т-632 | Т-633 | Т-634 | Т-635 | Т-636 | Т-637 | Т-638 | Т-639 | Т-640 | Т-641 | Т-642 | Т-643 | Т-644 | Т-645 | Т-646 | Т-647 | Т-648 | Т-649 | Т-650 | Т-651 | Т-652 | Т-653 | Т-654 | Т-655 | Т-656 | Т-657 | Т-658 | Т-659 | Т-660 | Т-661 | Т-662 | Т-663 | Т-664 | Т-665 | Т-666 | Т-667 | Т-668 | Т-669 | Т-670 | Т-671 | Т-672 | Т-673 | Т-674 | Т-675 | Т-676 | Т-677 | Т-678 | Т-679 | Т-680 | Т-681 | Т-682 | Т-683 | Т-684 | Т-685 | Т-686 | Т-687 | Т-688 | Т-689 | Т-690 | Т-691 | Т-692 | Т-693 | Т-694 | Т-695 | Т-696 | Т-697 | Т-698 | Т-699 | Т-700 | Т-701 | Т-702 | Т-703 | Т-704 | Т-705 | Т-706 | Т-707 | Т-708 | Т-709 | Т-710 | Т-711 | Т-712 | Т-713 | Т-714 | Т-715 | Т-716 | Т-717 | Т-718 | Т-719 | Т-720 | Т-721 | Т-722 | Т-723 | Т-724 | Т-725 | Т-726 | Т-727 | Т-728 | Т-729 | Т-730 | Т-731 | Т-732 | Т-733 | Т-734 | Т-735 | Т-736 | Т-737 | Т-738 | Т-739 | Т-740 | Т-741 | Т-742 | Т-743 | Т-744 | Т-745 | Т-746 | Т-747 | Т-748 | Т-749 | Т-750 | Т-751 | Т-752 | Т-753 | Т-754 | Т-755 | Т-756 | Т-757 | Т-758 | Т-759 | Т-760 | Т-761 | Т-762 | Т-763 | Т-764 | Т-765 | Т-766 | Т-767 | Т-768 | Т-769 | Т-770 | Т-771 | Т-772 | Т-773 | Т-774 | Т-775 | Т-776 | Т-777 | Т-778 | Т-779 | Т-780 | Т-781 | Т-782 | Т-783 | Т-784 | Т-785 | Т-786 | Т-787 | Т-788 | Т-789 | Т-790 | Т-791 | Т-792 | Т-793 | Т-794 | Т-795 | Т-796 | Т-797 | Т-798 | Т-799 | Т-800 | Т-801 | Т-802 | Т-803 | Т-804 | Т-805 | Т-806 | Т-807 | Т-808 | Т-809 | Т-810 | Т-811 | Т-812 | Т-813 | Т-814 | Т-815 | Т-816 | Т-817 | Т-818 | Т-819 | Т-820 | Т-821 | Т-822 | Т-823 | Т-824 | Т-825 | Т-826 | Т-827 | Т-828 | Т-829 | Т-830 | Т-831 | Т-832 | Т-833 | Т-834 | Т-835 | Т-836 | Т-837 | Т-838 | Т-839 | Т-840 | Т-841 | Т-842 | Т-843 | Т-844 | Т-845 | Т-846 | Т-847 | Т-848 | Т-849 | Т-850 | Т-851 | Т-852 | Т-853 | Т-854 | Т-855 | Т-856 | Т-857 | Т-858 | Т-859 | Т-860 | Т-861 | Т-862 | Т-863 | Т-864 | Т-865 | Т-866 | Т-867 | Т-868 | Т-869 | Т-870 | Т-871 | Т-872 | Т-873 | Т-874 | Т-875 | Т-876 | Т-877 | Т-878 | Т-879 | Т-880 | Т-881 | Т-882 | Т-883 | Т-884 | Т-885 | Т-886 | Т-887 | Т-888 | Т-889 | Т-890 | Т-891 | Т-892 | Т-893 | Т-894 | Т-895 | Т-896 | Т-897 | Т-898 | Т-899 | Т-900 | Т-901 | Т-902 | Т-903 | Т-904 | Т-905 | Т-906 | Т-907 | Т-908 | Т-909 | Т-910 | Т-911 | Т-912 | Т-913 | Т-914 | Т-915 | Т-916 | Т-917 | Т-918 | Т-919 | Т-920 | Т-921 | Т-922 | Т-923 | Т-924 | Т-925 | Т-926 | Т-927 | Т-928 | Т-929 | Т-930 | Т-931 | Т-932 | Т-933 | Т-934 | Т-935 | Т-936 | Т-937 | Т-938 | Т-939 | Т-940 | Т-941 | Т-942 | Т-943 | Т-944 | Т-945 | Т-946 | Т-947 | Т-948 | Т-949 | Т-950 | Т-951 | Т-952 | Т-953 | Т-954 | Т-955 | Т-956 | Т-957 | Т-958 | Т-959 | Т-960 | Т-961 | Т-962 | Т-963 | Т-964 | Т-965 | Т-966 | Т-967 | Т-968 | Т-969 | Т-970 | Т-971 | Т-972 | Т-973 | Т-974 | Т-975 | Т-976 | Т-977 | Т-978 | Т-979 | Т-980 | Т-981 | Т-982 | Т-983 | Т-984 | Т-985 | Т-986 | Т-987 | Т-988 | Т-989 | Т-990 | Т-991 | Т-992 | Т-993 | Т-994 | Т-995 | Т-996 | Т-997 | Т-998 | Т-999 | Т-1000 | Т-1001 | Т-1002 | Т-1003 | Т-1004 | Т-1005 | Т-1006 | Т-1007 | Т-1008 | Т-1009 | Т-1010 | Т-1011 | Т-1012 | Т-1013 | Т-1014 | Т-1015 | Т-1016 | Т-1017 | Т-1018 | Т-1019 | Т-1020 | Т-1021 | Т-1022 | Т-1023 | Т-1024 | Т-1025 | Т-1026 | Т-1027 | Т-1028 | Т-1029 | Т-1030 | Т-1031 | Т-1032 | Т-1033 | Т-1034 | Т-1035 | Т-1036 | Т-1037 | Т-1038 | Т-1039 | Т-1040 | Т-1041 | Т-1042 | Т-1043 | Т-1044 | Т-1045 | Т-1046 | Т-1047 | Т-1048 | Т-1049 | Т-1050 | Т-1051 | Т-1052 | Т-1053 | Т-1054 | Т-1055 | Т-1056 | Т-1057 | Т-1058 | Т-1059 | Т-1060 | Т-1061 | Т-1062 | Т-1063 | Т-1064 | Т-1065 | Т-1066 | Т-1067 | Т-1068 | Т-1069 | Т-1070 | Т-1071 | Т-1072 | Т-1073 | Т-1074 | Т-1075 | Т-1076 | Т-1077 | Т-1078 | Т-1079 | Т-1080 | Т-1081 | Т-1082 | Т-1083 | Т-1084 | Т-1085 | Т-1086 | Т-1087 | Т-1088 | Т-1089 | Т-1090 | Т-1091 | Т-1092 | Т-1093 | Т-1094 | Т-1095 | Т-1096 | Т-1097 | Т-1098 | Т-1099 | Т-1100 | Т-1101 | Т-1102 | Т-1103 | Т-1104 | Т-1105 | Т-1106 | Т-1107 | Т-1108 | Т-1109 | Т-1110 | Т-1111 | Т-1112 | Т-1113 | Т-1114 | Т-1115 | Т-1116 | Т-1117 | Т-1118 | Т-1119 | Т-1120 | Т-1121 | Т-1122 | Т-1123 | Т-1124 | Т-1125 | Т-1126 | Т-1127 | Т-1128 | Т-1129 | Т-1130 | Т-1131 | Т-1132 | Т-1133 | Т-1134 | Т-1135 | Т-1136 | Т-1137 | Т-1138 | Т-1139 | Т-1140 | Т-1141 | Т-1142 | Т-1143 | Т-1144 | Т-1145 | Т-1146 | Т-1147 | Т-1148 | Т-1149 | Т-1150 | Т-1151 | Т-1152 | Т-1153 | Т-1154 | Т-1155 | Т-1156 | Т-1157 | Т-1158 | Т-1159 | Т-1160 | Т-1161 | Т-1162 | Т-1163 | Т-1164 | Т-1165 | Т-1166 | Т-1167 | Т-1168 | Т-1169 | Т-1170 | Т-1171 | Т-1172 | Т-1173 | Т-1174 | Т-1175 | Т-1176 | Т-1177 | Т-1178 | Т-1179 | Т-1180 | Т-1181 | Т-1182 | Т-1183 | Т-1184 | Т-1185 | Т-1186 | Т-1187 | Т-1188 | Т-1189 | Т-1190 | Т-1191 | Т-1192 | Т-1193 | Т-1194 | Т-1195 | Т-1196 | Т-1197 | Т-1198 | Т-1199 | Т-1200 | Т-1201 | Т-1202 | Т-1203 | Т-1204 | Т-1205 | Т-1206 | Т-1207 | Т-1208 | Т-1209 | Т-1210 | Т-1211 | Т-1212 | Т-1213 | Т-1214 | Т-1215 | Т-1216 | Т-121 |
|---|--------------|--------|-------|-----|--------|-------|------|---------|--------|------------|---------------|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|
|---|--------------|--------|-------|-----|--------|-------|------|---------|--------|------------|---------------|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|

Fig. 3: Example of product table based on "comp clubs" search in Almaty

UnieAI prompt templates and small fine-tuned adapters are versioned, meaning that UNIA can push prompt updates to running copies and trigger cache invalidation of previous contexts. See Fig. 4 for the overall architecture illustration.

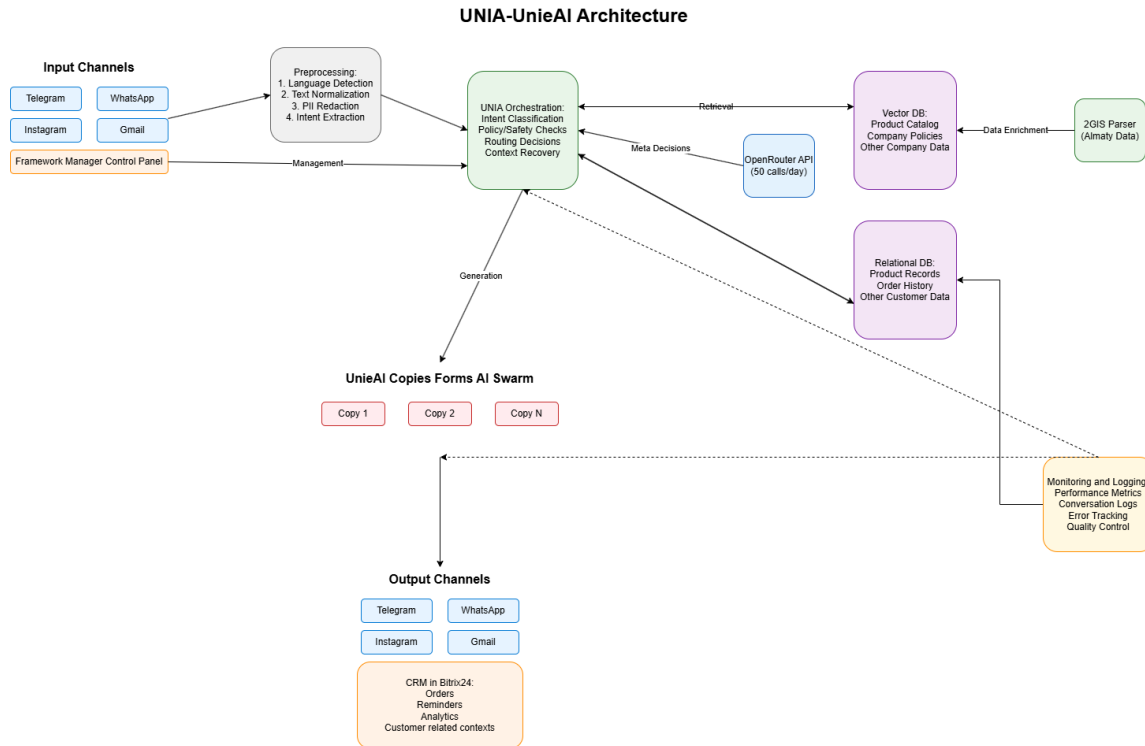


Fig. 4: Algorithm

### E. Operational Pipeline — Client Input

The end-to-end pipeline for a customer message is as follows:

- 1) Customer sends a message via Telegram/WhatsApp/Instagram/Gmail. The channel adapter normalizes incoming payloads and forwards the message to the preprocessing layer.
- 2) Language detection, text normalization, adapting markups and formatting and intent-slot extraction via lightweight embedding and translating models using masking.

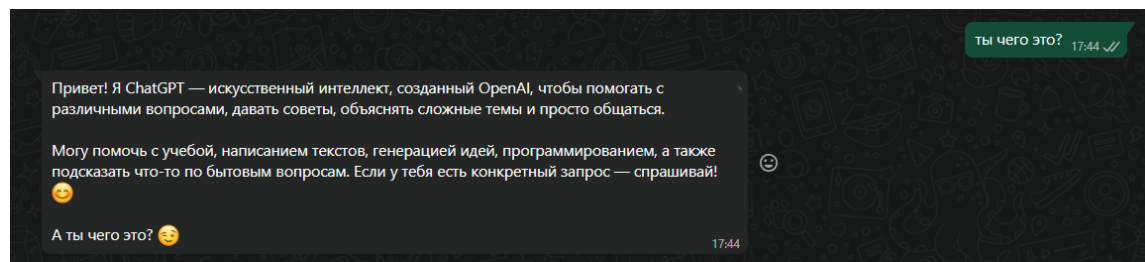


Fig. 5: Example of removing boldness of the messages. Whatsapp's boldness works differently from Telegram's one, meaning it was removed.

- 3) UNIA receives the preprocessed message and performs: policy, quality checks and safety checks by detecting attempts to distract or inject malicious context, routing UnieAI copies. UNIA also maintains conversation level context; in such cases it restores canonical context from stored transcripts.
- 4) If the UnieAI's response requires highly contextual knowledge like product info, warranty, pricing, UNIA issues a vector retrieval and fetches the required product information.
- 5) Augmentation can also be managed by UNIA by generating prompt constraints and combining user message, retrieved passages, company protocols using response templates. If translation is needed between Kazakh and Russian, the pipeline inserts translating layer.
- 6) The prompt is sent to the local quantized Nous Hermes 2 Mistral 7B instance used by UnieAI copies for final generation. For high risk decisions UNIA uses Deepseek via OpenRouter API.
- 7) After generated text validation UNIA may also correct critical language errors or append CRM tags and make a pre-order if the context forms so. The order tasks are handled by Bitrix24 for human employees. See Fig. 7 for example of verified orders.

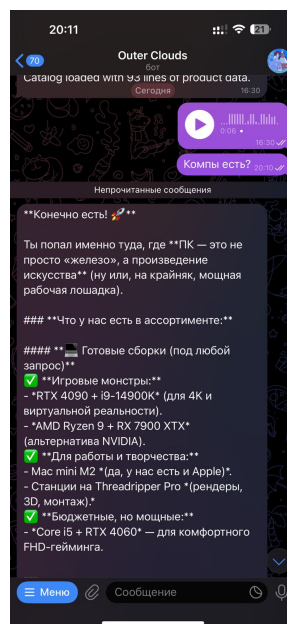


Fig. 6: Example of interacting with Unie send by one of Unibest's customers.

- 8) Then response is sent back through the channel to the customer and the full exchange is logged. See Fig. 6 for final UnieAI's response
- 9) If flow requires reminders, UNIA schedules timed reminders and associates them with the CRM record.

Fig. 7: UNIA makes orders and closes them automatically after verification. MoySklad and Kaspi are used.

#### F. Operational Pipeline — Manager Input

The pipeline for inputs from sales managers (campaign creation, overrides, audits):

- 1) Using framework console, UNIA's management console or Bitrix24 if the employee has higher positions operator creates a campaign spawning UnieAI copies into the "AI swarm". See Fig. 8 for UNIA's tasks for verification by manager in UNIA's management console

Fig. 8: UNIA creating tasks for delivery from the chats' data. Managers needs to verify it.

- 2) UNIA compiles specified by manager or company policies into rules, prompt templates and safety constraints.
- 3) UNIA instantiates specially configured UnieAI copies and assigns each copy a knowledge context. Copies are tracked and LLM resource usage is monitored. See Fig. 9 for UnieAI copies answering for different users in a group chat.

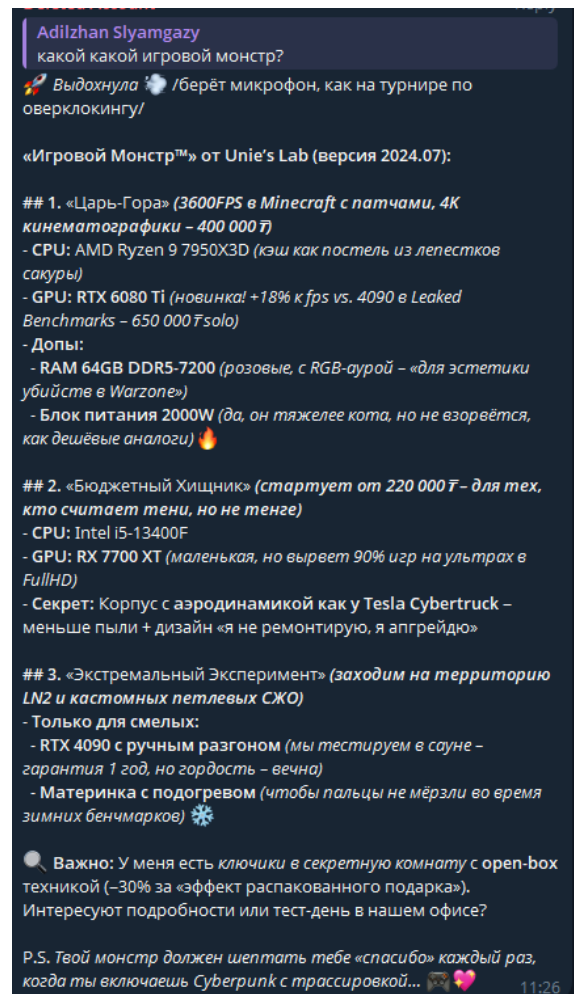


Fig. 9: Example of Unie answering to multiple users via of its own copies

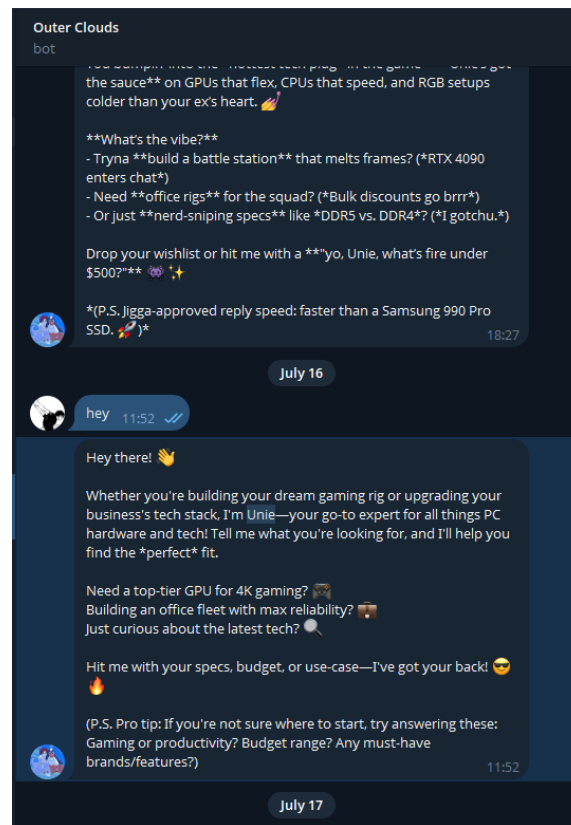


Fig. 10: Copy of Unie in Telegram platform

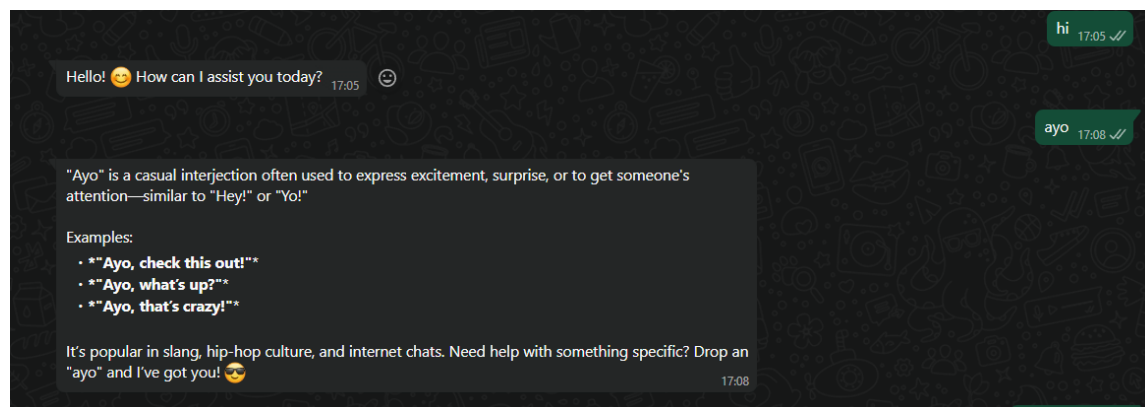


Fig. 11: Copy of Unie in Whatsapp platform

- 4) Managers can sample conversations, submit corrections and start prompt updates via UNIA. UNIA accepts feedback and applies template updates to selected copies using framework configuration. See Fig. 10, Fig. 11, Fig. 12 for more chat

functionality.

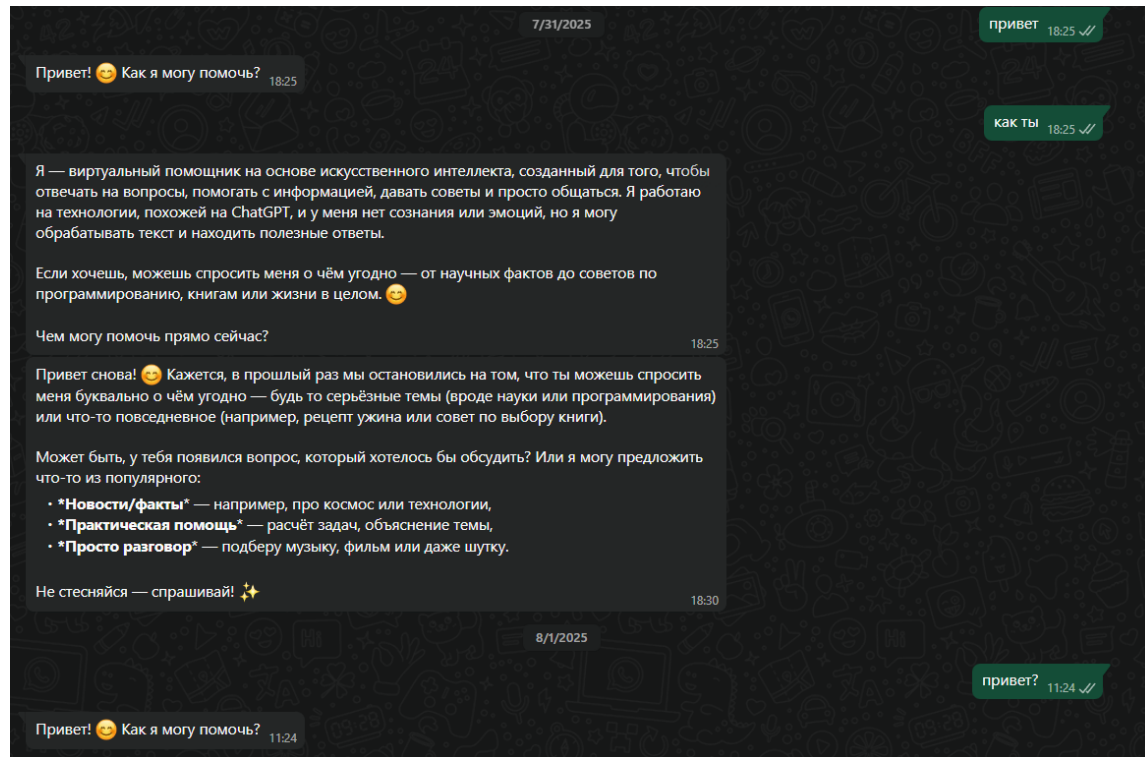


Fig. 12: Example of resending message to user after waiting 5 minutes

- 5) UNIA notifies assigned sales personnel via Bitrix or internal framework messaging when detects low-confidence responses or policy breaches.

### G. RAG and Embedding Flow

All textual knowledge in the knowledge base is tokenized and embedded using Sberbank's RU/KZ embedding model. For low-latency approximate nearest neighbor search, embeddings are stored in FAISS and in Qdrant for persistence and metadata filtering. Retrieval results include provenance metadata (session id, chunk offset, timestamp) to allow UNIA and sales managers to trace the origin of information.

### H. Error Handling, Fallbacks and Quota Management

Due to the OpenRouter API's 50 free calls per day budget, UNIA reserves these calls for non-local high risk decisions and uses UnieAI for the majority of generation tasks. Quota exhaustion logic falls back to local rule-based routing and preserved policies. Low-confidence outputs, policy violations or even hallucinations start prompt re-formulation and warn the sales manager which also ensures that incidents can be solved in short terms. For audition purposes the framework can render conversations via retrieval contexts generated by UNIA.

### I. Data Collection and Analysis Methods

Data collection was performed through systematic monitoring of system performance across multiple experimental runs. Key metrics included task completion times, resource utilization rates, system throughput, and error rates.

Received: December 20, 2025. Reviewed: September 26, 2026. Accepted: September 29, 2026. © 2026 Isakulov Zeiin. All rights reserved.

Additional metrics specific to the Unibest deployment were logged. Among them is the number of conversations per channel, average response latency (processing, UNIA layer, inference), conversion rate (from leading to making an order as the result), B2C review reminder and response rates, and the rate of UNIA-triggered corrections to UnieAI outputs. User-centered metrics such as answer relevance, customer satisfaction (short surveys) and required follow-up message counts were also collected.

All interactions, retrieved chat files and final assistant outputs were recorded using JSON files for offline fixes.

Data analysis focused on comparing performance between the AI assistants API approach and the proposed framework. The measured marginal effect of adding UNIA and UnieAI impacted business KPIs. Performance improvements were quantified using relative percentage differences.

For UNIA, analysis included classification and routing accuracy testing (precision and recall), compared to end-to-end latency. The changes in the assistants' output quality and prompt safety were also detected. Comparative prompt/response testing and inference techniques were used to enable better impact on task success and user satisfaction.

## IV. RESULTS AND DISCUSSION

### A. Experimental Findings

The experimental deployment of the framework with UNIA orchestration demonstrates significant improvements over baseline high-end local models. Inference products showed that the orchestrated system achieved almost 15% higher ordering conversion efficiency for B2B inquiries. In B2C operations, the automated review reminder achieved a customer engagement rate of almost 34%, defined as the proportion of customers who interacted with the reminder. This type of system also reduces the human resources required by one or two managers.

Architecture meets real-time interaction requirements by achieving average response latencies below 2.5 seconds for processing. UNIA's intervention in post-processing corrected almost 85% of initial UnieAI outputs. UNIA primarily ensures company policy during chatting, which substantially improved response relevance scores. See Table 5 for detailed performance comparison.

TABLE V: Performance Comparison of Framework Components

| Metric                   | Baseline | With UNIA |
|--------------------------|----------|-----------|
| B2B Conversion Rate      | 68%      | 83%       |
| Average Response Latency | 3.8s     | 2.3s      |
| Output Correction Rate   | N/A      | 85%       |
| Customer Engagement      | 30%      | 64%       |

### B. Comparative Analysis with Existing Approaches

Previous studies like on RepairLLaMA demonstrated efficiency of using specialized tools. Unlike many reviewed case studies that utilize high-end models for all tasks or highly specialized models only for one task, our results demonstrate that combination of lightweight local model for text generation and a more capable free demo model as an "conciseness adapter" is both effective and performant approach for multi-lingual highly contextual small and medium enterprises.

The performance improvements resulted from UNIA orchestration functionality. Comparing responses with and without UNIA, allows us to conclude that the increase in conversion rates directly caused by UNIA's routing and management. The efficient use of the limited OpenRouter budget is a causal result of the design decision to reserve it for high risk decisions.

### C. Limitations and Challenges

The main limitation is caused by the fact that evaluation was conducted primarily in a pilot deployment. The generalizability of the system to other contexts and enterprise remains unproven.

We speculate the performance gap between the orchestrated and non-orchestrated assistants would widen further if UNIA's context recovery mechanism will cause contextual collapse after repetitive re-generation of the company's context. That's why it is important to periodically check system quality. We also speculate that the embedding and retrieval solutions, while sufficient for Russian and Kazakh languages, could be also optimized by using future multilingual models that can handle several languages better on architectural levels.

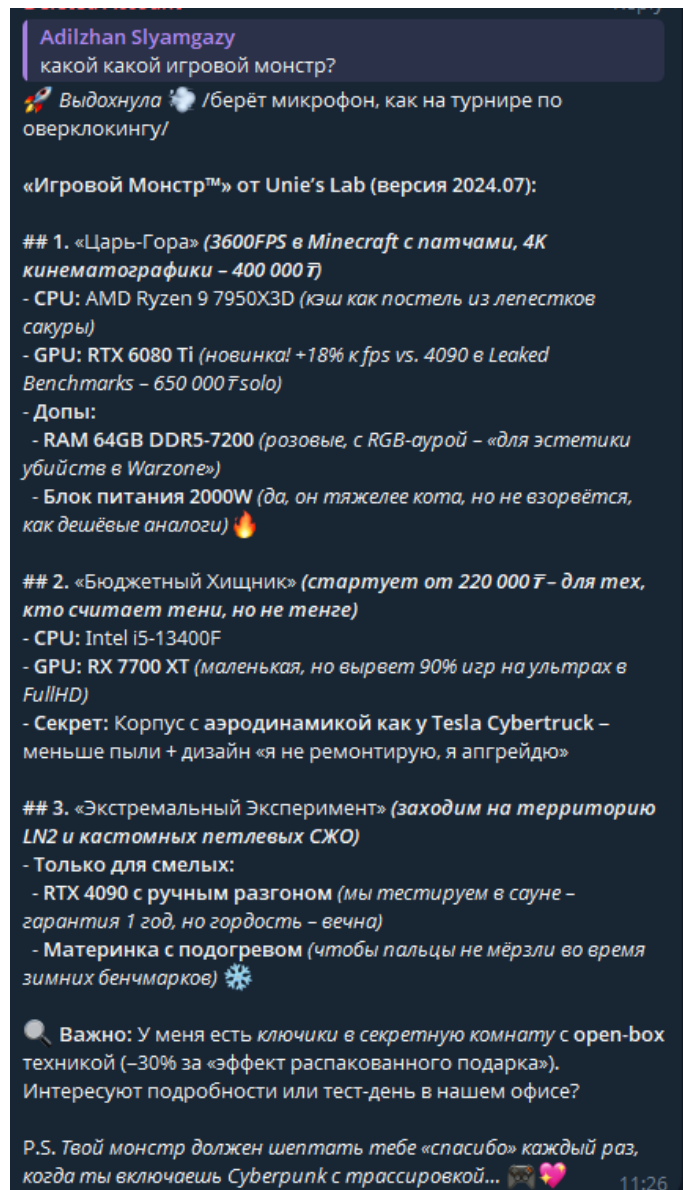


Fig. 13: Example of Unie answering to multiple users via of its own copies

Any AI assistant architecture including UnieAI, requires oversights and quality management otherwise it may cause high financial risks. UNIA is suitable to provide only parts of these oversights but can make decisions better. Therefore, implementing UNIA as a high-end orchestrating model and UnieAI as a specialized "AI Swarm" should lead to improvements in efficiency and quality for businesses with limited technological resources. The experimental results demonstrate improved accuracy and reduced errors, which validates the core hypothesis that optimizing models and dedicating decision making is possible and further standardized frameworks evaluate fitting AI-based sales automation in enterprise environments.

## V. CONCLUSION AND FUTURE WORK

This article explored several studies at the intersection of LLM with practical cases, focusing on their particular applications in small and medium sized enterprises with applied optimization methods. The study successfully designed and implemented the framework at key decision-making levels while preserving UnieAI's core purpose—text generation. The framework made local models more agile for sales automation by orchestrating UnieAI's copies on UNIA and by making UNIA a “conciseness adapter” for those copies. The results demonstrate the framework pipeline can significantly enhance business operations such as lead conversion, response quality and customer engagement while operating under significant computational constraints. The strategic partitioning of corporate tasks between a quantized local model and a budgeted and more powerful API model proves to be an efficient architecture for real-world deployment via a collection of AI assistants into a managed system, so called “AI swarm”. The project confirms that the marginal utility of adding AI assistants at such highly contextual layer is substantial.

The primary limitation of this work is its application in a specific city, where the key functionality providing quality were already implemented. Its dependence on external 2GIS and Bitrix24 services makes the framework applicable and easy to deploy on the huge part of the market but lacks independence if the services become unavailable or exceed budget limitations. The performance trade offs of using heavily quantized local models causes the quality issues. The assumptions regarding embedding quality and budget sufficiency validated in our example, but may not be applicable under different conditions.

Future work will mainly focus on expanding the deployment to other cities to generalize the framework across different markets. Additionally, it is important to deepen the investigation into the AI-based workflows to optimize the handoff process between UNIA and sales managers to reduce further operational overhead and the economic value of each UNIA function to make it more commercially interesting.

## VI. ACKNOWLEDGMENT

The author thanks the research participants and technical teams who contributed to this study. This research was supervised by the Department of Information Systems at SDU University. The author expresses their gratitude industry partners who provided access to the experimental infrastructure and case study environment.

## REFERENCES

- [1] A. Silva, S. Fang, and M. Monperrus, “RepairLLaMA: Efficient representations and fine-tuned adapters for program repair,” *arXiv preprint arXiv:2312.15698*, 2023. [Online]. Available: <https://arxiv.org/abs/2312.15698>
- [2] U. J. Umoga et al., “Exploring the potential of AI-driven optimization in enhancing network performance and efficiency,” *Magna Scientia Advanced Research and Reviews*, vol. 10, no. 1, pp. 368–378, 2024. [Online]. Available: <https://doi.org/10.30574/msarr.2024.10.1.0028>
- [3] B. Crawford et al., “Software project scheduling using the hyper-cube ant colony optimization algorithm,” *J. Comput. Inf. Technol.*, 2015. [Online]. Available: <https://hrca.hr/file/216665>
- [4] C. E. Love and H. F. Jordan, “An investigation of static versus dynamic scheduling,” University of Colorado at Boulder, Boulder, CO, USA, Tech. Rep., 1990. [Online]. Available: <https://doi.org/10.1109/ISCA.1990.134525>
- [5] A. Aldoseri, K. N. Al-Khalifa, and A. M. Hamouda, “A roadmap for integrating automation with process optimization for AI-powered digital transformation,” *Preprints.org*, 2023, doi: 10.20944/preprints202310.1055.v1. [Online]. Available: <https://doi.org/10.20944/preprints202310.1055.v1>
- [6] X. Hou, Y. Zhao, S. Wang, and H. Wang, “Model Context Protocol (MCP): Landscape, Security Threats, and Future Research Directions,” *arXiv preprint arXiv:2503.23278*, 2025. [Online]. Available: <https://arxiv.org/pdf/2503.23278.pdf>
- [7] M. Firat and S. Kuleli, “What if GPT4 Became Autonomous: The Auto-GPT Project and Use Cases,” *Journal of Emerging Computational Technologies*, vol. 3, no. 1, pp. 1–6, 2023. [Online]. Available: <https://doi.org/10.57020/ject.1297961>
- [8] H. Yang, S. Yue, and Y. He, “Auto-GPT for Online Decision Making: Benchmarks and Additional Opinions,” *arXiv preprint arXiv:2306.02224*, 2023. [Online]. Available: <https://arxiv.org/pdf/2306.02224.pdf>
- [9] A. J. Yepes, Y. You, J. Milczek, S. Laverde, and R. Li, “Financial Report Chunking for Effective Retrieval Augmented Generation,” *arXiv preprint arXiv:2402.05131*, 2024. [Online]. Available: <https://arxiv.org/pdf/2402.05131.pdf>
- [10] X. Wang, Z. Wang, X. Gao, F. Zhang, Y. Wu, Z. Xu, T. Shi, Z. Wang, S. Li, Q. Qian, R. Yin, C. Lv, X. Zheng, and X. Huang, “Searching for Best Practices in Retrieval-Augmented Generation,” in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Miami, Florida, USA, 2024, pp. 17716–17736, doi: 10.18653/v1/2024.emnlp-main.981. [Online]. Available: <https://aclanthology.org/2024.emnlp-main.981/>

- [11] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” *arXiv preprint arXiv:2005.11401*, 2021. [Online]. Available: <https://arxiv.org/pdf/2005.11401.pdf>
- [12] K. Grace, J. Salvatier, A. Dafoe, B. Zhang, and O. Evans, “Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts,” *Journal of Artificial Intelligence Research*, 2018. [Online]. Available: <https://doi.org/10.1613/jair.1.11222>
- [13] B. G. Buchanan and R. G. Smith, “Fundamentals of Expert Systems,” 1988, *Annual Review of Computer Science*, vol. 3, no. 1, pp. 23–58. [Online]. Available: <https://faculty.ist.psu.edu/vhonavar/Courses/ai100/expsysreview.pdf>
- [14] N. S. Moosavi, Q. Delfosse, K. Kersting, and I. Gurevych, “Adaptable Adapters,” *arXiv preprint arXiv:2205.01549*, 2022. [Online]. Available: <https://arxiv.org/pdf/2205.01549.pdf>
- [15] G. Deshmukh, S. De, C. Sehgal, J. S. Gupta, and S. Mittal, “Dressing the Imagination: A Dataset for AI-Powered Translation of Text into Fashion Outfits and A Novel KAN Adapter for Enhanced Feature Adaptation,” *arXiv preprint arXiv:2411.13901*, 2024. [Online]. Available: <https://arxiv.org/pdf/2411.13901.pdf>
- [16] O. Ram, Y. Levine, I. Dalmedigos, D. Muhlgay, A. Shashua, K. Leyton-Brown, and Y. Shoham, “In-Context Retrieval-Augmented Language Models,” *Transactions of the Association for Computational Linguistics*, vol. 11, pp. 1316–1331, 2023. [Online]. Available: [https://doi.org/10.1162/tacl\\_a\\_00605](https://doi.org/10.1162/tacl_a_00605)
- [17] F. M. Zanzotto, “Human-in-the-loop Artificial Intelligence,” *Journal of Artificial Intelligence Research*, vol. 64, 2019. [Online]. Available: <https://doi.org/10.1613/jair.1.11345>
- [18] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, “Towards a Better Understanding of Context and Context-Awareness,” in *Handheld and Ubiquitous Computing*, pp. 304–307, 1991. [Online]. Available: [https://link.springer.com/chapter/10.1007/3-540-48157-5\\_29](https://link.springer.com/chapter/10.1007/3-540-48157-5_29)
- [19] I. Shumailov, Z. Shumaylov, Y. Zhao, N. Papernot, R. Anderson, and Y. Gal, “AI models collapse when trained on recursively generated data,” *Nature*, vol. 631, pp. 755–759, 2024. [Online]. Available: <https://doi.org/10.1038/s41586-024-07566-y>
- [20] E. Peretz-Andersson, S. Tabares, P. Mikalef, and V. Parida, “Artificial intelligence implementation in manufacturing SMEs: A resource orchestration approach,” *International Journal of Information Management*, vol. 77, 102781, 2024. [Online]. Available: <https://doi.org/10.1016/j.ijinfomgt.2024.102781>